EXHIBIT 1

09:45:11 IN THE UNITED STATES DISTRICT COURT
IN AND FOR THE DISTRICT OF DELAWARE

- - -

XEROX CORPORATION, : CIVIL ACTION
:
Plaintiff, :
:
v. :
:
GOOGLE, INC., YAHOO! INC., RIGHT :
MEDIA INC., RIGHT MEDIA LLC, :
YOUTUBE, INC., YOUTUBE, LLC, :
: NO. 10-136 (LPS)
Defendants.

- - -

Wilmington, Delaware
Monday, February 28, 2011
ORAL ARGUMENT HEARING

- - -

BEFORE: HONORABLE **LEONARD P. STARK**, U.S.D.C.J.

APPEARANCES:

- - -

ASHBY & GEDDES, P.A.
BY: JOHN G. DAY, ESQ.

and

CRAVATH SWAINE & MOORE, LLP
BY: RICHARD J. STARK, ESQ., and
SCOTT A. LESLIE, ESQ., and
(New York, New York)

and

XEROX CORPORATION
BY: MICHELLE WAITES, ESQ.
(Norfolk, Connecticut)

Counsel for Xerox Corporation

Brian P. Gaffigan
Official Court Reporter

APPEARANCES: (Continued)

POTTER ANDERSON & CORROON, LLP
BY: DAVID E. MOORE, ESQ.

and

QUINN EMANUEL URQUHART OLIVER & HEDGES, LLP
BY: ANDREA PALLIOS ROBERTS, ESQ.
(Redwood Shores, California)

Counsel for Google, Inc., YouTube, Inc., and YouTube, LLC

MORRIS NICHOLS ARSHT & TUNNELL, LLP
BY: JACK B. BLUMENFELD, ESQ.

and

DAVIS POLK & WARDWELL, LLP
BY: ANTHONY I. FENWICK, ESQ.
(Menlo Park, California)

Counsel for Yahoo! Inc. and Right Media, LLC

- oOo -

P R O C E E D I N G S

(REPORTER'S NOTE: Oral argument hearing took place in open court, starting at 3:33 p.m.)

THE COURT: Good afternoon.

(The attorneys respond, "good afternoon, your Honor.")

THE COURT: Let's begin by having you all put your appearances on the record, please.

MR. DAY: Good afternoon, your Honor.

THE COURT: Good afternoon.

MR. DAY: John Day from Ashby & Geddes, Delaware counsel for Xerox Corporation. With me at counsel table, Richard Stark and Scott Leslie from Xerox's lead counsel, Cravath Swain & Moore in New York; and with us in the gallery is Michelle Waites who is Senior Patent Counsel at Xerox.

THE COURT: Welcome to all of you.

MR. STARK: Good afternoon, your Honor.

MR. BLUMENFELD: Good afternoon, your Honor.

THE COURT: Good afternoon.

MR. BLUMENFELD: Jack Blumenfeld from Morris Nichols with Yahoo! and Right Media. With me today is Tony Fenwick of Davis Polk; and with the Court's permission, Mr. Fenwick will be speaking for the defendants today.

THE COURT: That's fine.

MR. FENWICK: Good afternoon, your Honor.

THE COURT: Good afternoon.

MR. MOORE: Good afternoon, your Honor. David Moore from Potter Anderson on behalf of YouTube. With me is Andrea Roberts from Quinn Emanuel.

THE COURT: Welcome to you as well.

So we're here for argument on the defendants motion for a stay. I will hear first from the defendants. I have in mind roughly about 15 minutes per side.

Hopefully, we can get it done in that amount of time.

We'll proceed to hear first from Mr. Fenwick.

MR. FENWICK: Thank you, your Honor. I'll try to not repeat too much of the material that was in our briefs.

I think these motions for stay are often kind of put in a few general buckets. I think it's fair to put this one in the bucket of early in the litigation, early in the reexamination, noncompetitor case. And I submit that most of the arguments and perhaps all the arguments that have been made against the stay by the plaintiff have really been arguments of sort of general applicability.

I'd like to start with a couple of arguments of general applicability about cases in that sort of category.

The first is that we know that the PTO is not going to suspend the reexamination in this case at this stage. They had a statutory requirement to pursue the reexamination with special dispatch, so they're not going to suspend the reexamination. And that means that at a time when courts are short of judges, courts and the PTO are overburdened we read in the paper, the only way to avoid a lot of wasteful duplicative effort is for this court to stay the case. Otherwise, we will certainly have, for some period of time, two tracks of effort of folks looking at the same set of issues, which are, of course,

complex issues -- a patent that is something like 77 columns long, not necessarily simple stuff.

Second, Xerox asserts that given their time line that they think things will play out on, that the reexamination is essentially moot because we will get to a final decision here before we will get to a final decision in the reexamination, if there is no stay.

I would submit, your Honor, that if that is right, if they're right about that, then a stay is the only chance that my clients and the codefendants in the case have for these patent claims to be considered on a preponderance of the evidence standard; in other words, the validity of these claims considered on a preponderance of the evidence with respect to the art that is at play in front of the reexam, which the PTO has already said raises substantial new questions of patentability.

Otherwise, the only decision that is going to matter, according to their time line, is the decision of this Court in which we're facing a presumption against invalidity with respect to art that has never been considered.

It's also the case that this patent was issued after the Supreme Court *KSR* decision, so these claims have never been evaluated with respect to obviousness under the standard that is set forth by *KSR*, never been considered by the PTO. This is the one chance for the PTO to do that.

THE COURT: But in this case, while they're both at an early stage, the litigation was filed before the reexam; correct?

MR. FENWICK: That's correct.

THE COURT: So it's an interesting dilemma you pose because if you got there first, if you do get to the PTO first, you could potentially have the patent invalidated under a lesser burden than you would face in this court, but the patentee sued you first. Typically, the patentee gets to choose its forum and so maybe that is just a right that isn't fully effectuated for you in this case.

MR. FENWICK: Your Honor, that is a nice segue for me into some of the particulars in this case.

The plaintiff is a very big company. The licensing of the particular patent that is left in this case, the '979 patent, is a minuscule part of their business. So from a prejudice standpoint, it's hard to imagine a case in this general bucket where there would be less prejudice to the plaintiff.

You have referenced the plaintiff's right to bring a suit in district court. They have that right, just as the defendants have the right to file a reexamination petition. But that right comes with responsibilities, I would submit.

I think if we look at the particular way in

which this case has been pursued, we're in a better position to sort of discount that right at this point because the case was filed with two patents, one of which has already been withdrawn, a patent that should never have been asserted in the first case, a patent which we would submit was pretty clearly asserted for tactical reasons relating to venue because we've got two defendants that are headquartered in the Northern District of California.

The agent who was hired tried to license these patents for the plaintiff is headquartered in the Northern District of California. The only U.S. resident inventor in the case, named inventor on the '979 patent that is in the case is in the Bay area in California. The other one is in France.

And so on the '914 patent, which, again, we think should never have been a part of the case in the first place, a number of the inventors are in the Pennsylvania area, sort of local to the Court.

It's also noteworthy that the plaintiff amended the complaint to add a Yahoo! entity that is a East Coast entity. So I think a lot of the basic structure in the case at the outset was aimed at trying to secure venue here, again, with a patent that should never have been asserted.

So I would say the choice of forum issue can be discounted in this case for those reasons.

I think the decision to include that one patent which should never have been asserted casts some doubt or some shadow on the plaintiff's exercise of its rights here. In addition, your Honor, the way that the case has been pursued since its inception, not just the fact that we had to run around and produce a bunch of documents and do a bunch of analysis on a patent that is no longer in the case and should never have been in the case, but I think your Honor has now had a chance to peek under the hood a little bit and see the extent to which the plaintiff has actually progressed the case on the patent that remains, the extent to which the development or disclosure of infringement contentions has been either delayed or not pursued or held back, and finding ourselves where we are at this stage of the case, I would submit to your Honor that there is not a lot of deference owed to the plaintiff for having chosen this forum and having pursued its rights here.

So I'd also point out, your Honor, that much has been made of the number of years that it's purportedly going to take to complete the reexam. Your Honor is always free to reconsider the stay, but at some point it appears that it doesn't make sense. We can expect the petition to be in the central reexam unit of the PTO for two years, perhaps three years from the filing of the petition several months ago.

If what comes out of that as we head to the

appeal process at the PTO is a decision that is against the defendants, then your Honor will may very well want to take a look at that, at the stay at that point and say it looks like this is going the other way. Let's carry on with the case.

If, on the other hand, the central reexamination unit says these claims are invalid, then I don't think your Honor is going to feel too bad about having stayed the case while Xerox pursues its appellate course out of the PTO.

So what I would suggest, your Honor, is that in the bucket of early litigation, early reexamination motions for stay, if your Honor were to deny the stay in this case, I think the upshot of that would be that your Honor is really going to be giving a hard look at staying cases where they're sort of corner cases. They're the rare cases where the reexam was initiated before the litigation was initiated. In the vast majority of the cases, that is not going to be the situation.

If you look at cases that are early litigation, early in the reexam, where the parties are not competitors so there is no concern about prejudice, commercial prejudice, I would say you are not going to see a case in front of you that is a better case for a stay in that set of categories. Because, again, this is a big company. It's a little tiny part of their business, the licensing business. They're

not competitors with us. They now practice the patents. And, to me, all the stars are aligned in favor of the stay in this case, if your Honor is ever going to consider a stay, where a reexamination was filed after the initiation of litigation.

THE COURT: Okay. Thank you.

Mr. Stark.

MR. STARK: Thank you, your Honor.

Your Honor, I think the high level issue, really the main issue on any of these stay motions really has to do with the Court's role as an entity to resolve disputes, to bring cases to trial and thereby resolve disputes in a timely and efficient manner, as Judge Robinson said in one of the numerous cases the parties have cited.

And the question then framed by the stay motion is looking at the three factors that courts have looked at, and spelled out in your Honor's Cooper Notification decision, is granting a stay or denying a stay more consistent with the overall role of the court in efficient dispensation of justice.

And, Your Honor, I would submit in this case clearly all the factors line up in favor of denying the stay. We've touched on them all in the papers. But very briefly on simplification of the issues, we can see no simplification likely to come out of pursuing reexam here

largely for the reason that, as we said in the papers, reexam is going to go on for many, many years and totally interfere with the ability to resolve this case promptly in court.

Many of the defendants' defenses are not even available in the reexamination procedure. And this court, on the other hand, can hold a trial, will be ready for trial on the schedule your Honor entered about a year from now, can hold a trial on all issues and resolve all issues in that one shot. That, we submit, your Honor, very respectfully, is really the way that patent disputes ought to be resolved is the patent holder's right to bring an action to enforce its patent rights -- federally, statutorily granted patent rights in court, and to see those rights adjudicated.

The defendants here simply chose, long after the fact, after the filing of the complaint in this matter, to try to take that forum and shift it to the Patent Office, which we submit is not the fair result here. It's not the plaintiff's choice of forum. And it is their choice, frankly, to have waited that long to file for reexamination. These patents had been asserted and debated between the parties for years prior when defendants could have sought reexamination if they thought that was the appropriate course of action.

THE COURT: If I were to stay, though, and

wait until the PTO proceedings were completely done, there is at least a reasonable likelihood that the issues would be simplified on the back end, isn't there?

MR. STARK: Yes, your Honor. There is always the possibility in reexam situation that there could be some simplification of issues if one were to stay the litigation and if one were to wait however many years it is to get the final outcome from the Patent Office. But two things about that:

One is it will always only be on some of the issues because there are many issues, infringement, equitable defenses, Section 112 defenses and so on, that aren't available to be raised. Not even all the anticipation and obviousness issues can be raised in the Patent Office. So for that reason it really won't end up resolving everything. And,

The second issue is that, of course, it will take many years to reach that point; and by that time, we could have resolved this; and litigation will then act as an estoppel against the reexam.

So the ultimate simplification of issues is simply to resolve them here once and for all.

In terms of how far along this case is compared to Cooper Notification, this case is at least as far along. The time lines are actually very similar between the Cooper

Notification case and this case in terms of when reexam was filed, when the motion to stay was filed. There is a significant difference, however, in the schedules going forward in that this case will have had discovery finished and be ready for trial relatively faster than Cooper Notification was ready for trial or slated to be ready for trial.

In terms of Mr. Fenwick's points with respect to progress of the case, we certainly move this case right along in terms of getting all the documents exchanged from our side anyway, having gotten documents from the other side and reviewed them. We've gone through and confirmed in detail our infringement contentions.

It's true that we haven't had depositions but not for lack of trying. We noticed them and we haven't had any witnesses produced as of yet.

On the other hand, defendants have requested the prosecuting attorney of the patent application and both inventors, and we've granted, we've given dates for those coming up in March and the very beginning of April.

And so the case really has progressed quite substantially in the time that we've had. It's right on the cusp of proceeding right along to finish because we're right into the Markman proceedings now. We're exchanging our contentions on that front. In due course, we'll be

briefing. We have Markman hearings, as Your Honor knows, scheduled for May 19th. So we're right on a roll here of getting this case on its way to be resolved.

The reexamination, on the other hand, is still very much in its infancy. Though the reexamination was granted, there has been no Office action as of yet. So the reexamination on the '979 patent really has not even gotten off the ground as of yet. That is another factor potentially that makes this case relatively stronger in favor of denying a stay than Cooper Notification.

With regard to the prejudice or tactical advantage factor, your Honor, I would submit, first of all, that the factor is either undue prejudice to the nonmoving party Xerox or clear tactical advantage to the moving party, the defendants.

Taking the second one first. There is clearly a tactical advantage here to the defendants, and I would submit that is clearly why this motion was filed. In general, that is why defendants file motions such as this. But certainly six months into a case and trying to get a stay of six, seven, eight years I think would be impossible to characterize that as anything but a major tactical advantage for defendants and clearly why they are seeking to do that. That, in itself, is a substantial factor weighing in favor of denial of the stay.

With respect to prejudice to Xerox, the defendants have suggested, well, there can't be any prejudice if they're not direct competitors of the defendants. And I submit, your Honor, that that is really not quite what the cases say. The cases say that where there is direct competition, that is a factor in favor of finding undue prejudice, and that is a factor in favor of denying stay, but the lack of competition directly between them doesn't mean that there is a stay.

Clearly, there are cases that go both ways on that. And there clearly is prejudice to Xerox if it's denied the ability to enforce its patent for six, seven, eight years. That is a very substantial prejudice to their ability to license that patent and essentially effectuates the wasting away of the patent term. The patent has about ten years life on it. If enforcement were stayed for seven or eight years, that is a very substantial part of the ability, the term during which that patent could be enforced.

And that really has everything to do with the value of a patent. What the patent is, is the right to exclude. It isn't anything else. And to take that away from a patent holder, you have taken away a substantial part of the value of the patent. Anyone involved in these litigations could attest to the fact that the possibility

of an injunction is a large part of the value of a patent.

Witness the NTP v RIM case as a non-practicing party who was able to obtain an injunction. And, by the way, the eBay case rejected any notion that there is a categorical rule against a non-practicing party getting an injunction, that they may be entitled to such an injunction. NTP is an interesting example where the injunction clearly has a lot to do with the value of the patent.

Let the patent waste away and all you are talking about is past damages which, yes, can be recovered but the bargaining situation between the parties is very much altered if there is no possibility of an injunction because the term of the patent has wasted away. So that is clearly a very significant prejudice to Xerox, your Honor, as well as loss of evidence.

If the case is stayed for six, seven years, clearly peoples memories will fade, witnesses will be hard to track down or have left the companies or even have passed away by that point, and technologies will have changed as well. Six or seven years in the computer technology business is really the equivalent of eons. Many things will have changed. The technology that is at issue in this case may no longer be particularly relevant to Google, which, again, changed the negotiating landscape and the relative positions of the parties. So delay is very

significant and is a very significant element of prejudice in itself to Xerox here.

For all those reasons, your Honor, I submit that this case is very much even more in favor of denying a stay than was the case in Cooper Notification.

As to just a few of Mr. Fenwick's points. Mr. Fenwick suggested that there will be two tracks of efforts here because the PTO won't suspend its track during the pendency of the case.

True enough, but it was the defendants' choice to invoke the PTO procedure at this point, so I don't think that is something that can be held in their favor.

Mr. Fenwick also suggested that the patent in suit here is a minuscule part of Xerox's business.

I'm not so sure about that, but in any event, one could say that of patents that are at issue in many cases and that I think is an unfair characterization or it is unfair to use a characterization like that to downplay the value of a case. The patent rights, as I said, only exist for a certain amount of time and the Federal Patent Act gives the patentee the right to come into court and seek enforcement of that statute. To simply write the patent off as, well, we don't think it's a big deal I think kind of belittles the whole patent system in a way that I think is not appropriate.

Finally, Mr. Fenwick suggested that, well, if the CRU were to hold that all the claims were not patentable, or if it didn't hold that they were not patentable, the Court could consider a stay, but that may be years before we reach that point. Some of the statistics cited in the cases say that even the first stage, the examiner stage of the reexamination, lasts for, on average, 36 months, and after that, then you go on to the appeal stage. So we're still looking at years before getting any results in all that. And, again, the court is the place to get a full resolution of the issues.

For those reasons, your Honor, we respectfully request the motion be denied.

THE COURT: Go back to tactical advantage for a moment. What the defendants have done here, they have a statutory right to do. It so happens that the law at the moment at least is that is a different burden or a different standard of proof essentially to invalidate a patent at the PTO than it is here in court. I suppose that could be characterize that as seeking a tactical advantage, but there is nothing improper about what they are trying to do. They have a right to do it, don't they?

MR. STARK: Absolutely. They have a statutory right to seek a reexamination, but it is nevertheless a factor set out in case law for courts to consider whether

seeking that remedy is being done for tactical advantage in the litigation. And,

Here, as I said, there is no doubt that that is the case. They could have sought reexamination at a much earlier time, had they chosen to do so. Otherwise, if simply the fact they have a right to bring the reexam were enough to justify a stay, then virtually every case would be stayed.

In fact, Congress considered reenacting an automatic stay as part of the reexam procedure and chose not to do that, leaving it to the discretion of the court, which really brings me back to our story on this. It's really, I think on the motion in question, of what is more consistent with the court's role of dispensing justice in a timely and efficient manner.

In a case where a reexam was started very early or even before litigation and where it was likely to lead to decisions that would affect the course of the litigation and we're not much invested in the litigation, you can see where a stay might be granted.

But in a case we've already been at this for a year and the reexam was sought well into the case, and where we've only got the one patent at issue and the ability to get this case through to trial long before reexam could reach conclusion, then I would submit we're not going to

see any gain of efficiency or any timely administration of justice.

THE COURT: Okay. Thank you very much.

MR. STARK: Thank you, your Honor.

THE COURT: Mr. Fenwick.

MR. FENWICK: Your Honor, with respect to the time line of the PTO, in the central reexamination unit, the pre-appeal time line, statistically, if you look at that over all the examinations, the pendency there has been in the range of two to two and-a-half, approaching three years, depending on what set of data you look at.

The PTO, the CRU expedites petitions that are the subject of litigation and then it extra-expedites petitions that are the subjects of litigations that have been stayed. So the CRU has a goal of reaching an action closing prosecution at the CRU level within two years of the petition being filed. And the CRU, if you believe what you read, is making good progress toward bringing the pendencies down to that two-year period. There is good reason to think that with the expedition that will take place because of this patent is in litigation, has been stayed, that we could get to the appeal process in the PTO even before two years.

This is not a case that is even remotely in the ballpark of a permanent injunction. It's just not even close. And NTP is an interesting example for Mr. Stark to

bring up because that is a situation where a company was ultimately forced into a multimillion dollar settlement after a court case proceeded, put them right at the knife's edge only to have the patents invalidated by the PTO after they had to write a very big check.

There is a very substantial likelihood, even a probability of simplification here because we know that in approximately half of reexamination situations in a party's reexamination situation, all claims are cancelled.

The plaintiff here is simply not going to amend its claims, so we can assume without its ability to amend the claims, the likelihood is somewhere north of 50 percent that the claims are going to be cancelled. In that case, the simplification is we're done. We don't have to worry about the other invalidity issues. We don't have noninfringement issues. So the simplification is substantial.

As far as the process playing out six, seven, eight years in the PTO, we just don't know that. We don't know that. And, again, your Honor can take a look at this situation once the CRU has done its work, which again has never been done before. No one has ever looked at these references against this patent under KSR or not under KSR to assess the validity of the patent claims.

THE COURT: Okay. Thank you very much.

MR. STARK: Your Honor, just --
THE COURT: You can have the last word, if you wish.
MR. STARK: Just very briefly. Thank you, your Honor. Just on those last couple of points.
If the CRU takes two years, that is still far beyond the point where we would be ready for and, subject to your Honor's schedule, done with trial here.
On invalidation of the patents after the injunction had been granted. In NTP, it can't happen here because the PTO proceedings would be mooted by the effect of collateral estoppel under Section 317 of the patent code. So that is not applicable here. And,
Lastly, Mr. Fenwick asserts that no injunction is in the realm of possibility here. I think it's a bit too early to know that, frankly. We haven't been through the merits of this case. I would take a very different view of this. I think that is something that will be played out later when your Honor has had a chance to consider the merits.
Thank you, your Honor.
THE COURT: Okay. Thank you. We're going to take a short recess and then we'll come back in.
(Brief recess taken.)
THE COURT: Have a seat, please.

I'm going to give you the Court's ruling on defendants' motion to stay.
We had, of course, thoroughly reviewed the briefs before you all got here, and the argument gave us the additional consideration that we needed in order to reach a conclusion.
To dispense with any suspense, I'm going to deny the motion to stay. Let me explain some of my reasoning for that decision.
As we all know, whether or not to stay litigation pending a PTO reexamination is a matter left to the Court's discretion. In my view, each case needs to be evaluated individually. I think it's fair to say that on a discretionary decision such as this, reasonable minds very much could differ, and there is lots of good points on both sides.
In exercising its discretion, the Court must weigh the competing interests of the parties in an attempt to maintain an even balance. The three factors that are typically looked at have all been referred to in the briefing and here today in argument, and let me go through them for you.
The first, whether a stay will simplify the issues in trial of the case.
Here, I think that there is a reasonable

possibility of simplification, but there is also a near certainty that it will take quite a long number of years before that will occur.
The lawsuit involves many disputes related to the '979 patent, only some of which could possibly be resolved in the reexam.
The defendants raise invalidity defenses under Sections 101, 112 and 116, a marking defense pursuant to Section 287, and equitable defenses of estoppel and laches, and none of those could be resolved by the PTO.
There are, of course, other defenses under Sections 102 and 103 which could be resolved in the reexamination, and, of course, the reexamination proceeding will involve the validity of each of the claims of the '979 patent. So, again, I recognize there is a possibility of simplification, but, again, a certainty that it will take a long time and also relevant that the PTO can't resolve everything whereas the Court could.
The Yahoo! defendants I note, although not parties to the reexamination, agreed to be estopped by the PTO proceeding results to the same extent Google is, and that is helpful, but there are even some categories of prior art which cannot be considered in the reexaminations under Sections 301 and 311.
Relatedly, Xerox has represented that it will

not be seeking to amend its claims in the reexamination, which I guess raises the stakes that the path to simplification, if it were to be achieved, will require cancellation of the claims.
In terms of the potential of simplification, I guess the bottom line is it's very attractive to me to stay the case. Any chance of simplification, even any chance of significant delay with my docket and with the resources that this Court has is attractive, and I have considered it. But I just fundamentally think, as plaintiff's counsel argued, the Court is here to do a job. The plaintiff has brought the dispute to the Court. The plaintiff had the right to do so.
While we have limited resources, we have the resources to handle the case. While it's difficult and while it's going to take longer than any of us would wish, it's not going to take as long it takes the PTO. And so notwithstanding the possibility of simplification and the attractiveness of letting the PTO deal with it instead of us, simplification does not weigh strongly for a stay in this case.
In terms of status of litigation. While I think it is fair to say that the bucket is relatively early in litigation, relatively early in reexamination, because our process here in court is evidently much more accelerated

than the typical PTO process, even though we're early, we're already ahead of the PTO, and this case is about to pick up steam significantly.

As both parties know, a scheduling order has been entered. Under the scheduling order, the parties were bound to be at least substantially complete with document production and source code production some time last year. Fact discovery is now set to close in July of this year, and expert discovery in September of this year. Briefing on claim construction is about to begin. We have a claim construction hearing set for May. And the case will move along thereafter and be ready for trial within a reasonable time frame consistent with how this Court typically proceeds. Depositions have been noticed on both sides. So, again, the case is well underway. Notwithstanding the fact that it's fair to say it's early, it's still well underway.

The reexamination, also early, and really I think can't fairly be characterized as well underway. The request for reexamination was made in August of 2010, about six months after the case was initiated, and the reexamination was granted only in December of 2010, about three months ago and about ten months after the case here was initiated. The PTO has not yet issued an Office action. So the status of the litigation I think does not favor a stay. And,

Then the final set of issues or characteristics to consider are unfair prejudice or tactical advantage or disadvantage; and I thought about both sides' prejudice and tactical advantage for either side on this set of criteria. I think it's noteworthy here that granting a stay would effectively deprive the plaintiff of his choice to forum by transplanting this dispute at least for quite awhile to the PTO.

There would be significant delay, I think we all agree, if the case went to the PTO and was stayed in the meantime, which delay always creates a risk of stale evidence, faded memories, lost documents. Plaintiff's case for infringement is fact specific and depends largely on the minds or at least in large part on the minds of witnesses and how the accused products work today. And I think both very likely could be much harder to prove after a lengthy reexamination, even if that were only two or three years from now but it seems to me more likely to be longer than that.

Given that the technology here is a computer technology, I think that any delay hurts the plaintiff even more give that the defendants' products may very rapidly change, which could make it that much harder to prove what they looked like at a particular time.

By contrast, it seems less likely that the

defendants will be harmed in an evidentiary way by delay because their invalidity case will be based necessarily on prior art which I think will require less, will depend less on fact testimony that is time specific.

The plaintiff's rights also will be hindered during the pendency of a stay. Plaintiff represents that its business includes the development of technology which it subsequently tried to license and the cloud of reexamination but also the inability, if it were to happen, to get the case fully resolved at a trial, which would be the result of the stay, would hinder the plaintiff's efforts to license the technology, reducing unfairly in these circumstances its business interests.

All of this is, without making any decision, of course, today as to whether injunctive relief is or is not likely to be available at the end of this case. Even assuming the plaintiff wins the case, even assuming that injunctive relief is not likely, I still reach the same conclusion that I do in terms of how the stay factors play out, but I do want to have a side note I'm not making any decision today, of course, as to whether, if the plaintiff wins, they're going to be entitled to injunctive relief.

It is relevant, as defendants point out, that there is no direct competition between the plaintiff and the defendants with respect to this technology. But while that

is a factor that supports the defendants' position, it's clearly not dispositive here.

And in terms of tactical advantage and tactical disadvantage, I think it's also worth noting that the defendants were on notice from the plaintiff that these patents were being asserted against it for some time, I think in the nature of close to three years I believe, before it filed its reexamination. While I'm not faulting defendants for waiting and nor do I wish to invite a flood of reexams to the PTO, that decision has the consequence of placing us in that bucket of very early stage reexam and early stage but more advanced litigation which is relevant to the analysis today.

So for all of those reasons, having balanced all these considerations, having exercised the Court's discretion, having made an individualized case specific decision, the Court decides that it will deny defendants' motion to stay.

We have two outstanding discovery issues that with talked about on the phone. I want to give you my rulings on those as well.

First, we have defendants' request that plaintiff be made to supplement its responses to defendants' contention interrogatories. I'm going to grant defendants' request there.

Among other things, the Court has compared the level of specificity provided in defendants' responses to plaintiff's contention interrogatories relating to invalidity and compared those to the plaintiff's responses and agrees with defendant that defendants' responses are significantly more detailed.

While the Court recognizes that perhaps it's easier to be more detailed when you responding to invalidity contentions which are based on prior art, and, of course, the defense has put time into reviewing prior art in connection with the reexam, this case is moving along, as I've just indicated, and I think the time has come where plaintiff can, and must, do more than it has in response to the defendants' contention interrogatories.

The letter that plaintiff submitted to the Court with respect to this discovery dispute did do more in terms of identifying documents that have been produced in discovery and identifying with some specificity and some instances as to where claim limitations can be found to be being practiced in those documents. That is, plaintiff did better in its letter than it did in its currently standing responses to interrogatories. So we know it can do better and at this time, it need to do so.

We're going to give plaintiff until March 15th to supplement its responses. My hope very much is that that

response will be satisfy to the defendant at this point in the case. If that turns out not to be the case, then I'm not sure I will be speaking to all of you on the phone, again, in the future.

Then, finally, we have, I guess it's defendants' request to prevent plaintiff from proceeding with newly asserted claims. The Court is going to deny that request from the defendant. The Court will permit plaintiff to proceed with asserting claims 2, 3, 5, 10, and 19 in addition to the two previously asserted claims which were claims 1 and 18. This, too, is a discretionary decision.

Having looked at the patent and at the claims, it appears that the scope of the case will not be substantially expanded by the addition of these additional asserted claims. It won't surprise me if there are some additional claim terms that need to be construed as a result of the newly asserted claims but, in context, particularly given where we are with claim construction, it does not seem to me that it will overly burden the defendants or the Court to permit plaintiff to do this.

I am mindful in that regard, as I have already noted with the reexam, that defendants undoubtedly have spent extensive time studying the patent and the prior art in order to formulate their very lengthy reexamination papers in which Google asserts that all of the claims,

including the newly asserted ones here, are anticipated by prior art references and the Yahoo! defendant has agreed to be bound by that so presumably it has reviewed that submission and at least begun some level of study of the newly asserted claims.

I will mention timing is definitely a little bit troubling, particularly we were on the phone about a week before the plaintiff asserted these new claims and there was no hint of it. But I have counsel's representation that it was not something that was under consideration by plaintiff a week before. It was asserted, and I have no basis to not believe that representation. So I'm denying, again, the defendants' request.

I know that the joint claim construction chart had initially been due today, but I stayed that obligation until giving you this ruling today. My thought was that you all could get the joint claim construction chart done by a week from today, but if that poses a substantial hardship to anyone, I would be happy to hear that and try to fix it.

Mr. Stark, how does that sound?

MR. STARK: Perfect by us, your Honor. No problem.

THE COURT: And from the defense?

MR. FENWICK: Your Honor, if I could ask for ten days?

THE COURT: Ten days?

MR. FENWICK: Yes.

THE COURT: Any objection, Mr. Stark?

MR. STARK: No objection.

THE COURT: Okay. We'll make it due in ten days.

I think that resolves everything that we had outstanding.

Is there anything further at this time, Mr. Stark?

MR. STARK: No, your Honor. Thank you very much.

THE COURT: And, Mr. Fenwick, anything else?

MR. FENWICK: No, your Honor.

THE COURT: All right. Thank you all very much.

(Hearing ends at 4:39 p.m.)

D

E

F

G

H

I

J

Q

R

S

T

U

V

W

X

Y

EXHIBIT 2

IN THE UNITED STATE DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

XEROX CORPORATION Plaintiff, v. GOOGLE INC., YAHOO! INC., and RIGHT MEDIA LLC Defendants.	**C.A. No. 1:10-cv-00136-LPS** **HIGHLY CONFIDENTIAL-OUTSIDE COUNSEL ONLY**

DEFENDANT GOOGLE INC.'S FOURTH SUPPLEMENTAL OBJECTIONS AND RESPONSES TO XEROX'S FIRST SET OF INTERROGATORIES TO DEFENDANTS (NOS. 7)

Pursuant to Federal Rules of Civil Procedure 26 and 33, Defendant Google Inc. hereby further objects and respond in writing to Interrogatory No. 7 of Plaintiff Xerox Corporation's First Set of Interrogatories to Defendants.

GENERAL OBJECTIONS

Google and YouTube make the following general objections to each and every definition, instruction, and interrogatory made in Xerox's First Interrogatories to Defendants. Each of these objections is incorporated into the Specific Objections set forth below, whether or not separately set forth therein. By responding to any of the interrogatories or failing to specifically refer to or specify any particular General Objection in response to a particular interrogatory, Google and YouTube do not waive any of these General Objections, nor admit or concede the appropriateness of any purported interrogatory or any assumptions contained therein.

1. Nothing in these responses should be construed as waiving rights or objections that might otherwise be available to Google and YouTube nor should Google and YouTube's responses to any of these interrogatories be deemed an admission of relevancy, materiality, or admissibility in evidence of the interrogatory or the response thereto.

2. Google and YouTube object to each interrogatory to the extent that it seeks the disclosure of information protected from disclosure by the attorney-client privilege, the attorney work product doctrine or any other applicable privilege or protection as provided by law. Google and YouTube will not produce such privileged or protected information, and any inadvertent disclosure of any privileged or protected information should not be deemed a waiver of any privilege.

3. Google and YouTube object to each interrogatory, and to the definitions and instructions, to the extent they purport to impose upon Google and YouTube obligations broader than, or inconsistent with, the Federal Rules of Civil Procedure or the Local Rules and Orders of this Court.

4. Google and YouTube object to each interrogatory, and to the definitions and instructions, to the extent that they are overbroad, vague and ambiguous, unduly burdensome and oppressive, in purporting to require Google and YouTube to search facilities and inquire of employees other than those facilities and employees that could reasonably be expected to have responsive information, or produce information outside a relevant time period or unrelated to the asserted claims of the patent-in-suit. In particular, Google and YouTube object to Xerox's definition of "personalized search" as vague, ambiguous, and overbroad. Google and YouTube will not produce documents and information that are irrelevant, immaterial or not reasonably

calculated to lead to the discovery of admissible evidence. Google and YouTube also will not produce information that is not in its possession, custody or control.

5. Google and YouTube object to each interrogatory to the extent it seeks information already in Xerox's possession or equally available to Xerox from other sources that are more convenient, less burdensome and/or less expensive.

6. Google and YouTube object to each interrogatory and to the definitions and instructions included therewith pursuant to Federal Rule of Civil Procedure 26(b)(2)(i) to the extent that they purport to require the disclosure of information that is more readily available and/or more appropriately obtainable through other means of discovery.

7. Google and YouTube object to each interrogatory to the extent that it is compound and/or is comprised of subparts constituting more than one interrogatory, particularly in view of Xerox's instructions with respect to each "subpart" of each interrogatory as each subpart properly counts as separate interrogatories against the limit of interrogatories for Xerox in this case.

8. Google and YouTube object to these interrogatories to the extent that such interrogatories, when properly counted, exceed the limit for interrogatories available to Xerox in this case.

9. Google and YouTube object to each interrogatory, and to the definitions and instructions included therewith, to the extent they seek proprietary, trade secret or other confidential or competitively sensitive business information. Subject to Local Rule 26.2, Google and YouTube will only produce such relevant, non-privileged information subject to adequate protections for Google and YouTube's confidential, trade secret and/or proprietary business or technical information via a protective order entered by the Court in this action.

10. Google and YouTube object to each interrogatory, and to the definitions and instructions included therewith, to the extent that they purport to Require Google and YouTube to disclose private or personally-identifiable information of its users.

11. Google and YouTube object to each interrogatory, and to the definitions and instructions included therewith, to the extent that they purport to require Google and YouTube to disclose information that is subject to any protective order, privacy interest, contractual obligation, or other confidentiality obligation owed to any third party.

12. Google and YouTube object to each interrogatory to the extent that such interrogatory prematurely seeks the production of information and documents in advance of the dates set by the Federal Rules of Civil Procedure, the Local Rules, or any orders entered by this Court.

13. Google and YouTube object to each interrogatory as premature and unduly burdensome to the extent that it seeks information likely to depend on construction of claim terms and/or expert analysis of the patent-in-suit, the deadlines for which have not yet been set.

14. Google and YouTube object to each interrogatory as premature and unduly burdensome to the extent that it seeks discovery regarding non-infringement of any claim(s) of the patent-in-suit for which Xerox has not provided a substantive contention that Google and/or YouTube practice every element of such claim(s).

17. Google and YouTube object to each interrogatory as premature and unduly burdensome to the extent that it seeks discovery before Xerox pleads facts sufficient to define each and every accused instrumentality and how they could plausibly infringe the patent-in-suit.

18. Google and YouTube object to each interrogatory as unduly burdensome to the extent it seeks information about every version or release of purportedly accused technology or

functionality. The burden and expense associated with producing such information grossly outweighs its benefit and relevance.

19. Google and YouTube object to Xerox's definitions of the terms "Content Matching Products," "Google Content Matching Products," and "Accused Products" as vague, overbroad, unduly burdensome, and oppressive.

20. Google and YouTube object to Xerox's definition of the term "Google Maps" as vague, overbroad, unduly burdensome, and oppressive, particularly to the extent it encompasses products, services and software that display "information related to maps, addresses, directions, points of interest and/or businesses."

21. Google and YouTube object to Xerox's definition of the term "Google Video" as vague, overbroad, unduly burdensome, and oppressive, particularly to the extent it encompasses products, services and software that display "information related to videos."

22. Google and YouTube object to Xerox's definitions of the term "Youtube.com" as vague, overbroad, unduly burdensome, and oppressive, particularly to the extent it encompasses products, services and software that display "information related to videos."

23. Google and YouTube object to Xerox's definitions of the term "Predecessor Product," as vague, overbroad, unduly burdensome, and oppressive. In particular, it is not clear what "subsequent product, service, facility and/or computer software program" refers to. To the extent it is meant to refer to the accused products as defined elsewhere in Xerox's requests, Google and YouTube object on the ground that it cannot be expected to identify every "product, service, facility and/or computer software product" any part of which was "directly or indirectly used" in the creation of any accused product, regardless of relevance. The burden and expense associated with producing such information grossly outweighs its benefit and relevance.

24. Google and YouTube object to Xerox's definition of the term "Related Products," as vague, overbroad, unduly burdensome, and oppressive. Google and YouTube cannot be expected to identify all "products, service, facilities and/or computer software product" that "in any manner include, reference, utilize, call or invoke any of the Accused Products," regardless of relevance. The burden and expense associated with producing such information grossly outweighs its benefit and relevance.

25. Google and YouTube object to Xerox's definition of the term "'979 Accused Products" as vague, overbroad, unduly burdensome, and oppressive, particularly to the extent that it incorporates Xerox's overbroad definition of the term "Google Content Matching Products."

26. Google and YouTube object to Xerox's definition of the term "'994 Accused Products" as vague, overbroad, unduly burdensome, and oppressive, particularly to the extent that it incorporates Xerox's overbroad definitions of the terms "Google Maps," "Google Video," and "YouTube.com."

27. Google and YouTube object to each interrogatory, definition, and instruction to the extent the burden or expense of the proposed discovery outweighs its likely benefit, considering the needs of the case, the amount in controversy, the importance of the issues at stake in the action, and the importance of the discovery in resolving the issues.

28. Google and YouTube respond to these interrogatories based upon its current understanding and reserves the right to supplement its responses if any additional information is identified at a later time and to make any additional objections that may become apparent.

29. Each of Google and YouTube's responses to these interrogatories are made subject to and without waiving, limiting, or intending to waive:

A. each of the above-stated general objections and reservations;

B. the right to object on the grounds of competency, privilege, relevancy, or materiality, or any other proper grounds, to the use of the documents or information, for any purpose, in whole or in part, in any subsequent step or proceeding in this action or any other action;

C. the right to object on any and all grounds, at any time, to other discovery requests involving or relating to the subject matter of the present litigation; and

D. the right at any time to revise, correct, and add to or clarify any of the responses herein.

30. By responding to these interrogatories, Google and YouTube do not waive or intend to waive, but expressly reserves, all of its statements, reservations, and objections, both general and specific, set forth in these responses, even though Google and YouTube may in some instances disclose information over the statements, reservations, and objections contained herein.

31. Pursuant to the Court's May 11, 2010 Order bifurcating the issues of infringement and invalidity from the issues of willfulness and damages, Google and YouTube will not be providing documents or information related to the issues of willfulness or damages until the commencement of bifurcated discovery on those issues.

STATEMENT ON SUPPLEMENTATION

Google and YouTube's investigation in this action is ongoing, and Google and YouTube reserve the right to rely on and introduce information in addition to any information provided herein at the trial of this matter or in other related proceedings. Google and YouTube have yet to receive complete discovery responses from Xerox. Google and YouTube anticipate that facts they learn later in the litigation may be responsive to one or more of the interrogatories and

Google and YouTube reserve their right to supplement these interrogatories at appropriate points throughout this litigation without prejudice and/or to otherwise make available to Xerox such information. Google and YouTube also reserve the right to change, modify or enlarge the following responses based on additional information, further analysis, and/or in light of events in the litigation such as rulings by the Court. Google and YouTube reserve the right to rely on or otherwise use any such amended response for future discovery, trial or otherwise.

SPECIFIC OBJECTIONS AND RESPONSES

Google and YouTube expressly incorporate the above objections as though set forth fully in response to each of the following individual interrogatories, and, to the extent that they are not raised in the particular response, Google and YouTube do not waive those objections.

INTERROGATORIES

INTERROGATORY NO. 7:

If you contend that any claim of the Patents in Suit is invalid and/or unenforceable, specify each claim that you contend is invalid and/or unenforceable and describe in full for each such claim the basis for your contention, identifying all prior art, all documents and all facts that you believe support your contention.

RESPONSE TO INTERROGATORY NO. 7:

Google and YouTube incorporate here in response to this interrogatory their General Objections above by this reference. Google and YouTube object to this interrogatory on the ground that it is compound and/or is comprised of subparts constituting more than one interrogatory. Google and YouTube further object to this interrogatory as premature as Xerox

has not yet set forth its allegations of infringement or identified all of the claims it intends to assert against Google and YouTube.

SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 7:

Subject to the foregoing general and specific objections, Google and YouTube further respond as follows:

The '994 Patent:

The '994 Patent is invalid under 35 U.S.C. § 101 to the extent that it attempts to cover unpatentable abstract ideas. *See Bilski*. *See Bilski v. Kappos*, 561 U.S. __, slip op. at 3 (2010).

The asserted claims of the '994 Patent are invalid under 35 U.S.C. § 102 and/or § 103 because at least the following prior art references anticipate the claims or render them obvious, alone or in combination:

Patents or Patent Applications:

US 5,367,619 (Diapaolo)

US 5,649,192 (Stucky)

US 5,987,440 (O'Neil)

US 5,077,666 (Brimm)

US 6,141,694 (Gardner)

Publications:

Rennison, Galaxy of News: An Approach to Visualizing and Understanding Expansive News Landscape, Proceedings of the 7th annual ACM symposium on User interface software and technology (1994)

Systems in Prior Public Use (beyond those already listed):

The Internet Movie Database

Google incorporates by reference herein the identification by other defendants of any Prior Art as invalidating claims of the '994 Patent under § 102 and/or § 103, to the extent such Prior Art is not specifically identified above. Google reserves the right to use any of the listed references in support of an argument based on a disclosed system in prior use.

Based on Plaintiff's apparent construction of the claims of the '994 patent (as expressed in its response to Google and YouTube.com's Interrogatory No. 2), and based at least upon the use of the terms "performing data analysis operations," "generate data and analysis results," "independently storing the knowledge, in the form of documents," "document database," "validating the accuracy of the knowledge," "making the stored knowledge available across a network," "managing the flow of information," "integration of the data and analysis results with the documents," "updating the documents," and "a change in the data or analysis results" the claims of the '994 Patent are invalid under 35 U.S.C. § 112 for indefiniteness, non-enablement, and inadequate written description.

The '979 Patent:

The '979 Patent is invalid under 35 U.S.C. § 101 to the extent that it attempts to cover unpatentable abstract ideas. *See Bilski*. *See Bilski v. Kappos*, 561 U.S. __, slip op. at 3 (2010).

The asserted claims of the '979 Patent are invalid under 35 U.S.C. § 102 and/or § 103 because at least the following prior art references anticipate the claims or render them obvious, alone or in combination:

Patents or Patent Applications:

US 6,546,386 (Black)

US 7,225,180 (Donaldson)

US 6,236,768 (Rhodes)

US 5,893,092 (Driscoll)

US 6,363,378 (Conklin)

US 6,947,920 (Alpha)

US 7,047,242 (Ponte)

US 7,089,236 (Stibel)

US 5,488,725 (Turtle)

US 5,748,954 (Mauldin)

US 5,963,940 (Liddy)

US 6,038,561 (Snyder)

US 6,161,084 (Messerly)

US 6,519,586 (Anick)

US 2003/0014405 (Shapiro)

US 2002/0052898 (Schilit)

US 5,321,833 (Chang)

PCT/US00/41713 (publication no: WO 20 01/44992A1) (YellowBrix)

Publications:

Pazzani, et al., Syskill & Webert: Identifying interesting web sites, AAAI-96 Proceedings (1996)

Salton, Another Look at Automatic Text-Retrieval Systems, Comm. of ACM (1986)

Google incorporates by reference herein the identification by other defendants of any Prior Art as invalidating claims of the '979 Patent under § 102 and/or § 103, to the extent such

Prior Art is not specifically identified above. Google reserves the right to use any of the listed references in support of an argument based on a disclosed system in prior use.

The '979 Patent may also be invalid under 35 U.S.C. §§ 102(f) and 116 for failing to include all inventors of the claimed subject matter, pending further investigation.

Google and YouTube.com reserve the right to supplement this response as their investigation continues.

SECOND SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 7:

Subject to the foregoing general and specific objections, Google further supplements its response to this Interrogatory as follows:

This Supplemental Response addresses only claims 1 and 18 of the '979 Patent. On February 9, 2011, Xerox alleged for the first time that certain Google products infringe claims 2, 3, 5, 10, and 19 of the '979 Patent. Google has objected to the assertion of these additional claims. Google will supplement this Response in due course to address these additional claims if and to the extent that the Court allows Xerox to expand its infringement case to encompass them. Xerox has indicated that it is dropping all claims based on the '994 Patent, and, accordingly, this Supplemental Response does not address the '994 Patent.

Exemplary claim charts under 35 U.S.C. § 102(a), (b), (e), (f) and/or (g), and/or § 103 are attached as Charts A-1 through A-13 for the claims 1 and 18 of the '979 patent, identified by plaintiff Xerox Inc. in its Response to Google Interrogatory Nos. 1 and 2. Google incorporates herein the discussion of prior art references and the invalidity arguments under 35 U.S.C. sections 102 and/or 103 set forth in its Corrected Request for Inter Partes Reexamination of the '979 Patent filed with the United States Patent and Trademark Office on or about September 8,

2010. Google further directs Xerox to all subsequent proceedings in connection with the reexamination of the '979 Patent.

Google expressly reserves the right to amend the disclosures herein should Xerox provide any information that it failed to provide in its infringement contentions or should Xerox amend its infringement contentions. Further, because Google has not yet completed its search for and analysis of relevant prior art, Google reserves the right to revise, amend, and/or supplement the information provided herein, including identifying and relying on additional references, should Google's further search and analysis yield additional information or references, consistent with the Federal Rules of Civil Procedure. Moreover, Google reserves the right to revise its ultimate contentions concerning the invalidity of the claims of the '979 patent, which may change depending upon the Court's construction of the claims of the '979 patent, any findings as to the priority date of the '979 patent, and/or positions that Xerox or its expert witness(es) may take concerning claim interpretation, infringement, and/or invalidity issues. Google further reserves the right to supplement its contentions to the extent that Xerox is permitted to assert additional claims of the '979 patent against Google.

Prior art not included in this disclosure, whether known or not known to Google, may become relevant. In particular, Google is currently unaware of the extent, if any, to which Xerox will contend that limitations of the asserted claims are not disclosed in the prior art identified by Google. To the extent that such an issue arises, Google reserves the right to identify other references that would have made the addition of the allegedly missing limitation to the disclosed device or method obvious.

Google's claim charts cite to particular teachings and disclosures of the prior art as applied to features of the asserted claims. However, persons having ordinary skill in the art

generally may view an item of prior art in the context of other publications, literature, products, and understanding. As such, the cited portions are only examples, and Google reserves the right to rely on un-cited portions of the prior art references and on other publications and expert testimony as aids in understanding and interpreting the cited portions, as providing context thereto, and as additional evidence that the prior art discloses a claim limitation. Google further reserves the right to rely on un-cited portions of the prior art references, other publications, and testimony to establish bases for combinations of certain cited references that render the asserted claims obvious.

The references discussed in the claim charts may disclose the elements of the asserted claims explicitly and/or inherently, and/or they may be relied upon to show the state of the art in the relevant time frame. The suggested obviousness combinations are provided in the alternative to Google's anticipation contentions and are not meant to suggest that any reference included in the combinations is not by itself anticipatory

For purposes of this interrogatory response, Google identifies prior art references and provides element-by-element claim charts based in part on the apparent constructions of the asserted claims advanced by Xerox. Nothing stated herein shall be treated as an admission or suggestion that Google agrees with Xerox regarding either the scope of any of the asserted claims or the claim constructions advanced by it in its infringement contentions or anywhere else. Moreover, nothing in this interrogatory response shall be treated as an admission that Google's accused technology meets any limitation of the claims.

Depending on the Court's construction of the claims of the '979 patent, and/or positions that Xerox or its expert witness(es) may take concerning claim interpretation, infringement, and/or invalidity issues, different charted prior art references may be of greater or lesser

relevance and different combinations of these references may be implicated. Given this uncertainty, the charts may reflect alternative applications of the prior art against the asserted claims.

Google further intends to rely on inventor admissions concerning the scope of the prior art relevant to the '979 patent found in, *inter alia:* the patent prosecution histories for the '979 patent and related patents and/or patent applications; any deposition testimony of the named inventors; and the papers filed and any evidence submitted by Xerox in conjunction with this litigation.

Discovery is ongoing, and Google's prior art investigation and third party discovery is therefore not yet complete. Google reserves the right to present additional items of prior art under 35 U.S.C. § 102(a), (b), (e), (f) and/or (g), and/or § 103 located during the course of discovery or further investigation. For example, Google expects to issue subpoenas to third parties believed to have knowledge, documentation and/or corroborating evidence concerning some of the prior art listed in the Interrogatory response and/or additional prior art. These third parties include without limitation the authors, inventors, or assignees of the references listed in the Interrogatory response. In addition, Google reserves the right to assert invalidity under 35 U.S.C. § 102(c) or (d) to the extent that discovery or further investigation yield information forming the basis for such claims.

Further, based on Google's present understanding of the asserted claims of the '979 patent that Google believes Xerox to be asserting based on Xerox's proposed constructions and it infringement Contentions, Google believes that the charted references anticipate the claims of the '979 patent as shown in the references' respective charts. However, if the finder of fact determines that some element of a given claim was not disclosed by an anticipation reference,

that reference in combination with the knowledge and skill of a person of ordinary skill in the art at the time of the alleged invention and/or other prior art disclosing the allegedly missing limitations would have rendered each of the asserted claims obvious.

The Supreme Court has held that the combination of familiar elements according to known methods is likely to be obvious when it does no more than yield predictable results. *KSR Intl Co. v. Teleflex Inc.,* 550 U.S. 398, 127 S. Ct. 1727, 1739 (2007). When a work is available in one field of endeavor, design incentives and other market forces can prompt variations of it, either in the same field or a different one. *Id.* at 1740. For the same reason, if a technique has been used to improve one device, and a person of ordinary skill in the art would recognize that it would improve similar devices in the same way, using the technique is obvious unless its actual application is beyond his or her skill. *Id.*

In order to determine whether there is an apparent reason to combine the known elements in the fashion claimed by the patent at issue, a court can look to interrelated teachings of multiple patents; the effects of demands known to the design community or present in the marketplace; and the background knowledge possessed by a person having ordinary skill in the art. *Id.* at 1740-41. For example, obviousness can be demonstrated by showing there existed at the time of invention a known problem for which there was an obvious solution encompassed by the patent's claims. *Id.* at 1743. Any need or problem known in the field of endeavor at the time of invention and addressed by the patent can provide a reason for combining the elements in the manner claimed. *Id.* Common sense also teaches that familiar items may have obvious uses beyond their primary purposes, and in many cases a person of ordinary skill will be able to fit the teachings of multiple patents together like pieces of a puzzle. *Id.*

Thus, the motivation to combine the teachings of the prior art references disclosed herein is found in the references themselves and/or: (1) the nature of the problem being solved, (2) the express, implied and inherent teachings of the prior art, (3) the knowledge of persons of ordinary skill in the art, (4) the fact that the prior art is generally directed towards providing personalized information services to a user, and/or (5) the predictable results obtained in combining the different elements of the prior art.

Based on Google's present understanding of the asserted claims of the '979 patent and the constructions that Google believes Xerox to be asserting based on Xerox's proposed constructions and its Infringement Contentions, the asserted claims of the '979 patent are obvious in light of the combinations outlined in the attached charts. Each of these combinations yields predictable results.

THIRD SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 7:

Subject to the foregoing general and specific objections, Google further supplements its response to this Interrogatory as follows:

The asserted claims of the '979 Patent are invalid under 35 U.S.C. § 102 and/or § 103 because at least the following additional prior art references anticipates the claims or renders them obvious, alone or in combination:

Publications (beyond those already listed):

Oracle Corp., *Oracle Text, An Oracle Technical White Paper* (May 2001)

(available at

http://ugweb.cs.ualberta.ca/~c391/W08/resources/oracle_text.pdf)

Oracle Corp., *Oracle8i interMedia Text Reference, Release 2* (8.1.6) (December 1999) (http://download.oracle.com/docs/cd/A87860_01/doc/inter.817/a77063.pdf)

Oracle Corp., *Oracle8i interMedia Text Migration, Release 2* (8.1.6) (December 1999) (available at http://download.oracle.com/docs/cd/A87860_01/doc/inter.817/a77061.pdf)

Oracle 8i inter Media Text 8.1.5 - Technical Overview (1999) (available at http://www.oracle.com/technetwork/database/enterprise-edition/imt-815-083189.html)

Systems in Prior Public Use (beyond those already listed):

Oracle Text

Supplemental exemplary claim charts under 35 U.S.C. § 102(a), (b), (e), (f) and/or (g), and/or § 103 are attached as Charts B-1 through B-16 for claims 1, 2, 3, 5, 10, 18, and 19 of the '979 patent, identified by plaintiff Xerox Inc. in its Response to Google Interrogatory Nos. 1 and 2, and are incorporated here. Google further reserves the right to supplement this interrogatory with additional references and charts as Google's investigation continues.

Google expressly reserves the right to amend the disclosures herein should Xerox provide any information that it failed to provide in its infringement contentions or should Xerox amend its infringement contentions. Further, because Google has not yet completed its search for and analysis of relevant prior art, Google reserves the right to revise, amend, and/or supplement the information provided herein, including identifying and relying on additional references, should Google's further search and analysis yield additional information or references, consistent with the Federal Rules of Civil Procedure. Moreover, Google reserves the right to revise its ultimate

contentions concerning the invalidity of the claims of the '979 patent, which may change depending upon the Court's construction of the claims of the '979 patent, any findings as to the priority date of the '979 patent, and/or positions that Xerox or its expert witness(es) may take concerning claim interpretation, infringement, and/or invalidity issues. Google further reserves the right to supplement its contentions to the extent that Xerox is permitted to assert additional claims of the '979 patent against Google.

Prior art not included in this disclosure, whether known or not known to Google, may become relevant. In particular, Google is currently unaware of the extent, if any, to which Xerox will contend that limitations of the asserted claims are not disclosed in the prior art identified by Google. To the extent that such an issue arises, Google reserves the right to identify other references that would have made the addition of the allegedly missing limitation to the disclosed device or method obvious.

Google's claim charts cite to particular teachings and disclosures of the prior art as applied to features of the asserted claims. However, persons having ordinary skill in the art generally may view an item of prior art in the context of other publications, literature, products, and understanding. As such, the cited portions are only examples, and Google reserves the right to rely on un-cited portions of the prior art references and on other publications and expert testimony as aids in understanding and interpreting the cited portions, as providing context thereto, and as additional evidence that the prior art discloses a claim limitation. Google further reserves the right to rely on un-cited portions of the prior art references, other publications, and testimony to establish bases for combinations of certain cited references that render the asserted claims obvious.

The references discussed in the claim charts may disclose the elements of the asserted claims explicitly and/or inherently, and/or they may be relied upon to show the state of the art in the relevant time frame. The suggested obviousness combinations are provided in the alternative to Google's anticipation contentions and are not meant to suggest that any reference included in the combinations is not by itself anticipatory

For purposes of this interrogatory response, Google identifies prior art references and provides element-by-element claim charts based in part on the apparent constructions of the asserted claims advanced by Xerox. Nothing stated herein shall be treated as an admission or suggestion that Google agrees with Xerox regarding either the scope of any of the asserted claims or the claim constructions advanced by it in its infringement contentions or anywhere else. Moreover, nothing in this interrogatory response shall be treated as an admission that Google's accused technology meets any limitation of the claims.

Depending on the Court's construction of the claims of the '979 patent, and/or positions that Xerox or its expert witness(es) may take concerning claim interpretation, infringement, and/or invalidity issues, different charted prior art references may be of greater or lesser relevance and different combinations of these references may be implicated. Given this uncertainty, the charts may reflect alternative applications of the prior art against the asserted claims.

As noted in the specification of the ‘979 Patent, “Many products provide various solutions for individual aspects of the overall problem of knowledge management: anticipatory services, unstructured information management, and visualization of information and knowledge.” (‘979 Patent, col. 2:11-15.) Google further intends to rely on admissions of the inventors and of Xerox concerning the scope of this prior art, including at least the following:

Defining an Organized Classification of Document Content

The specification of the '979 Patent cites DMOZ as an example of a known ontology that could be used to classify document content. ('979 Patent, col. 41:5-9, 41:59-60.) [redacted]

Classifying Document Content

The '979 Patent discloses several known techniques for classifying document, including a probabilistic model, fuzzy model, latent semantic indexing, and a vector space mode. ('979 Patent, col. 41:52 – col. 46:67.) [redacted]

Entity Extraction

The '979 Patent states that "[e]ntities can be recognized using a variety of known techniques." ('979 Patent, col. 10:52-55.) [redacted]

[redacted]

Automatically Generating Contextualized Queries

[redacted]

Restricting a Search to a Category of Information

The '979 Patents lists three examples of information retrieval systems to which a contextualized query can be submitted: Google, Yahoo, and Northern Lights. ('979 Patent, col. 49:49-54[redacted]

[redacted]

Limiting the Number of Terms in a Query to a Predefined Number

It was well known that some prior art search engines restricted the number of characters in a query. Limiting the number of terms added to a query is therefore an obvious solution to this known issue.

Discovery is ongoing, and Google's prior art investigation and third party discovery is therefore not yet complete. Google reserves the right to present additional items of prior art under 35 U.S.C. § 102(a), (b), (e), (f) and/or (g), and/or § 103, or other relevant evidence, located during the course of discovery or further investigation. For example, Google expects to issue subpoenas to third parties believed to have knowledge, documentation and/or corroborating evidence concerning some of the prior art listed in the Interrogatory response and/or additional prior art. These third parties include without limitation the authors, inventors, or assignees of the references listed in the Interrogatory response. In addition, Google reserves the right to assert invalidity under 35 U.S.C. § 102(c) or (d) to the extent that discovery or further investigation yield information forming the basis for such claims.

Further, based on Google's present understanding of the asserted claims of the '979 patent that Google believes Xerox to be asserting based on Xerox's proposed constructions and it infringement Contentions, Google believes that the charted references anticipate the claims of the '979 patent as shown in the references' respective charts. However, if the finder of fact determines that some element of a given claim was not disclosed by an anticipation reference, that reference in combination with the knowledge and skill of a person of ordinary skill in the art at the time of the alleged invention and/or other prior art disclosing the allegedly missing limitations would have rendered each of the asserted claims obvious.

The Supreme Court has held that the combination of familiar elements according to known methods is likely to be obvious when it does no more than yield predictable results. *KSR Intl Co. v. Teleflex Inc.,* 550 U.S. 398, 127 S. Ct. 1727, 1739 (2007). When a work is available in one field of endeavor, design incentives and other market forces can prompt variations of it, either in the same field or a different one. *Id.* at 1740. For the same reason, if a technique has been used to improve one device, and a person of ordinary skill in the art would recognize that it would improve similar devices in the same way, using the technique is obvious unless its actual application is beyond his or her skill. *Id.*

In order to determine whether there is an apparent reason to combine the known elements in the fashion claimed by the patent at issue, a court can look to interrelated teachings of multiple patents; the effects of demands known to the design community or present in the marketplace; and the background knowledge possessed by a person having ordinary skill in the art. *Id.* at 1740-41. For example, obviousness can be demonstrated by showing there existed at the time of invention a known problem for which there was an obvious solution encompassed by the patent's claims. *Id.* at 1743. Any need or problem known in the field of endeavor at the time of invention and addressed by the patent can provide a reason for combining the elements in the manner claimed. *Id.* Common sense also teaches that familiar items may have obvious uses beyond their primary purposes, and in many cases a person of ordinary skill will be able to fit the teachings of multiple patents together like pieces of a puzzle. *Id.*

Thus, the motivation to combine the teachings of the prior art references disclosed herein is found in the references themselves and/or: (1) the nature of the problem being solved, (2) the express, implied and inherent teachings of the prior art, (3) the knowledge of persons of ordinary skill in the art, (4) the fact that the prior art is generally directed towards providing personalized

information services to a user, and/or (5) the predictable results obtained in combining the different elements of the prior art.

Based on Google's present understanding of the asserted claims of the '979 patent and the constructions that Google believes Xerox to be asserting based on Xerox's proposed constructions and its Infringement Contentions, the asserted claims of the '979 patent are obvious in light of the combinations outlined in the attached charts. [redacted] Numerous techniques for preventing information overload were already well known by the time the application for the '979 Patent was filed. [redacted] In the '979 Patent itself, the technique disclosed in, for example, Figure 39,

is restricting a search to certain categories of documents using Google's known category-searching technology.

Xerox's improper and expansive interpretation of the "to restrict a search" element of claims 1 and 18 to include adding any search criteria, such as keywords, would also encompass well-known techniques for achieving more precise search results. [redacted]

Further, automating known search-narrowing techniques was also well-known at the time. The '979 Patent, for example, discloses prior art systems like Watson, Autonomy, Zapper.com, and Flyswat; all used document content to generate queries for additional, related information. (*See* '979 Patent at col. 2:10-33.) [redacted]

Google further contents that the '979 patent is invalid under 35 U.S.C. § 102(f) for failing to name all inventors of the '979 patent. [redacted]

Google further contends that some or all of the asserted claims of the '979 patent are invalid under 35 U.S.C. § 112. The following contentions are subject to revision and amendment pursuant to Federal Rule of Civil Procedure 26(e) and the Orders of record in this matter to the extent appropriate in light of further investigation and discovery regarding the defenses, the Court's construction of the claims at issue, and/or the review and analysis of expert witnesses. To the extent that the following contentions reflect constructions of claim limitations consistent with Xerox's infringement contentions and supplements thereto, or Xerox's proposed claim constructions and claim construction briefing, no inference is intended nor should any be drawn that Google agrees with Xerox's claim constructions. Google offers such contentions in response to Xerox's claim construction theories without prejudice to any position it may ultimately take as to any claim construction issues.

Claims 1 and 18 of the '979 patent, and all corresponding dependent claims, are invalid for indefiniteness pursuant to 35 U.S.C. § 112. Specifically, the use of the phrase "selected

document content" lacks any antecedent basis and is insolubly ambiguous. Although the phrase "selected document" appears at various points in the '979 patent specification, at no point does the specification indicate how any document content is selected, by whom such a selection is made, or any other information regarding the selection of document content.

Claims 1 and 18 of the '979 patent, and all corresponding dependent claims, are invalid for failure to disclose the best mode contemplated by the inventors of carrying out their invention pursuant to 35 U.S.C. § 112. [redacted]

Google further reserves the right to further amend its response concerning bases of invalidity under 35 U.S.C. § 112 in response to positions taken by Xerox or the construction of claim terms by the Court. For example, in the event that Xerox asserts that the "defining an organized classification of document content" element of claim 1 of the '979 patent includes generating a classification scheme, that element would be invalid as insufficiently described and not enabled. Other terms may similarly be invalid if construed more broadly than can be supported by the '979 Patent specification, and Google reserves its right to amend these contentions in the event other terms are so construed.

FOURTH SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 7:

Subject to the foregoing general and specific objections, Google further supplements its response to this Interrogatory as follows:

The '979 Patent is invalid under 35 U.S.C. § 101 to the extent that it attempts to cover unpatentable abstract ideas. *See Bilski v. Kappos*, 561 U.S. __, slip op. at 3 (2010). The claims of the '979 patent purport to cover abstract ideas such as defining a classification scheme, identifying salient aspects of a given document, sorting a document by category, and formulating queries to for a search restricted to a category. The claims of the '979 patent are not tied to a particular machine or mechanism, do not transform anything in the document or information retrieval system, and do not include or require specific steps for accomplishing the broad ideas recited in the claims. The claims of the '979 patent reflect nothing more than an automated version of a contextualized search, which was the commonsense approach to limit search results to relevant information employed regularly by users of Internet search engines. [redacted]

The asserted claims of the '979 Patent are invalid under 35 U.S.C. § 102 and/or § 103 because at least the following additional prior art references anticipates the claims or renders them obvious, alone or in combination:

Publications (beyond those already listed):

Alexander Pretschner and Susan Gauch, "Ontology Based Personalized Search," Proceedings of the 11th IEEE Int'l Conf. on Tools with Artificial Intelligence, pp. 391-98 (Nov. 1999).

Second supplemental exemplary claim charts under 35 U.S.C. § 102(a), (b), (e), (f) and/or (g), and/or § 103 are attached as Charts B-1 through B-17 for claims 1, 2, 3, 5, 10, 18, and

19 of the '979 patent, identified by plaintiff Xerox Inc. in its Response to Google Interrogatory Nos. 1 and 2, and are incorporated here. Google further reserves the right to supplement this interrogatory with additional references and charts as Google's investigation continues.

QUINN EMANUEL URQUHART
& SULLIVAN, LLP

By: */s/ David A. Perlson*

Charles K. Verhoeven
David A. Perlson
50 California St., 22nd Floor
San Francisco, CA 94111
Tel.: (415) 875-6600

Andrea Pallios Roberts
555 Twin Dolphin Dr., 5th Floor
Redwood Shores, CA 94065

Attorneys for Defendant Google Inc.

OF COUNSEL:

Richard L. Horwitz (#2246)
David E. Moore (#3983)
POTTER ANDERSON & CORROON LLP
Hercules Plaza, 6th Floor
1313 N. Market Street
Wilmington, DE 19801
Tel: (302) 984-6000
rhorwitz@potteranderson.com
dmoore@potteranderson.com

Dated: May 10, 2011

CERTIFICATE OF SERVICE

I, Matthew D. Cannon, hereby certify that on May 11, 2011, true and correct copies of the within document were served on the following counsel of record at the addresses and in the manner indicated:

VIA ELECTRONIC MAIL

Lawrence C. Ashby
John G. Day
Lauren E. Maguire
ASHBY & GEDDES
500 Delaware Avenue, 8th Floor
Wilmington, DE 19899
lashby@ashby-geddes.com
jday@ashby-geddes.com
lmaguire@ashby-geddes.com
Attorneys for Plaintiff Xerox Corporation

Richard J. Stark
Andrei Harasymiak
Peter A. Emmi
Scott A. Leslie
Allison M. Snyder
CRAVATH, SWAINE & MOORE LLP
Worldwide Plaza
825 Eighth Avenue
New York, NY 10019
rstark@cravath.com
aharasymiak@cravath.com
pemmi@cravath.com
sleslie@cravath.com
asnyder@cravath.com
Attorneys for Plaintiff Xerox Corporation

/s/ Matthew D. Cannon
Matthew D. Cannon

Second Supplemental Chart B-1

Claim chart of U.S. Patent Application Publication No. 2002/0147738 to Reader ("Method and Apparatus for Finding Patent-Relevant Web Documents")

as prior art to

U.S. Patent No. 6,778,979

This chart is based upon Xerox's apparent construction of the claims, and is not an admission that those constructions are correct or appropriate.

'979 Patent	READER
Claim 1	
1. A method for automatically generating a query from selected document content, comprising:	READER at para. 1: "Patent professionals often search for publications relevant to patents. Searches typically arise in two contexts: when looking for "prior art" publications that might invalidate a patent and when looking for publications that might disclose an infringement of a patent." READER at para. 2: ("An ever-increasing number of publications are being published on the Internet, for example, "white papers" published on companies' public websites. Thus, the Internet has become a more and more important resource for patent professionals looking for publications relevant to patents." READER at para 2: "However, patent professionals have for the most part relied on general Internet search techniques, such as applying keywords to general-purpose Internet search engines, to discover patent-relevant publications on the Internet." READER at para 4: "The present invention provides a highly automated search technique for discovering patent-relevant publications on the Internet. The high level of automation may be achieved with the expedient of a search client resident on an end-user station that initiates linked searches for patent data and Internet publication data in a manner transparent to a user. From the user's perspective, a patent-identifying attribute, such as an inventor name, assignee name or patent number, input on an end-user station automatically returns Internet publication data, such as Uniform Resource Locators (URLs) of Web documents. The invention thereby allows a user to find patent-relevant publications on the Internet by merely inputting a patent-identifying attribute. A patent-identifying attribute may be a patent family-identifying attribute, such as an inventor name or assignee name. Or a patent identifying-attribute may be a single patent-identifying attribute, such as a patent number. Or a patent identifying-attribute may be a patent claim-

'979 Patent	READER
	identifying attribute, such as a patent claim number. A basic method for finding patent-relevant documents published on the Internet in accordance with the present invention comprises the steps of: inputting a patent-identifying attribute on an end-user station; identifying patent data from the patent-identifying attribute; identifying Internet publication data from the patent data; and outputting the Internet publication data on the end-user station." *See also* Reader at para. 0015, Fig. 4. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Black, Abstract, 1:64-67, 2:43, 1:47-48, Fig. 1, 4:18-20, 1:50-54, 2:10-14, 2:57-61, 2:42-45, 2:26-29, 3:29-35, 3:37-45, 3:55-57 Wieser, p. 5, lines 8-12, p. 2 lines 16-19 Rhodes, 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein, p. 406 Horowitz, 2:42-51, 3:20-23 Fleming, Abstract, Fig. 2 Ford, 5:38, 5:30-39 Kraft, Abstract Apte, 6:60-67 Henkin, 2:42-49, 44:8-19 Pretschner, p. 1
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	READER at para. 0014: "Patent server 330 has patent database 332 and website database 334 resident thereon. Patent database 332 has entries stored thereon associating patent-identifying attributes, such as inventor names, assignee names and patent numbers, with patent classifications and patent language, such as patent claim text. Entries may include full-text patents. Website database 334 has entries stored thereon associating patent classifications with company website identifiers, such as URLs of company home pages." READER at para. 0015: "The patent classification may be a U.S. or

'979 Patent	READER
	international patent classification." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Mase, Abstract, Fig. 1, p. 377-378 Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55 Donaldson, 15:38-46, 17:17-19, 19:51-56, Fig. 8c HyPursuit, Fig. 3, p. 181, 184 Finkelstein, p. 410 Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40 Apte, 9:33-45 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19 Pretschner, pp. 1-2
automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	READER at para. 0013: "Abstraction of Web document-identifying attributes from the patent language search result may be accomplished by any of numerous algorithms well known in the art. Abstraction may involve, for example, reduction of a full-text patent claim to keywords separated by Boolean operators, which keywords and operators may be selected taking into account the syntactic and lexico-semantic interdependency of the words (i.e,. context) of the full-text claim." READER at para. 0015: "Search client 314 extracts a company website identifier from the CW [Company Website] search result

'979 Patent	READER
	and abstracts Web document-identifying (WDI) attributes from the patent language portion of the PC-PL search result (435)" To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Wieser, p. 14, line 19 to p. 15 line 5 Rhodes, 13:19-34, 1:56-2:6, 2:15-54 Black, 2:26-29, 2:42-45, 3:29-35 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7 Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Ford, 11:4-11, 12:52-63 Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B OracleText White Paper, p. 8, 9 Pretschner, p. 2
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	READER at para. 0015: "A user of end-user station 310 inputs at least one patent-identifying (PI) attribute on user interface 312 (405). Search client 314 forms a patent-identifying search query using the one or more patent-identifying attributes (410). In this regard, search client 314 forms a search query targeted, when applied to patent database 332, to retrieve a patent classification/patent language search result that includes pairs of patent classifications and patent language from one or more patents relevant to the one or more patent-identifying attributes. The patent classification may be a U.S. or international patent classification. The patent-identifying search query is transmitted via network interface 316 and network 320 from end-user station 310 to patent server 330. Patent server 330 applies the patent-identifying search query to patent database 332 to generate patent classification/patent language (PC-PL) search result (415). Patent server 330 transmits the patent classification/patent

'979 Patent	READER
	language search result to end-user station 310. End-user station 310, particularly search client 314, extracts a patent classification (PC) attribute from the patent classification portion of the PC-PL search result (420) and forms a company website-identifying (CWI) search query using the patent classification attribute (425)." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Mase, p. 377-379, Fig. 1 Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 5:12-28 Black, 2:10-18, 2:57-61 Donaldson, Figs. 9b and 9c, 18:47-63, 20:3-18, 20:19-29 HyPursuit, p. 181, 182, 185, 186, 191 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17 Kraft, 11:16-40 Apte, 9:33-37 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information	READER at para. 0015: "Patent server 330 transmits the patent classification/patent language search result to end-user station 310. End-user station 310, particularly search client 314, extracts a patent classification attribute (PC) attribute from the patent classification portion of the PC-PL search result (420) and forms a company website-identifying (CWI) search query using the patent

'979 Patent	READER
retrieval system identified by the assigned classification label.	classification attribute (425). In this regard, end-user station 310 forms a search query targeted, when applied on patent server 330, to retrieve a company website search result that includes one or more company website identifiers, such as URLs of company home pages, relevant to the patent classification attribute. End-user station 310 transmits the CWI search query to patent server 330. Patent server 330 applies the CWI search query to website database 334 to generate company website (CW) search result (430). The CW search result is transmitted to end-user station 310. Search client 314 extracts a company website identifier from the CW search result and abstracts Web document-identifying (WDI) attributes from the patent language portion of the PC-PL search result (435). Search client 314 passes the company website identifier and WDI attributes to search agent 318 (440). Using the company website identifier and well known DNS addressing, search agent 318 contacts the appropriate one of Web hosts 340 and, using well known "Web crawler" techniques, searches the totality of full-text documents published on the associated company website for Web document language relevant to the WDI attributes (445). Upon completion of the search, search agent 318 generates a Web document (WD) search result including Web document identifiers, such as URLs, of the relevant Web documents (450). Search agent 318 passes the Web document search result to search client 314 (455). Search client 314 extracts Web document identifiers from the Web document search result (460) and outputs the Web document identifiers on user interface 312." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Rhodes, 5:12-28 Black, 3:37-45, 3:55-57 Donaldson, 18:47-54, Figs. 9b and 9c, 18:55-63, 20:3-29 HyPursuit, p. 181, 182, 184, 185, 186 Finkelstein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8

'979 Patent	READER
	Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17 Kraft, 13:40-50, 13:64-67, 14:43-52 Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44 Pretschner, pp. 1-2
Claim 5	
5. The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	READER at para. 0015: "Search client 314 forms a patent-identifying search query using the one or more patent-identifying attributes (410). In this regard, search client 314 forms a search query targeted, when applied to patent database 332, to retrieve a patent classification/patent language search result that includes pairs of patent classifications and patent language from one or more patents relevant to the one or more patent-identifying attributes. The patent classification may be a U.S. or international patent classification." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Wieser, p. 18, line 24 to p. 19, line 1; p. 22, lines 3-15; Fig. 15 Rhodes, 11:62-12:4, 12:41-49, 13:1-8, 13:15-19 Donaldson, 15:38-46, 17:17-19, Fig. 8c HyPursuit, p. 180, col. 2; p. 184, col. 1; Fig. 3 Horowitz, 5:65-6:4 Henkin, 3:1-3, 14:4-9, 14:25-36, 27:30-34, Fig. 3, Fig. 23 Finkelstein, p. 410 OracleText White Paper, p. 19 Pretschner, p. 2
Claim 10	
10. The method according to	READER at para. 0015: "The patent classification may be a U.S. or

'979 Patent	READER
claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary.	international patent classification" To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Wieser, p. 18, line 24 to p. 19, line 1; p. 22, lines 3-15; Fig. 15 Rhodes, 4:45-55 Mase, p. 377, col. 2; p. 379, col. 1 Donaldson, 16:5-9 HyPursuit, p. 185, col. 1 Horowitz, 3:37-39, 5:50-55 Fleming, 6:45-60, 12:3-11, Fig. 9 OracleText White Paper, p. 19 Pretschner, p. 2
Claim 18	
18. An article of manufacture for use in a computer system, comprising: a memory;	READER at para. 0014. *See* claim 1 above. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Mase, p. 382 col. 1. Wieser, p. 4, line 26 to p. 5 line 1 Rhodes, Fig. 1, 9:18-34 Black, 4:58-67, 1:50-57, 4:30-43. Donaldson, Fig. 4, 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189

'979 Patent	READER
	Finkelstein, p. 406 Horowitz, 5:25-30, Figs. 2, 9 Fleming, Fig. 1, 4:34-41 Ford, 3:34-39, Fig. 3 Kraft, 1:23-27, Fig. 1 Apte, Figs. 3, 4, claim 33 Henkin, 45:37-38, Fig. 22 Pretschner, p. 7
instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	READER at para. 0014. *See* claim 1 above. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 5, lines 8-12, p. 2 lines 16-19; p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22; p. 14, line 19 to p. 15 line 5 Rhodes, 9:18-34, 10:1-17 Black, 4:58-67, 1L50-57 Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 col. 2. Finkelstein, p. 406 Horowitz, 5:30-33, Fig. 2 at 103 Fleming, 4:34-41

'979 Patent	READER
	Ford, 13:10-19 Kraft, 1:23-27 Henkin, 45:46-50, Fig. 22 Pretschner, pp. 1, 7
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	*See* claim 1[a] above. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Mase, p. 382 col. 1. Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55 Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 Finkelstein, p. 410 Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40 Apte, 6:60-67 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19 Pretschner, pp. 1-2

'979 Patent	READER
automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	*See* claim 1[b] above. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Wieser, p. 14, line 19 to p. 15 line 5 Rhodes, 13:19-34, 1:56-2:6, 2:15-54 Black, 1:50-57 and 4:58-67 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7 Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Ford, 11:4-11, 12:52-63 Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B OracleText White Paper, p. 8, 9 Pretschner, p. 2
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	*See* claim 1[c] above. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014, 0015 Mase, para. 382 col. 1. Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 5:12-28

'979 Patent	READER
	Donaldson, 7:22-32, 7:34-36, 13:62-14:6 HyPursuit, p. 189 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17 Kraft, 11:16-40 Apte, 9:33-37 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	*See* claim 1[d] above. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014, 0015, Fig. 4 Mase, p. 382, col. 1 Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Rhodes, 5:12-28 Donaldson, 7:22-32, 7:34-36, 13:62-14:6. HyPursuit, p. 189 Finkestein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8

'979 Patent	READER
	Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17 Kraft, 13:40-50, 13:64-67, 14:43-52 Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44 Pretschner, p. 1-2

Second Supplemental Chart B-2

Claim chart of PCT Application Pub. No. WO 01/44992 to Wieser et al. ("Context Matching System and Method")

as prior art to

U.S. Patent No. 6,778,979

This chart is based upon Xerox's apparent construction of the claims, and is not an admission that those constructions are correct or appropriate.

'979 Patent	WIESER
Claim 1	
1. A method for automatically generating a query from selected document content, comprising:	WIESER, p. 5, lines 8-12 (the client 12 generates a query composed of all or a portion of a document (e.g., a web page) and sends the query to a match server 14). WIESER, p. 2, lines 16-19: "Another object of the present invention is to provide a system and method which automatically and contextually matches products, advertisements or other content (hereinafter referred to as 'offers') to the content on a web page that a user has selected in real-time" To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0001, 0002, 0004, 0015, Fig. 4. Black, Abstract, 1:64-67, 2:43, 1:47-48, Fig. 1, 4:18-20, 1:50-54, 2:10-14, 2:57-61, 2:42-45, 2:26-29, 3:29-35, 3:37-45, 3:55-57 Rhodes, 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein, p. 406 Horowitz, 2:42-51, 3:20-23 Fleming, Abstract, Fig. 2 Ford, 5:38, 5:30-39 Kraft, Abstract Apte, 6:60-67 Henkin, 2:42-49, 44:8-19

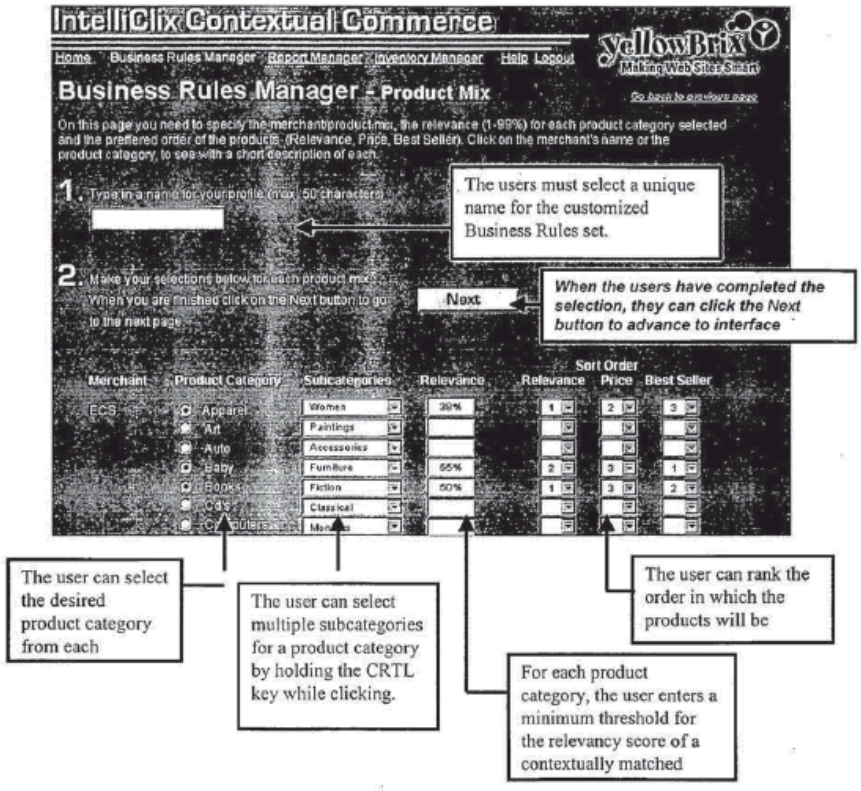
'979 Patent	WIESER
	Pretschner, p. 1
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	WIESER, p. 15, lines 8-11: "According to one embodiment, the present invention provides a novel approach to representing textual documents as high dimensional vectors. Such an approach provides an efficient means of indexing document collections, allowing retrieval of document (querying) based on keywords, **grouping related documents (categorization)**. Additionally, this method supports such contextual queries and document groupings" WIESER, p. 17, lines 15-32: "After the vector generations, the document vector, or feature vector, must be compared with database vectors, or feature vectors. A naïve approach to product matching would be to compare the document vector to the vectors for every product in the database. This process becomes burdensome as the number of products in the database grows. Fortunately, the database vectors are not smoothly distributed throughout the vector space, but rather, tend to "clump" together, leaving vast empty spaces between the clumps, or clusters as they are commonly known. For any given cluster, there is a sphere that bounds every point in the cluster. For all the various clusters, one could compute the center point and the radius of the sphere which bounds the cluster. Then, when one wishes to find the products that match a given document vector, one need only compare the document to the products in the cluster whose bounding sphere contains the document vector (or the nearest spheres, if no sphere contains the document vector). Thus, the computation is reduced from comparing the document vector to all product vectors to simply comparing the document vector to the center vectors for the spheres, followed by comparison to the product vectors for the products in the matching spheres." WIESER, p. 18, lines 16-22: "In accordance with an embodiment of the present invention, the query context vector is compared to the center vector of each cluster and the clusters with center vectors closest to the query context vector are selected. It is appreciated that these cluster IDs are then used to narrow the scope of products returned by the original metadata query as issued by the client 12. In other words, this narrowing qualification is added to the original metadata query to form an SQL query." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.:

'979 Patent	WIESER
	Reader, para. 0013, 0014, 0015 Mase, Abstract, Fig. 1, p. 377-378 Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55 Donaldson, 15:38-46, 17:17-19, 19:51-56, Fig. 8c HyPursuit, Fig. 3, p. 181, 184 Finkelstein, p. 410 Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40 Apte, 9:33-45 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19 Pretschner, pp. 1-2
automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	WIESER, p. 14, line 19 to p. 15, line 5: "the contextual matching engine 210 is composed of three subsystems: (1) the contextual matching server 300 …The contextual matching server 300 generates a query context vector, or feature vector, using a vector generation algorithm. Generally, vector based generation algorithms have certain features in common: (1) they all characterize documents based on the presence of keywords; (2) they all associate vectors with these keywords; and (3) they all form document vectors by combining the vectors of the keywords present in the document" To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0015 Rhodes, 13:19-34, 1:56-2:6, 2:15-54

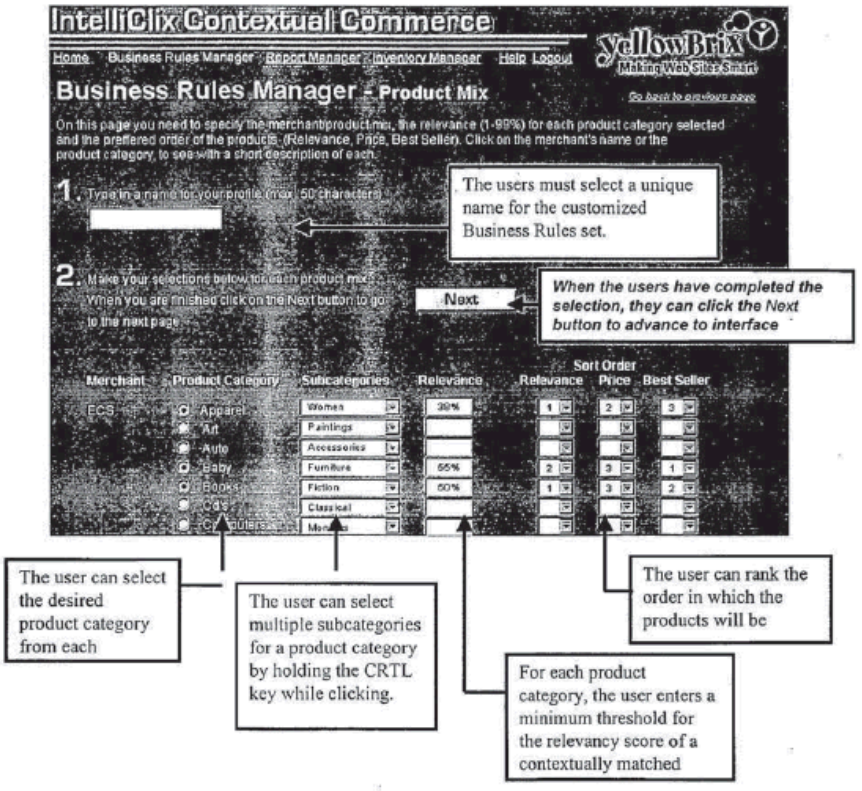
'979 Patent	WIESER
	Black, 2:26-29, 2:42-45, 3:29-35 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7 Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Ford, 11:4-11, 12:52-63 Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B OracleText White Paper, p. 8, 9 Pretschner, p. 2
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	WIESER, p. 17, lines 15-32: "After the vector generations, the document vector, or feature vector, must be compared with database vectors, or feature vectors. A naïve approach to product matching would be to compare the document vector to the vectors for every product in the database. This process becomes burdensome as the number of products in the database grows. Fortunately, the database vectors are not smoothly distributed throughout the vector space, but rather, tend to "clump" together, leaving vast empty spaces between the clumps, or clusters as they are commonly known. For any given cluster, there is a sphere that bounds every point in the cluster. For all the various clusters, one could compute the center point and the radius of the sphere which bounds the cluster. Then, when one wishes to find the products that match a given document vector, one need only compare the document to the products in the cluster whose bounding sphere contains the document vector (or the nearest spheres, if no sphere contains the document vector). Thus, the computation is reduced from comparing the document vector to all product vectors to simply comparing the document vector to the center vectors for the spheres, followed by comparison to the product vectors for the products in the matching spheres." WIESER, p. 18, lines 16-22: "In accordance with an embodiment of the present invention, the query context vector is compared to the center vector of each cluster and the clusters with center vectors closest to the query context vector are selected. It is appreciated that

'979 Patent	WIESER
	these cluster IDs are then used to narrow the scope of products returned by the original metadata query as issued by the client 12. In other words, this narrowing qualification is added to the original metadata query to form an SQL query. Preferably, the contextual matching server 300 returns N most relevant to the client or the E-commerce applet 12, along with their associated relevance" To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Mase, p. 377-379, Fig. 1 Rhodes, Table 2, 5:12-28 Black, 2:10-18, 2:57-61 Donaldson, Figs. 9b and 9c, 18:47-63, 20:3-18, 20:19-29 HyPursuit, p. 181, 182, 185, 186, 191 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17 Kraft, 11:16-40 Apte, 9:33-37 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information	WIESER, p. 17, lines 15-32: "After the vector generations, the document vector, or feature vector, must be compared with database vectors, or feature vectors. A naïve approach to product matching would be to compare the document vector to the vectors for every product in the database. This process becomes burdensome as the number of products in the database grows. Fortunately, the database

'979 Patent	WIESER
retrieval system identified by the assigned classification label.	vectors are not smoothly distributed throughout the vector space, but rather, tend to "clump" together, leaving vast empty spaces between the clumps, or clusters as they are commonly known. For any given cluster, there is a sphere that bounds every point in the cluster. For all the various clusters, one could compute the center point and the radius of the sphere which bounds the cluster. Then, when one wishes to find the products that match a given document vector, one need only compare the document to the products in the cluster whose bounding sphere contains the document vector (or the nearest spheres, if no sphere contains the document vector). Thus, the computation is reduced from comparing the document vector to all product vectors to simply comparing the document vector to the center vectors for the spheres, followed by comparison to the product vectors for the products in the matching spheres." WIESER, p. 18, lines 14-22: "The contextual matching server 300 then compares the query context vector to pre-determined item context vectors to narrow the search to focus on products that are most likely relevant to the selected text. In accordance with an embodiment of the present invention, the query context vector is compared to the center vector of each cluster and the clusters with center vectors closest to the query context vector are selected. It is appreciated that these cluster IDs are then used to narrow the scope of products returned by the original metadata query as issued by the client 12. In other words, this narrowing qualification is added to the original metadata query to form an SQL query." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015, Fig. 4. Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Rhodes, 5:12-28 Black, 3:37-45, 3:55-57 Donaldson, 18:47-54, Figs. 9b and 9c, 18:55-63, 20:3-29 HyPursuit, p. 181, 182, 184, 185, 186 Finkelstein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8

'979 Patent	WIESER
	Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17 Kraft, 13:40-50, 13:64-67, 14:43-52 Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44 Pretschner, pp. 1-2
Claim 5	
5. The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	WIESER, Fig. 15:  WIESER, p. 18, line 24 to p. 19, line 1: "Preferably, the user can, as an option, set the matching process to return results based on the default matching, or based upon user-defined business rules 240. If the user selects results based on general topic matching, then the system will employ the predefined (default) business rules 240. Whereas, if the user selects results based on business rules 240, then the system will pass the text to the vector generation subsystem 310. The vector generation subsystem 310 employs user-defined business

'979 Patent	WIESER
	rules 240 to return results that are filtered, i.e., biased towards the business rules or instructions 240. In accordance with an embodiment of the present invention, the business rules 240 are applied on the front end of the contextual analysis process to define the inventory sources to match to and the minimum relevancy score acceptable for contextual matches." WIESER, p. 22, lines 3-15: "For example, a screen shot of the Business Rules Manager for customizing business rules 240 shown in Fig. 15. … Each merchant will have the applicable Product Categories in a separate column. These categories are designed to be options so there are duplicate categories among merchants. Each category has a corresponding subcategory select box. The selection box allows the users to choose more than one subcategory for each corresponding category. After all of the desired subcategories are selected, the users can enter the Relevance score. This is a percentage that determines the minimum relevancy score of the product to the context of the content." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Rhodes, 11:62-12:4, 12:41-49, 13:1-8, 13:15-19 Donaldson, 15:38-46, 17:17-19, Fig. 8c HyPursuit, p. 180, col. 2; p. 184, col. 1; Fig. 3 Horowitz, 5:65-6:4 Henkin, 3:1-3, 14:4-9, 14:25-36, 27:30-34, Fig. 3, Fig. 23 Finkelstein, p. 410 OracleText White Paper, p. 19 Pretschner, p. 2
Claim 10	
10. The method according to claim 1, wherein each class in the organized classification of document content has associated	WIESER, Fig. 15:

'979 Patent	WIESER
therewith a characteristic vocabulary.	 WIESER, p. 18, line 24 to p. 19, line 1: "Preferably, the user can, as an option, set the matching process to return results based on the default matching, or based upon user-defined business rules 240. If the user selects results based on general topic matching, then the system will employ the predefined (default) business rules 240. Whereas, if the user selects results based on business rules 240, then the system will pass the text to the vector generation subsystem 310. The vector generation subsystem 310 employs user-defined business rules 240 to return results that are filtered, i.e., biased towards the business rules or instructions 240. In accordance with an embodiment of the present invention, the business rules 240 are applied on the front end of the contextual analysis process to define the inventory sources to match to and the minimum relevancy score acceptable for contextual matches." WIESER, p. 22, lines 3-15: "For example, a screen shot of the Business Rules Manager for customizing business rules 240 shown in Fig. 15. … Each merchant will have the applicable Product Categories in a separate column. These categories are designed to be options so there are duplicate categories among merchants. Each category has a corresponding subcategory select box. The selection box allows the users to choose more than one subcategory for each

'979 Patent	WIESER
	corresponding category. After all of the desired subcategories are selected, the users can enter the Relevance score. This is a percentage that determines the minimum relevancy score of the product to the context of the content." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Wieser, p. 18, line 24 to p. 19, line 1; p. 22, lines 3-15; Fig. 15 Rhodes, 4:45-55 Mase, p. 377, col. 2; p. 379, col. 1 Donaldson, 16:5-9 HyPursuit, p. 185, col. 1 Horowitz, 3:37-39, 5:50-55 Fleming, 6:45-60, 12:3-11, Fig. 9 OracleText White Paper, p. 19 Pretschner, p. 2
Claim 18	
18. An article of manufacture for use in a computer system, comprising: a memory;	WIESER, p. 4, line 26 to p. 5, line 1: describing system "that is readily implemented by presently available communication apparatus and electronic components. The invention finds ready application in virtually all commercial communications and/or computer networks including but not limited to world wide web (Internet), intranet, local area network (LAN), wide area network (WAN), wireless network and wired cable transmission systems." *See* claim 1 above. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014.

'979 Patent	WIESER
	Mase, p. 382 col. 1. Rhodes, Fig. 1, 9:18-34 Black, 4:58-67, 1:50-57, 4:30-43. Donaldson, Fig. 4, 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 Finkelstein, p. 406 Horowitz, 5:25-30, Figs. 2, 9 Fleming, Fig. 1, 4:34-41 Ford, 3:34-39, Fig. 3 Kraft, 1:23-27, Fig. 1 Apte, Figs. 3, 4, claim 33 Henkin, 45:37-38, Fig. 22 Pretschner, p. 7
instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	*See* claim 1 above. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Rhodes, 9:18-34, 10:1-17 Black, 4:58-67, 1L50-57 Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 col. 2. Finkelstein, p. 406

'979 Patent	WIESER
	Horowitz, 5:30-33, Fig. 2 at 103 Fleming, 4:34-41 Ford, 13:10-19 Kraft, 1:23-27 Henkin, 45:46-50, Fig. 22 Pretschner, pp. 1, 7
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	*See* claim 1[a] above. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Mase, p. 382 col. 1. Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55 Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 Finkelstein, p. 410 Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40 Apte, 6:60-67 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19

'979 Patent	WIESER
	Pretschner, pp. 1-2
automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	*See* claim 1[b] above. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Rhodes, 13:19-34, 1:56-2:6, 2:15-54 Black, 1:50-57 and 4:58-67 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7 Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Ford, 11:4-11, 12:52-63 Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B OracleText White Paper, p. 8, 9 Pretschner, p. 2
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	*See* claim 1[c] above. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014, 0015 Mase, para. 382 col. 1. Rhodes, Table 2, 5:12-28 Donaldson, 7:22-32, 7:34-36, 13:62-14:6

'979 Patent	WIESER
	HyPursuit, p. 189 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17 Kraft, 11:16-40 Apte, 9:33-37 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	*See* claim 1[d] above. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014, 0015, Fig. 4 Mase, p. 382, col. 1 Rhodes, 5:12-28 Donaldson, 7:22-32, 7:34-36, 13:62-14:6. HyPursuit, p. 189 Finkestein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8 Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17

'979 Patent	WIESER
	Kraft, 13:40-50, 13:64-67, 14:43-52 Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44 Pretschner, p. 1-2

Second Supplemental Chart B-3

Claim chart for U.S. Patent No. 6,236,768 to Rhodes et al. ("Method and Apparatus for Automated, Context-Dependent Retrieval of Information")

as prior art to

U.S. Patent No. 6,778,979

This chart is based upon Xerox's apparent construction of the claims, and is not an admission that those constructions are correct or appropriate.

'979 Patent	RHODES
Claim 1	
1. A method for automatically generating a query from selected document content, comprising:	RHODES, 1:56-2:6: "The RA works in two stages. First, the user's collection of text documents is indexed into a database saved in a vector format. These form the reservoir of documents from which later suggestions of relevance are drawn; that is, stored documents will later be "suggested" as being relevant to a document currently being edited or read. The store documents can be any sort of text document (notes, Usenet entries, webpages, e-mail, etc.). This indexing is usually performed automatically every night, and the index files are stored in a database. After the database is created, the other stage of the RA is run from Emacs, periodically taking a sample of text from the working buffer. The RA finds documents "similar" to the current sample according to word similarities; that is, the more times a word in the current sample is duplicated in a candidate database document, the greater will be assumed the relevance of that database document. The RA displays one-line summaries of the best few documents at the bottom of the Emacs window." RHODES, 10:42-51: "Analysis module 133 first indexes all the documents in a corpus of data (which, again, are stored as files mass storage device 106, which is assumed for explanatory purposes to be a hard disk), and writes indices to disk. Unlike the RA, the invention preferably keeps several vectors for each document. These include not only the wordvec vector for text (if any) in the document but also vectors for meta-information, e.g., subject, people, time, date, day of week, location, etc." RHODES, 12:53-57: "4. Determination of relevance For each element of each discrete vector in a query - the generation and vectorization of which is described below - the algorithm used by the RA may be used to determine relevance to documents in the corpus."

'979 Patent	RHODES
	RHODES, 13:1-8: "5. Weighted addition of vectors The result of the foregoing operations is a single similarity value for each type of meta-information. These values are associated with each document in the indexed corpus, and are used to compute the overall similarity using bias values for query and document types, by the following formula: Query biases = bq pq sq lq dq etc. (i.e., body_query_bias, person_query_bias, etc.)" RHODES, 13:15-19: "Each vector similarity is multiplied by its respective bias and the resulting biased similarity is summed, to produce an overall similarity between zero and one." RHODES, 13:42-47: "Analysis module 133 supplies a ranked list of the most relevant documents, which may be continually, intermittently, or upon request presented to the user over display 126. If desired, or upon user command, the list may be pruned to include only documents whose relevance level exceeds a predetermined threshold." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0001, 0002, 0004, 0015, Fig. 4. Black, Abstract, 1:64-67, 2:43, 1:47-48, Fig. 1, 4:18-20, 1:50-54, 2:10-14, 2:57-61, 2:42-45, 2:26-29, 3:29-35, 3:37-45, 3:55-57 Wieser, p. 5, lines 8-12, p. 2 lines 16-19 Finkelstein, p. 406 Horowitz, 2:42-51, 3:20-23 Fleming, Abstract, Fig. 2 Ford, 5:38, 5:30-39 Kraft, Abstract

‘979 Patent	RHODES
	Apte, 6:60-67 Henkin, 2:42-49, 44:8-19 Pretschner, p. 1
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	RHODES, Table 2: (int) (width*uns int) (int) (uns int) (uns int) (uns int) NUM_WORDS, WORDCODE-1, NUM_DOCS=N1, DOC-1, DOC-2, . . . , DOC-N1, WORDCODE-2, NUM_DOCS=N2, DOC-1, DOC-2, . . . , DOC-N2, etc. RHODES, 2:15-24: "Briefly, the concept behind the indexing scheme used in RA is that any given document may be represented by a multidimensional vector, each dimension or entry of which corresponds to a single word and is equal in magnitude to the number of times that word appears in the document. ... The advantages gained by this representation are relatively speedy disk retrieval, and an easily computed quantity indicating similarity between two documents: the dot product of their (normalized) vectors." RHODES, 4:20-27: "Experience with the RA has shown that actually performing a dot product with each indexed document is prohibitively slow for large databases. In preferred implementations, therefore, document vectors are not stored; instead, word vectors are stored. The "wordvec" file contains each word appearing in the entire indexed corpus of documents followed by a list of each document that contains that particular word." RHODES, 4:45-55: "Each word in the wordvec is represented by a unique numerical code, the "width" indicating the number of integers in the code (the RA uses two integers per code). The NUM_DOCS field indicates the number of documents containing the word specified by the associated wordcode. The word-count variables DOC-1, DOC-2,..., DOC-N1 each correspond to a document containing the word, and reflect the number of occurrences of the word divided by the total number of words in the document." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious.

'979 Patent	RHODES
	See, e.g.: Reader, para. 0013, 0014, 0015 Mase, Abstract, Fig. 1, p. 377-378 Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Donaldson, 15:38-46, 17:17-19, 19:51-56, Fig. 8c HyPursuit, Fig. 3, p. 181, 184 Finkelstein, p. 410 Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40 Apte, 9:33-45 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19 Pretschner, pp. 1-2
automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	RHODES, 13:19-34: "Analysis module 133 preferably generates queries autonomously from the current document in document buffer 140 or by reference to a current context. In the former case, analysis module 133 classifies the document either by its header or by reference to a template, and extracts the appropriate meta-information. In the latter case, the user's physical or interpersonal surroundings furnish the meta-information upon which the query is based. It is not necessary for the documents searched or identified to correspond in type to a current document. Furthermore, the query may not be limited to meta-information. Instead, the invention may utilize both a meta-information component (with relevance to candidate documents determined as discussed above) and a text component (with relevance determined in accordance with RA)."

‘979 Patent	RHODES
	RHODES, 1:56-2:6: "The RA works in two stages. First, the user's collection of text documents is indexed into a database saved in a vector format. These form the reservoir of documents from which later suggestions of relevance are drawn; that is, stored documents will later be "suggested" as being relevant to a document currently being edited or read. The store documents can be any sort of text document (notes, Usenet entries, webpages, e-mail, etc.). This indexing is usually performed automatically every night, and the index files are stored in a database. After the database is created, the other stage of the RA is run from Emacs, periodically taking a sample of text from the working buffer. The RA finds documents "similar" to the current sample according to word similarities; that is, the more times a word in the current sample is duplicated in a candidate database document, the greater will be assumed the relevance of that database document. The RA displays one-line summaries of the best few documents at the bottom of the Emacs window." RHODES, 2:15-54: "Briefly, the concept behind the indexing scheme used in RA is that any given document may be represented by a multidimensional vector, each dimension or entry of which corresponds to a single word and is equal in magnitude to the number of times that word appears in the document. ... RA creates vectors in three steps: ... Step 1: Remove stop words ... Step 2: Stem words ... Step 3: Make the document vector." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0015 Wieser, p. 14, line 19 to p. 15 line 5 Black, 2:26-29, 2:42-45, 3:29-35 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7

'979 Patent	RHODES
	Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Ford, 11:4-11, 12:52-63 Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B OracleText White Paper, p. 8, 9 Pretschner, p. 2
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	RHODES: (int) (width*uns int) (int) (uns int) (uns int) (uns int) NUM_WORDS, WORDCODE-1, NUM_DOCS=N1, DOC-1, DOC-2, . . . , DOC-N1, WORDCODE-2, NUM_DOCS=N2, DOC-1, DOC-2, . . . , DOC-N2, etc. RHODES, 5:12-28: "Accordingly, for each word in the query vector, the RA first looks up the word in the word offset file, and from that the word's entry is looked up in the wordvec file. An array of document similarities is used to maintain a running tally of documents and their similarities, in terms of numbers of word matches, to the query vector. The array is sorted by similarity, with the most similar documents at the top of the list. Similarity is computed for each word in the query vector by taking the product of the query-vector entry and the weight of each document in the corresponding wordvec file. To normalize this product, it is then divided by the query-vector magnitude (computed in the same manner as the document magnitude) and also by the document magnitude. The final value is added to the current running-total similarity for that document, and the process is repeated for the next word in the query. In summary, the query vector is analyzed wordcode by wordcode, with the similarities array indicating the relevance to the query of each document." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious.

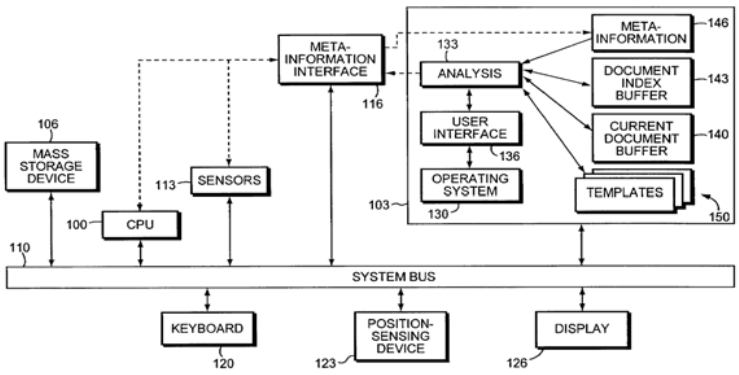
‘979 Patent	RHODES
	See, e.g.: Reader, para. 0015 Mase, p. 377-379, Fig. 1 Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Black, 2:10-18, 2:57-61 Donaldson, Figs. 9b and 9c, 18:47-63, 20:3-18, 20:19-29 HyPursuit, p. 181, 182, 185, 186, 191 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17 Kraft, 11:16-40 Apte, 9:33-37 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	RHODES, 5:12-28, "Accordingly, for each word in the query vector, the RA first looks up the word in the word offset file, and from that the word's entry is looked up in the wordvec file. An array of document similarities is used to maintain a running tally of documents and their similarities, in terms of numbers of word matches, to the query vector. The array is sorted by similarity, with the most similar documents at the top of the list. Similarity is computed for each word in the query vector by taking the product of the query-vector entry and the weight of each document in the corresponding wordvec file. To normalize this product, it is then divided by the query-vector magnitude (computed in the same manner as the document magnitude) and also by the document magnitude.

'979 Patent	RHODES
	The final value is added to the current running-total similarity for that document, and the process is repeated for the next word in the query. In summary, the query vector is analyzed wordcode by wordcode, with the similarities array indicating the relevance to the query of each document." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015, Fig. 4. Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Black, 3:37-45, 3:55-57 Donaldson, 18:47-54, Figs. 9b and 9c, 18:55-63, 20:3-29 HyPursuit, p. 181, 182, 184, 185, 186 Finkelstein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8 Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17 Kraft, 13:40-50, 13:64-67, 14:43-52 Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44 Pretschner, pp. 1-2
Claim 2	
2. The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding	RHODES, 13:19-34: "Analysis module 133 preferably generates queries autonomously from the current document in document buffer 140 or by reference to a current context. In the former case, analysis module 133 classifies the document

'979 Patent	RHODES
the set of entities in the selected document content.	either by its header or by reference to a template, and extracts the appropriate meta-information. In the latter case, the user's physical or interpersonal surroundings furnish the meta-information upon which the query is based. It is not necessary for the documents searched or identified to correspond in type to a current document. Furthermore, the query may not be limited to meta-information. Instead, he invention may utilize both a meta-information component (with relevance to candidate documents determined as discussed above) and a text component (with relevance determined in accordance with RA). RHODES, 11:10-31: "2. After the document is identified, different fields are extracted, again based on the template. For example, the email template continues: Delimiter {startline, "From"} Format {{anyorder {startline, "From: ", PERSON, ".backslash.n"} {startline, "Date: ", DATE, ".backslash.n"} optional {startline, "Subject: ", SUBJECT, ".backslash.n"}} ".backslash.n.backslash.n", BODY} } Bias 21100000 The delimiter command explicitly identifies the separator between one document of this template type and another, should they both reside in the same file. (For example, a plain e-mail archive may contain several pieces of mail in the same file, all separated by the word "From" plus a space at the start of a line.) The remainder of the template specifies that the "From:" line contains the person or people associated with this document. and the line starting with "Date:" contains the date/timestamp of the document." RHODES, 10:42-51: "Analysis module 133 first indexes all the documents in a corpus of data (which, again, are stored as files mass storage device 106, which is assumed for explanatory purposes to be a hard disk), and writes indices to disk. Unlike the RA, the invention preferably keeps several vectors for each document. These include not only the wordvec vector for text (if any) in the document but also vectors for metainformation, e.g., subject, people, time, date, day of week, location, etc.)" RHODES, 12:41-49: "Any meta-information that can be represented by text (e.g., subject lines, room names, people

'979 Patent	RHODES
	names, bodies of text, etc.) is encoded in accordance with the above scheme. Like the body of text, each word in these text strings is encoded separately and added to a vector. Vectors of discrete (text) data are all stored in one file, but the vectors are still conceptually distinct and are distinguished by their type bits. The file format for discrete type information is the same as the wordvec file format." RHODES, 13:1-8: "5. Weighted addition of vectors The result of the foregoing operations is a single similarity value for each type of meta-information. These values are associated with each document in the indexed corpus, and are used to compute the overall similarity using bias values for query and document types, by the following formula: Query biases = bq pq sq lq dq etc. (i.e., body_query_bias, person_query_bias, etc.)." RHODES, 13:15-19: "Each vector similarity is multiplied by its respective bias and the resulting biased similarity is summed, to produce an overall similarity between zero and one." RHODES, 13:42-47: "Analysis module 133 supplies a ranked list of the most relevant documents, which may be continually, intermittently, or upon request presented to the user over display 126. If desired, or upon user command, the list may be pruned to include only documents whose relevance level exceeds a predetermined threshold." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Black, 1:64-2:9, 3:37-45, 3:55-57 Donaldson, 18:47-63, 20:3-29, Fig. 9b, Fig. 9c HyPursuit, p. 186, col. 2 Horowitz, 7:19-25, 7:46-49 Kraft, 10:32-36, 12:13-21, 13:2-50 Henkin, 27:65-28:5 Stibel, 2:42-47, 9:21-43, 10:30-36, 11:56-66, 12:13-21, Table 2 Pretschner, p. 2
Claim 5	
5. The method according to claim 1, wherein the organized	RHODES, 12:41-49: "Any meta-information that can be represented by text (e.g., subject lines, room names, people

'979 Patent	RHODES
classification of document content is defined using a hierarchical organization.	names, bodies of text, etc.) is encoded in accordance with the above scheme. Like the body of text, each word in these text strings is encoded separately and added to a vector. Vectors of discrete (text) data are all stored in one file, but the vectors are still conceptually distinct and are distinguished by their type bits. The file format for discrete type information is the same as the wordvec file format." RHODES, 11:62-12:4: "In accordance with the present invention, each type of meta-information is placed in its own vector, and a single vector represents each type of meta-information supported by the invention. The final entry in the template file is the bias number for the particular type of file, which ranks the fields of the file in terms of importance. In the e-mail example above, the bias means that the body of the e-mail is most important, person and date fields are secondary (in a ratio of 2 to 1 to1), and no other fields are used to compute similarity." RHODES, 13:1-8: "5. Weighted addition of vectors The result of the foregoing operations is a single similarity value for each type of meta - information. These values are associated with each document in the indexed corpus, and are used to compute the overall similarity using bias values for query and document types, by the following formula: Query biases = bq pq sq lq dq etc. (i.e., body_query_bias, person_query_bias, etc.)." RHODES, 13:15-19: "Each vector similarity is multiplied by its respective bias and the resulting biased similarity is summed, to produce an overall similarity between zero and one." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Wieser, p. 18, line 24 to p. 19, line 1; p. 22, lines 3-15; Fig. 15 Donaldson, 15:38-46, 17:17-19, Fig. 8c HyPursuit, p. 180, col. 2; p. 184, col. 1; Fig. 3 Horowitz, 5:65-6:4 Henkin, 3:1-3, 14:4-9, 14:25-36, 27:30-34, Fig. 3, Fig. 23

‘979 Patent	RHODES
	Finkelstein, p. 410 OracleText White Paper, p. 19 Pretschner, p. 2
Claim 10	
10. The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary.	RHODES, 4:45-55: “Each word in the wordvec is represented by a unique numerical code, the “width” indicating the number of integers in the code (the RA uses two integers per code). The NUM_DOCS field indicates the number of documents containing the word specified by the associated wordcode. The word-count variables DOC‐1, DOC‐2, ⋯ , DOC‐N1 each correspond to a document containing the word, and reflect the number of occurrences of the word divided by the total number of words in the document.” To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Wieser, p. 18, line 24 to p. 19, line 1; p. 22, lines 3-15; Fig. 15 Mase, p. 377, col. 2; p. 379, col. 1 Donaldson, 16:5-9 HyPursuit, p. 185, col. 1 Horowitz, 3:37-39, 5:50-55 Fleming, 6:45-60, 12:3-11, Fig. 9 OracleText White Paper, p. 19 Pretschner, p. 2
Claim 18	
18. An article of manufacture for use in a computer system, comprising: a memory;	RHODES, Fig. 1:  RHODES, 9:18-34: "Refer now to FIG. 1, which illustrates, in block-diagram form, a hardware platform incorporating a representative, generalized embodiment of the invention. As

'979 Patent	RHODES
	indicated therein, the system includes a central-processing unit ("CPU") 100, which perform operations on and interacts with a main system memory 103 and components thereof. System memory 103 typically includes volatile or random- access memory ("RAM") for temporary storage of information, including buffers, executing programs, and portions of the computer's basic operating system. The platform typically also includes read-only memory ("ROM") for permanent storage of the computer's configuration and additional portions of the basic operating system, and at least one mass storage device 106, such a hard disk and/or CD-ROM drive. All components of the platform are interconnected by and communicate over, a bidirectional system bus 110." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 4, line 26 to p. 5 line 1 Black, 4:58-67, 1:50-57, 4:30-43. Donaldson, Fig. 4, 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 Finkelstein, p. 406 Horowitz, 5:25-30, Figs. 2, 9 Fleming, Fig. 1, 4:34-41 Ford, 3:34-39, Fig. 3 Kraft, 1:23-27, Fig. 1 Apte, Figs. 3, 4, claim 33 Henkin, 45:37-38, Fig. 22

'979 Patent	RHODES
	Pretschner, p. 7
instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	RHODES, 9:18-34: "The main system memory 103 stores instructions for operating a method for automatically generating a query from selected document content. Analysis module 133 of the main system memory 103 directs execution of the instructions to automatically generate a query from selected document content." RHODES, 10:1-17: "The main memory 103 contains a group of modules that control the operation of CPU 100 and its interaction with the other hardware components. These modules are implemented as executable machine instructions, running (by means of CPU 100) as active processes effectively capable of interacting (i.e., exchanging data and control commands) as illustrated. An operating system 130 directs the execution of low-level, basic system functions such as memory allocation, file management, and operation of mass storage devices 106. At a higher level, an analyzer module 133 directs execution of the primary functions performed by the invention, as discussed below; and instructions defining a user interface 136 allow straightforward interaction over display 126. User interface 136 generates words or graphical images on display 126 to facilitate user action and examination of documents, and accepts user commands from keyboard 120 and/or position-sensing device 123." *See* claim 1 above. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 5, lines 8-12, p. 2 lines 16-19; p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22; p. 14, line 19 to p. 15 line 5 Black, 4:58-67, 1L50-57

'979 Patent	RHODES
	Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 col. 2. Finkelstein, p. 406 Horowitz, 5:30-33, Fig. 2 at 103 Fleming, 4:34-41 Ford, 13:10-19 Kraft, 1:23-27 Henkin, 45:46-50, Fig. 22 Pretschner, pp. 1, 7
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	*See* claim 1[a] above. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Mase, p. 382 col. 1. Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 Finkelstein, p. 410 Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40

'979 Patent	RHODES
	Apte, 6:60-67 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19 Pretschner, pp. 1-2
automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	*See* claim 1[b] above. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Wieser, p. 14, line 19 to p. 15 line 5 Black, 1:50-57 and 4:58-67 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7 Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Ford, 11:4-11, 12:52-63 Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B OracleText White Paper, p. 8, 9 Pretschner, p. 2
automatically categorizing the	*See* claim 1[c] above.

'979 Patent	RHODES
selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014, 0015 Mase, para. 382 col. 1. Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Donaldson, 7:22-32, 7:34-36, 13:62-14:6 HyPursuit, p. 189 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17 Kraft, 11:16-40 Apte, 9:33-37 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	*See* claim 1[d] above. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014, 0015, Fig. 4

'979 Patent	RHODES
	Mase, p. 382, col. 1 Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Donaldson, 7:22-32, 7:34-36, 13:62-14:6. HyPursuit, p. 189 Finkestein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8 Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17 Kraft, 13:40-50, 13:64-67, 14:43-52 Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44 Pretschner, p. 1-2
Claim 19	
19. The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	*See* claim 2 above. RHODES, 10:1-17: "The main memory 103 contains a group of modules that control the operation of CPU 100 and its interaction with the other hardware components. These modules are implemented as executable machine instructions, running (by means of CPU 100) as active processes effectively capable of interacting (i.e., exchanging data and control commands) as illustrated. An operating system 130 directs the execution of low-level, basic system functions such as memory allocation, file management, and operation of mass storage devices 106. At a higher level, an analyzer module 133 directs execution of the primary functions performed by the invention, as discussed below; and instructions defining a user interface 136 allow straightforward interaction over display 126. User interface 136 generates words or graphical images on display 126 to facilitate user action and examination of documents, and accepts user commands from keyboard 120 and/or position-sensing device 123."

‘979 Patent	RHODES
	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Black 1:50-57, 4:58-67 Donaldson 7:22-32, 13:62-14:6 HyPursuit, p. 189, col. 2 Horowitz, 7:19-25, 7:46-49 Kraft, 10:32-36, 12:13-21, 13:2-50 Henkin, 3:1-3, 14:4-9, 14:25-36, 27:30-34, Fig. 3, Fig. 23 Stibel, 2:42-47, 9:21-43, 10:30-36, 11:56-66, 12:13-21, Table 2 Pretschner, p. 2

Second Supplemental Chart B-4

Claim chart for U.S. Patent Application Publication No. 2002/0147738 to Reader ("Method and Apparatus for Finding Patent-Relevant Web Documents") taken in view of "Experimental Simulation for Automatic Patent Categorization" by Mase (Advances in Production Management Systems, 1996)

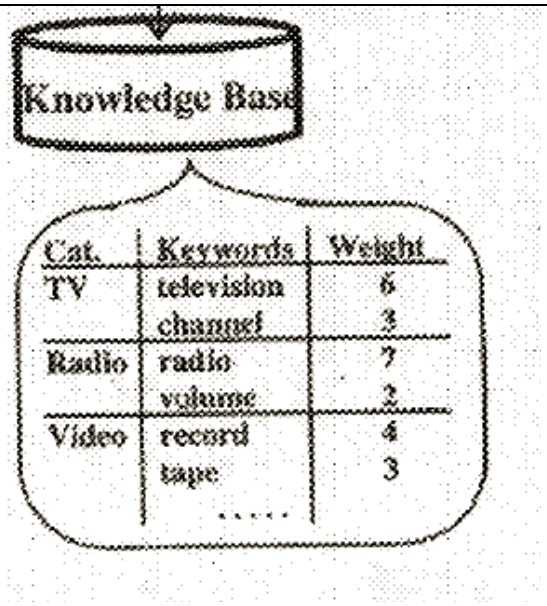
as prior art to

U.S. Patent No. 6,778,979

This chart is based upon Xerox's apparent construction of the claims, and is not an admission that those constructions are correct or appropriate.

'979 Patent	READER AND MASE
Claim 1	
1. A method for automatically generating a query from selected document content, comprising:	Reader, para. 0001: "Patent professionals often search for publications relevant to patents. Searches typically arise in two contexts: when looking for "prior art" publications that might invalidate a patent and when looking for publications that might disclose an infringement of a patent." Reader, para. 0002: "An ever-increasing number of publications are being published on the Internet, for example, "white papers" published on companies' public websites. Thus, the Internet has become a more and more important resource for patent professionals looking for publications relevant to patents." Reader, para. 0002: "However, patent professionals have for the most part relied on general Internet search techniques, such as applying keywords to general-purpose Internet search engines, to discover patent-relevant publications on the Internet." Reader, para. 0004: "The present invention provides a highly automated search technique for discovering patent-relevant publications on the Internet. The high level of automation may be achieved with the expedient of a search client resident on an end-user station that initiates linked searches for patent data and Internet publication data in a manner transparent to a user. From the user's perspective, a patent-identifying attribute, such as an inventor name, assignee name or patent number, input on an end-user station automatically returns Internet publication data, such as Uniform Resource Locators (URLs) of Web documents. The invention thereby allows a user to find patent-relevant publications on the Internet by merely inputting a patent-identifying attribute. A patent-identifying attribute may be a patent family-identifying attribute,

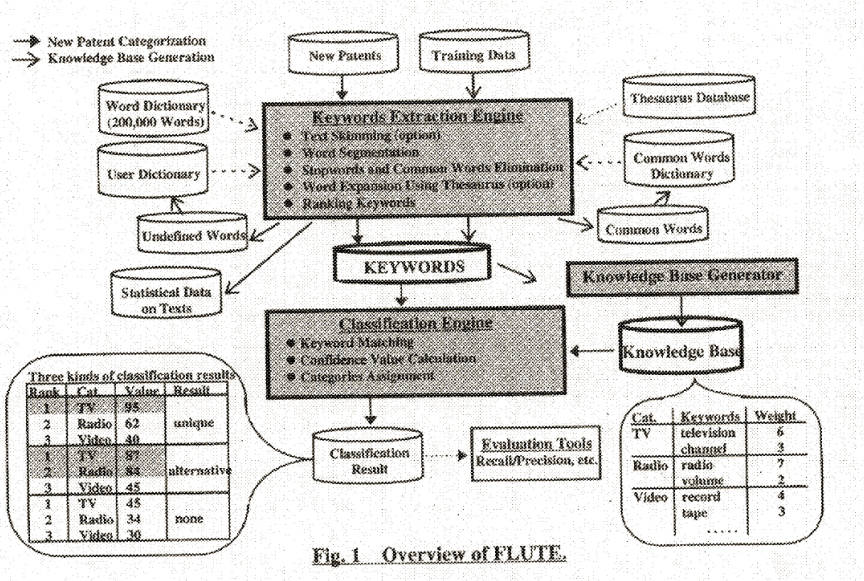
'979 Patent	READER AND MASE
	such as an inventor name or assignee name. Or a patent identifying-attribute may be a single patent-identifying attribute, such as a patent number. Or a patent identifying-attribute may be a patent claim-identifying attribute, such as a patent claim number. A basic method for finding patent-relevant documents published on the Internet in accordance with the present invention comprises the steps of: inputting a patent-identifying attribute on an end-user station; identifying patent data from the patent-identifying attribute; identifying Internet publication data from the patent data; and outputting the Internet publication data on the end-user station. *See also* Reader, para. 0015 and Fig. 4. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Black, Abstract, 1:64-67, 2:43, 1:47-48, Fig. 1, 4:18-20, 1:50-54, 2:10-14, 2:57-61, 2:42-45, 2:26-29, 3:29-35, 3:37-45, 3:55-57 Wieser, p. 5, lines 8-12, p. 2 lines 16-19 Rhodes, 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein, p. 406 Horowitz, 2:42-51, 3:20-23 Fleming, Abstract, Fig. 2 Ford, 5:38, 5:30-39 Kraft, Abstract Apte, 6:60-67 Henkin, 2:42-49, 44:8-19 Pretschner, p. 1
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a	Mase, Fig. 1 and p. 378:

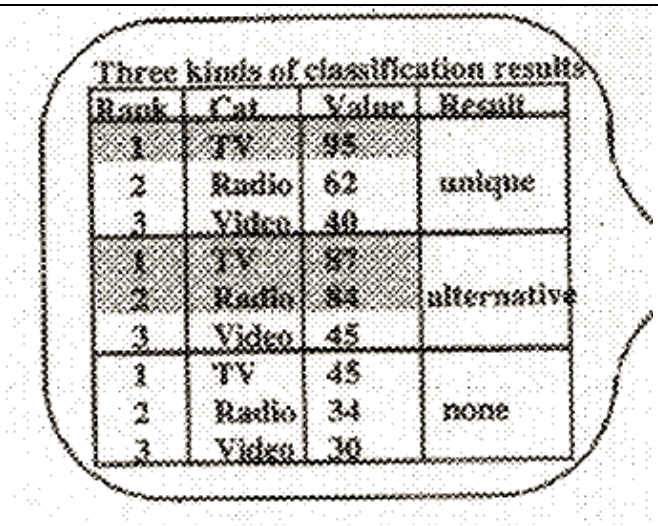
'979 Patent	READER AND MASE
classification label; each classification label corresponding to a category of information in an information retrieval system;	 Mase, Abstract: "This paper describes keywords-based patent categorization using our text classification support tool called FLUTE and discusses a simulation study applied to 154,000 patents. FLUTE automatically generates a classification knowledge base from sample patent texts. Experimental simulation results show that FLUTE is powerful enough to support the patent classification work of indexing experts." Mase, p. 378, col. 1: "(1) Automatic knowledge base initialization FLUTE is able to obtain word statistics from electronic documents. It first derives keyword candidates for each document. Then, it removes stop-words which are obviously not keywords ("thing", "is", etc.) and common words appearing over every category ("invention and "patent" appear in all patent documents). Next, it identifies keywords for each category by applying weights. Finally, it generates classification rules, which include a certain factor. Since the KB structure is simple, a person can verify and modify the knowledge base." Mase, p. 377 col. 2: "At present most patents are applied electronically. If an intelligent system could read a patent, recognize its purpose, and finally categorize it, the patent management process could be restructured. To classify patents into appropriate categories, the system would have to have a powerful knowledge base, which is difficult to build. We have been researching automatic patent categorization as a submitted research from Industrial Property Cooperation Center.

'979 Patent	READER AND MASE
	This paper presents our text classification support tool called FLUTE (Mase, et al. 1996a), the customization of FLUTE for patent categorization and its evaluation using 154,000 patents" Mase, p. 377 col. 1: "Obtaining patents is important for any organization that needs to maintain intellectual propery rights. In order to manage the review, search, and citation of patents at the patent office efficiently, patent applications should be appropriately indexed as soon as possible." Mase, p. 378: "(2) Classification certainty FLUTE presents three kinds of classification results according to values corresponding to the degree of confidence (certainty). If one category's certainty is high, FLUTE presents a unique solution. If the certainties of more than two categories are high, FLUTE offers alternatives. Otherwise FLUTE offers no solution and asks an expert to intervene." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55 Donaldson, 15:38-46, 17:17-19, 19:51-56, Fig. 8c HyPursuit, Fig. 3, p. 181, 184 Finkelstein, p. 410 Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40 Apte, 9:33-45 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55

'979 Patent	READER AND MASE
	OracleText White Paper, p. 11, 18, 19 Pretschner, pp. 1-2
automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	Reader, para. 0013: "Abstraction of Web document-identifying attributes from the patent language search result may be accomplished by any of numerous algorithms well known in the art. Abstraction may involve, for example, reduction of a full-text patent claim to keywords separated by Boolean operators, which keywords and operators may be selected taking into account the syntactic and lexico-semantic interdependency of the words (i.e,. context) of the full-text claim." Reader, para. 0015: "Search client 314 extracts a company website identifier from the CW [Company Website] search result and abstracts Web document-identifying (WDI) attributes from the patent language portion of the PC-PL search result (435)" To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Wieser, p. 14, line 19 to p. 15 line 5 Rhodes, 13:19-34, 1:56-2:6, 2:15-54 Black, 2:26-29, 2:42-45, 3:29-35 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7 Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Ford, 11:4-11, 12:52-63 Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B OracleText White Paper, p. 8, 9 Pretschner, p. 2
automatically categorizing the selected document content using	Mase, p. 377: "Obtaining patents is important for any organization that needs to maintain intellectual property rights. In order to

'979 Patent	READER AND MASE
the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	manage the review, search, and citation of patents at the patent office efficiently, patent applications should be appropriately indexed as soon as possible. However, there are problems, as follows: - There are approximately 3,000 patent categories. This makes it impossible for any indexing expert to perform categorization work over every technical field. – In Japan, there were 370,000 patent applications in 1994. Despite such a large number, it is difficult to increase the number of indexing experts. – Most patents include over 5,000 words. In Japan, patent categorization requires experts to read all documents, which is time consuming work. – To maintain an applicant's proprietary information, only persons entrusted with confidentiality can be assigned to indexing work. At present, most patents are applied electronically. If an intelligent system could read a patent, recognize its purpose, and finally categorize it, the patent management process could be restructured. To classify patents into appropriate categories, the system would have to have a powerful knowledge base, which is difficult to build." Mase, p. 377 col. 2: "We have been researching automatic patent categorization as a submitted research from Industrial Property Cooperation Center. This paper presents our text classification support tool called FLUTE (Mase, et al. 1996a), the customization of FLUTE for patent categorization, and its evaluation using 154,000 patents." Mase, Fig 1:

‘979 Patent	READER AND MASE
	 Fig. 1 Overview of FLUTE. Mase, p. 379 col. 1: “The classification engine calculates the similarity between inputted text and each category by comparing the keywords extracted from the inputted text with those stored in the KB.” Mase, p. 378: “(2) *Classification certainty* FLUTE presents three kinds of classification results according to values corresponding to the degree of confidence (certainty). If one category’s certainty is high, FLUTE presents a unique solution. If the certainties of more than two categories are high, FLUTE offers alternatives. Otherwise FLUTE offers no solution and asks an expert to intervene.” Mase, Fig. 1:

‘979 Patent	READER AND MASE
	 To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 5:12-28 Black, 2:10-18, 2:57-61 Donaldson, Figs. 9b and 9c, 18:47-63, 20:3-18, 20:19-29 HyPursuit, p. 181, 182, 185, 186, 191 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17 Kraft, 11:16-40 Apte, 9:33-37 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A

'979 Patent	READER AND MASE
	OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	Reader, para. 0015: "Patent server 330 transmits the patent classification/patent language search result to end-user station 310. End-user station 310, particularly search client 314, extracts a patent classification attribute (PC) attribute from the patent classification portion of the PC-PL search result (420) and forms a company website-identifying (CWI) search query using the patent classification attribute (425). In this regard, end-user station 310 forms a search query targeted, when applied on patent server 330, to retrieve a company website search result that includes one or more company website identifiers, such as URLs of company home pages, relevant to the patent classification attribute. Enduser station 310 transmits the CWI search query to patent server 330. Patent server 330 applies the CWI search query to website database 334 to generate company website (CW) search result (430). The CW search result is transmitted to end-user station 310. Search client 314 extracts a company website identifier from the CW search result and abstracts Web document-identifying (WDI) attributes from the patent language portion of the PC-PL search result (435). Search client 314 passes the company website identifier and WDI attributes to search agent 318 (440). Using the company website identifier and well known DNS addressing, search agent 318 contacts the appropriate one of Web hosts 340 and, using well known "Web crawler" techniques, searches the totality of full-text documents published on the associated company website for Web document language relevant to the WDI attributes (445). Upon completion of the search, search agent 318 generates a Web document (WD) search result including Web document identifiers, such as URLs, of the relevant Web documents (450). Search agent 318 passes the Web document search result to search client 314 (455). Search client 314 extracts Web document identifiers from the Web document search result (460) and outputs the Web document identifiers on user interface 312." *See also* Fig. 4. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Rhodes, 5:12-28

'979 Patent	READER AND MASE
	Black, 3:37-45, 3:55-57 Donaldson, 18:47-54, Figs. 9b and 9c, 18:55-63, 20:3-29 HyPursuit, p. 181, 182, 184, 185, 186 Finkelstein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8 Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17 Kraft, 13:40-50, 13:64-67, 14:43-52 Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44 Pretschner, pp. 1-2
Claim 5	
5. The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	Reader, para. 0015: "Search client 314 forms a patent-identifying search query using the one or more patent-identifying attributes (410). In this regard, search client 314 forms a search query targeted, when applied to patent database 332, to retrieve a patent classification/patent language search result that includes pairs of patent classifications and patent language from one or more patents relevant to the one or more patent-identifying attributes. The patent classification may be a U.S. or international patent classification." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Wieser, p. 18, line 24 to p. 19, line 1; p. 22, lines 3-15; Fig. 15 Rhodes, 11:62-12:4, 12:41-49, 13:1-8, 13:15-19 Donaldson, 15:38-46, 17:17-19, Fig. 8c HyPursuit, p. 180, col. 2; p. 184, col. 1; Fig. 3 Horowitz, 5:65-6:4

'979 Patent	READER AND MASE
	Henkin, 3:1-3, 14:4-9, 14:25-36, 27:30-34, Fig. 3, Fig. 23 Finkelstein, p. 410 OracleText White Paper, p. 19 Pretschner, p. 2
Claim 10	
10. The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary.	Reader, para. 0015: "The patent classification may be a U.S. or international patent classification." Mase, p. 377, col. 2: "FLUTE adopts a keywords-based classification approach, and is applicable to system building satisfying the following premises: • Categories are pre-defined and exclusive of each other. • Each category has characteristic keywords." Mase, p. 379, col. 1: "The classification engine calculates the similarity between inputted text and each category by comparing the keywords extracted from the inputted text with those stored in the KB." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Wieser, p. 18, line 24 to p. 19, line 1; p. 22, lines 3-15; Fig. 15 Rhodes, 4:45-55 Donaldson, 16:5-9 HyPursuit, p. 185, col. 1 Horowitz, 3:37-39, 5:50-55 Fleming, 6:45-60, 12:3-11, Fig. 9 OracleText White Paper, p. 19 Pretschner, p. 2
Claim 18	

'979 Patent	READER AND MASE
18. An article of manufacture for use in a computer system, comprising: a memory;	*See* claim 1 above. Reader, para. 0014. Mase, p. 382 col. 1. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Wieser, p. 4, line 26 to p. 5 line 1 Rhodes, Fig. 1, 9:18-34 Black, 4:58-67, 1:50-57, 4:30-43. Donaldson, Fig. 4, 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 Finkelstein, p. 406 Horowitz, 5:25-30, Figs. 2, 9 Fleming, Fig. 1, 4:34-41 Ford, 3:34-39, Fig. 3 Kraft, 1:23-27, Fig. 1 Apte, Figs. 3, 4, claim 33 Henkin, 45:37-38, Fig. 22 Pretschner, p. 7
instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	*See* claim 1 above. Reader, para. 0014. Mase, p. 382 col. 1. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.:

'979 Patent	READER AND MASE
	Wieser, p. 5, lines 8-12, p. 2 lines 16-19; p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22; p. 14, line 19 to p. 15 line 5 Rhodes, 9:18-34, 10:1-17 Black, 4:58-67, 1L50-57 Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 col. 2. Finkelstein, p. 406 Horowitz, 5:30-33, Fig. 2 at 103 Fleming, 4:34-41 Ford, 13:10-19 Kraft, 1:23-27 Henkin, 45:46-50, Fig. 22 Pretschner, pp. 1, 7
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	*See* claim 1[a] above. Reader, para. 0014. Mase, p. 382 col. 1. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55 Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 Finkelstein, p. 410 Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5

'979 Patent	READER AND MASE
	Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40 Apte, 6:60-67 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19 Pretschner, pp. 1-2
automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	*See* claim 1[b] above. Reader, para. 0014. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Wieser, p. 14, line 19 to p. 15 line 5 Rhodes, 13:19-34, 1:56-2:6, 2:15-54 Black, 1:50-57 and 4:58-67 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7 Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Ford, 11:4-11, 12:52-63 Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B

'979 Patent	READER AND MASE
	OracleText White Paper, p. 8, 9 Pretschner, p. 2
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	*See* claim 1[c] above. Reader, para. 0014. Mase, para. 382 col. 1. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 5:12-28 Donaldson, 7:22-32, 7:34-36, 13:62-14:6 HyPursuit, p. 189 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17 Kraft, 11:16-40 Apte, 9:33-37 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of	*See* claim 1[d] above. Reader, para. 0014. Mase, p. 382, col. 1

'979 Patent	READER AND MASE
information in the information retrieval system identified by the assigned classification label.	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015, Fig. 4 Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Rhodes, 5:12-28 Donaldson, 7:22-32, 7:34-36, 13:62-14:6. HyPursuit, p. 189 Finkestein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8 Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17 Kraft, 13:40-50, 13:64-67, 14:43-52 Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44 Pretschner, p. 1-2

Second Supplemental Chart B-5

Claim chart of U.S. Patent No. 6,546,386 to Black et al. ("Brilliant Query System") taken in view of U.S. Patent No. 7,225,180 to Donaldson et al. ("Filtering Search Results")

as prior art to

U.S. Patent No. 6,778,979

This chart is based upon Xerox's apparent construction of the claims, and is not an admission that those constructions are correct or appropriate.

'979 Patent	BLACK AND DONALDSON
Claim 1	
1. A method for automatically generating a query from selected document content, comprising:	Black, 1:64-67. Black , Abstract: "[a] system for conducting queries from any document displayed on any computer device." Black, 2:43, 1:47-48: "automatic ... statistical and empirical analysis" of a body of selected content that may be "text, such as magazine articles, news stories or any other text" where the text can be an online article. *See also id.* Fig. 1, 4:18-20. Black, 1:50-54: "Brilliant queries require a preparation process that analyzes any text to enhance and generate a set of suggested searches based on that analysis and certain pre-set user parameters." Black, 2:10-14: "The hook is the concept, primary subject matter or main topic for a body of text. The hook is used to define a query as narrowly as possible on a particular topic for a selected information source." Black, 2:57-61: "Automatic Generation of the Hook - One embodiment of the brilliant query to enable an automatic process for generating brilliant queries for a body of text, is to determine the hook by extracting the highest frequency proper names from the text body." Black, 2:42-45: "A brilliant query requires a list of keywords that are generated by automatic ... statistical and empirical analysis of the body of content to be enhanced or a comparable body of content." Black, 2:26-29: "Keywords are simply a collection of words, generated automatically..., that are deemed to be indicative of the topic matter or one of the topics for a given content selection." Black, 3:29-35: "Automatic Generation of Keywords - A word frequency analysis is done on all of the text, with stopwords excluded, and the resulting words, by order of frequency are compared to a pre-selected

'979 Patent	BLACK AND DONALDSON
	keyword list. Those that match, based upon a desired frequency become keywords to be combined with the hook to form focused, optimal queries." Black, 3:37-45: "Once the keywords have been selected and the hook for a body of text has been determined or automatically generated, the searches are created by generating a link for every keyword extracted from the body of text and combining it with the hook in a search that results in a result set that is the logical intersection of the results generated by the hook and the keyword. Basically, each entry in the list of search results must contain both the hook and the keyword and not just one or the other." Black, 3:55-57: "After the hook and the keywords have been established the query is conducted selecting one of the hook-keyword sets." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0001, 0002, 0004, 0015, Fig. 4. Wieser, p. 5, lines 8-12, p. 2 lines 16-19 Rhodes, 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein, p. 406 Horowitz, 2:42-51, 3:20-23 Fleming, Abstract, Fig. 2 Ford, 5:38, 5:30-39 Kraft, Abstract Apte, 6:60-67 Henkin, 2:42-49, 44:8-19 Pretschner, p. 1
defining an organized classification of document content with each class in the organized classification of document content having associated therewith	Donaldson, 15:38-46, 17:17-19, 19:51-56: disclosing categorizing Web pages/Web sites using a categorization scheme. Donaldson, 17:17-19, Fig. 8c: "Each category may include a listing of sub-categories 865 and web sites 875 within those categories."

'979 Patent	BLACK AND DONALDSON
a classification label; each classification label corresponding to a category of information in an information retrieval system;	Donaldson, 15:38-46: "For example, in one implementation, the hierarchy of category identifiers may include a hierarchy of category names, where groups of the category names are linked together in a hierarchical relationship. In this instance, names in the hierarchy represent categories, the names of which are linked together using sub-categories. The hierarchy of category identifiers also may include other related information, such as a list of web sites that are related to the category by name, description, or otherwise." Donaldson, 19:51-65: "Each electronic information store may contain content that has been classified and stored based on a specified type or types of classification criteria. For instance, the first electronic information store 992 may include content classified as non-offensive and the second electronic information store 994 may include content classified as offensive. Other types of content classification criteria may be implemented in addition to or separate from criteria based on offensive and non-offensive classifications. Other criteria that may be used, for example, include medical and non-medical, legal and non-legal, and sports and non-sports. In one implementation, the first electronic information includes contents relating to non-offensive web sites, and the second electronic information includes contents relating to offensive web sites." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Mase, Abstract, Fig. 1, p. 377-378 Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55 HyPursuit, Fig. 3, p. 181, 184 Finkelstein, p. 410 Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40

'979 Patent	BLACK AND DONALDSON
	Apte, 9:33-45 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19 Pretschner, pp. 1-2
automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	Black, 2:26-29: "Keywords are simply a collection of words, generated automatically … , that are deemed to be indicative of the topic matter or one of the topics for a given content selection." Black, 2:42-45: "A brilliant query requires a list of keywords that are generated by automatic ... statistical and empirical analysis of the body of content to be enhanced or a comparable body of content." Black, 3:29-35: "Automatic Generation of Keywords - A word frequency analysis is done on all of the text, with stopwords excluded, and the resulting words, by order of frequency are compared to a pre-selected keyword list. Those that match, based upon a desired frequency become keywords to be combined with the hook to form focused, optimal queries." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0015 Wieser, p. 14, line 19 to p. 15 line 5 Rhodes, 13:19-34, 1:56-2:6, 2:15-54 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7 Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Ford, 11:4-11, 12:52-63 Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B

'979 Patent	BLACK AND DONALDSON
	OracleText White Paper, p. 8, 9 Pretschner, p. 2
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	Black, 2:10-18: "The hook is the concept, primary subject matter or main topic for a body of text. The hook is used to define a query as narrowly as possible on a particular topic for a selected information source. To determine a "hook", a content layer must exist for which a context can be determined. There must be a perceivable structure to the information source and each content entry must have an associated context or place or places within the structure of the information source." Black, 2:57-61: "Automatic Generation of the Hook - One embodiment of the brilliant query to enable an automatic process for generating brilliant queries for a body of text, is to determine the hook by extracting the highest frequency proper names from the text body." Donaldson, Figs. 9b and 9c, 18:47-54: "Classifying the search term (922) generally includes classifying the received search term among one or more categories, with a first category and a second category being described and shown for illustrative purposes. If several search terms are grouped as a single string, the search terms may be collectively classified as a single string based on the grouping of search terms, or they may be classified individually based on each individual search term." Donaldson, 18:55-63: Comparing the search terms (step 924) generally includes comparing the search term to first electronic information within a first electronic information store when the search term is classified within the first category. By contrast, comparing the search term (step 926) generally includes comparing the search term to the second electronic information within the second electronic information store to determine whether matches exist when the search term is classified within the second category." Donaldson, 20:3-18: "The following describes an example applying the described search methods of FIG. 9b to this implementation. A user of a client system enters a search term (step 910). The search term is classified as either being offensive or non-offensive (step 922). If the term is classified as being non-offensive, then only the contents of the first electronic information store are searched (924) and results from the search are communicated for display to the user (step 930). In this example, the first electronic information store only contains contents that previously have been classified as non-offensive. If the search term entered by the user is classified as being offensive, the contents of either the second electronic information store or both the first and second electronic information stores are searched (step 926) and the results are communicated

'979 Patent	BLACK AND DONALDSON
	for display to the user (step 930)." Donaldson, 20:19-29: "The described filtering of results between offensive content and non-offensive content based on the classification of the search term may allow a web host to implement a parental type of control in determining what search results are displayed to the user. Because the offensive and non-offensive contents are stored in different electronic information stores, the ability to restrict access is enhanced. For instance, parental control can be exercised by blocking the access of a user to one or more electronic information stores. Other forms of data filtering also are enabled through this process and related techniques." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Mase, p. 377-379, Fig. 1 Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 5:12-28 HyPursuit, p. 181, 182, 185, 186, 191 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17 Kraft, 11:16-40 Apte, 9:33-37 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2
automatically formulating the query to restrict a	Black, 3:37-45: "Once the keywords have been selected and the hook for a body of text has been determined or automatically generated, the searches

'979 Patent	BLACK AND DONALDSON
search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	are created by generating a link for every keyword extracted from the body of text and combining it with the hook in a search that results in a result set that is the logical intersection of the results generated by the hook and the keyword. Basically, each entry in the list of search results must contain both the hook and the keyword and not just one or the other." Black, 3:55-57: "After the hook and the keywords have been established the query is conducted selecting one of the hook-keyword sets." Donaldson, 18:47-54: "Classifying the search term (922) generally includes classifying the received search term among one or more categories, with a first category and a second category being described and shown for illustrative purposes. If several search terms are grouped as a single string, the search terms may be collectively classified as a single string based on the grouping of search terms, or they may be classified individually based on each individual search term." Donaldson, Figs. 9b and 9c, 18:55-63: "Comparing the search terms (step 924) generally includes comparing the search term to first electronic information within a first electronic information store when the search term is classified within the first category. By contrast, comparing the search term (step 926) generally includes comparing the search term to the second electronic information within the second electronic information store to determine whether matches exist when the search term is classified within the second category." Donaldson, 20:3-18: "The following describes an example applying the described search methods of FIG. 9b to this implementation. A user of a client system enters a search term (step 910). The search term is classified as either being offensive or non-offensive (step 922). If the term is classified as being non-offensive, then only the contents of the first electronic information store are searched (924) and results from the search are communicated for display to the user (step 930). In this example, the first electronic information store only contains contents that previously have been classified as non-offensive. If the search term entered by the user is classified as being offensive, the contents of either the second electronic information store or both the first and second electronic information stores are searched (step 926) and the results are communicated for display to the user (step 930)." Donaldson, 20:19-29: "The described filtering of results between offensive content and non-offensive content based on the classification of the search term may allow a web host to implement a parental type of control in determining what search results are displayed to the user. Because the offensive and non-offensive contents are stored in different electronic information stores, the ability to restrict access is enhanced. For instance,

'979 Patent	BLACK AND DONALDSON
	parental control can be exercised by blocking the access of a user to one or more electronic information stores. Other forms of data filtering also are enabled through this process and related techniques." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015, Fig. 4. Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Rhodes, 5:12-28 HyPursuit, p. 181, 182, 184, 185, 186 Finkelstein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8 Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17 Kraft, 13:40-50, 13:64-67, 14:43-52 Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44 Pretschner, pp. 1-2
Claim 5	
5. The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	Donaldson, 17:17-19, Fig. 8c: "Each category may include a listing of sub-categories 865 and web sites 875 within those categories." Donaldson, 15:38-46: "("For example, in one implementation, the hierarchy of category identifiers may include a hierarchy of category names, where groups of the category names are linked together in a hierarchical relationship. In this instance, names in the hierarchy represent categories, the names of which are linked together using sub‐categories. The hierarchy of category identifiers also may include other related information, such as a list of web sites that are related to the category by name, description, or otherwise." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art

'979 Patent	BLACK AND DONALDSON
	renders this claim element obvious. See, e.g.: Reader, para. 0015 Wieser, p. 18, line 24 to p. 19, line 1; p. 22, lines 3-15; Fig. 15 Rhodes, 11:62-12:4, 12:41-49, 13:1-8, 13:15-19 HyPursuit, p. 180, col. 2; p. 184, col. 1; Fig. 3 Horowitz, 5:65-6:4 Henkin, 3:1-3, 14:4-9, 14:25-36, 27:30-34, Fig. 3, Fig. 23 Finkelstein, p. 410 OracleText White Paper, p. 19 Pretschner, p. 2
Claim 10	
10. The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary.	Donaldson, 16:5-9: "Comparing the search term with terms related to one or more categories to determine whether matches exist (step 830) may include using information related to categories, such as a name of a web site corresponding to a category, a description of the web site, or other related terms." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Wieser, p. 18, line 24 to p. 19, line 1; p. 22, lines 3-15; Fig. 15 Rhodes, 4:45-55 Mase, p. 377, col. 2; p. 379, col. 1 HyPursuit, p. 185, col. 1 Horowitz, 3:37-39, 5:50-55 Fleming, 6:45-60, 12:3-11, Fig. 9 OracleText White Paper, p. 19

'979 Patent	BLACK AND DONALDSON
	Pretschner, p. 2
Claim 18	
18. An article of manufacture for use in a computer system, comprising: a memory;	*See* claim 1 above. Black, 4:58-67, 1:50-57, 4:30-43. Donaldson, Fig. 4, 13:62-14:6, 6:53-7:10, 3:45-58 To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 4, line 26 to p. 5 line 1 Rhodes, Fig. 1, 9:18-34 HyPursuit, p. 189 Finkelstein, p. 406 Horowitz, 5:25-30, Figs. 2, 9 Fleming, Fig. 1, 4:34-41 Ford, 3:34-39, Fig. 3 Kraft, 1:23-27, Fig. 1 Apte, Figs. 3, 4, claim 33 Henkin, 45:37-38, Fig. 22 Pretschner, p. 7
instructions stored in the memory for operating a method for automatically generating a query from selected document content,	*See* claim 1 above. Black, 4:58-67, 1L50-57

'979 Patent	BLACK AND DONALDSON
comprising:	Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 5, lines 8-12, p. 2 lines 16-19; p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22; p. 14, line 19 to p. 15 line 5 Rhodes, 9:18-34, 10:1-17 HyPursuit, p. 189 col. 2. Finkelstein, p. 406 Horowitz, 5:30-33, Fig. 2 at 103 Fleming, 4:34-41 Ford, 13:10-19 Kraft, 1:23-27 Henkin, 45:46-50, Fig. 22 Pretschner, pp. 1, 7
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	*See* claim 1[a] above. Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Mase, p. 382 col. 1. Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55

'979 Patent	BLACK AND DONALDSON
	HyPursuit, p. 189 Finkelstein, p. 410 Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40 Apte, 6:60-67 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19 Pretschner, pp. 1-2
automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	*See* claim 1[b] above. Black, 1:50-57 and 4:58-67 To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Wieser, p. 14, line 19 to p. 15 line 5 Rhodes, 13:19-34, 1:56-2:6, 2:15-54 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7 Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Ford, 11:4-11, 12:52-63 Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32

'979 Patent	BLACK AND DONALDSON
	Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B OracleText White Paper, p. 8, 9 Pretschner, p. 2
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	*See* claim 1[c] above. Donaldson, 7:22-32, 7:34-36, 13:62-14:6 To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014, 0015 Mase, para. 382 col. 1. Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 5:12-28 HyPursuit, p. 189 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17 Kraft, 11:16-40 Apte, 9:33-37 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2
automatically formulating the query to restrict a search at the information	*See* claim 1[d] above. Black, 1:50-57, 4:58-67

'979 Patent	BLACK AND DONALDSON
retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	Donaldson, 7:22-32, 7:34-36, 13:62-14:6. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014, 0015, Fig. 4 Mase, p. 382, col. 1 Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Rhodes, 5:12-28 HyPursuit, p. 189 Finkestein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8 Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17 Kraft, 13:40-50, 13:64-67, 14:43-52 Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44 Pretschner, p. 1-2

Second Supplemental Chart B-6

Claim Chart of U.S. Patent No. 6,546,386 to Black et al. ("Brilliant Query System") taken in view of "HyPursuit: A Hierarchical Network Search Engine that Exploits Content-Link Hypertext Clustering" (Proceedings of 7th ACM conference on Hypertext, Copyright 1996)

as prior art to

U.S. Patent No. 6,778,979

This chart is based upon Xerox's apparent construction of the claims, and is not an admission that those constructions are correct or appropriate.

'979 Patent	BLACK AND HYPURSUIT
Claim 1	
1. A method for automatically generating a query from selected document content, comprising:	Black, 1:64-67. Black , Abstract: "[a] system for conducting queries from any document displayed on any computer device." Black, 2:43, 1:47-48: "automatic ... statistical and empirical analysis" of a body of selected content that may be "text, such as magazine articles, news stories or any other text" where the text can be an online article. *See also id.* Fig. 1, 4:18-20. Black, 1:50-54: "Brilliant queries require a preparation process that analyzes any text to enhance and generate a set of suggested searches based on that analysis and certain pre-set user parameters." Black, 2:10-14: "The hook is the concept, primary subject matter or main topic for a body of text. The hook is used to define a query as narrowly as possible on a particular topic for a selected information source." Black, 2:57-61: "Automatic Generation of the Hook - One embodiment of the brilliant query to enable an automatic process for generating brilliant queries for a body of text, is to determine the hook by extracting the highest frequency proper names from the text body." Black, 2:42-45: "A brilliant query requires a list of keywords that are generated by automatic ... statistical and empirical analysis of the body of content to be enhanced or a comparable body of content." Black, 2:26-29: "Keywords are simply a collection of words, generated automatically..., that are deemed to be indicative of the topic matter or one of the topics for a given content selection." Black, 3:29-35: "Automatic Generation of Keywords - A word frequency

'979 Patent	BLACK AND HYPURSUIT
	analysis is done on all of the text, with stopwords excluded, and the resulting words, by order of frequency are compared to a pre-selected keyword list. Those that match, based upon a desired frequency become keywords to be combined with the hook to form focused, optimal queries." Black, 3:37-45: "Once the keywords have been selected and the hook for a body of text has been determined or automatically generated, the searches are created by generating a link for every keyword extracted from the body of text and combining it with the hook in a search that results in a result set that is the logical intersection of the results generated by the hook and the keyword. Basically, each entry in the list of search results must contain both the hook and the keyword and not just one or the other." Black, 3:55-57: "After the hook and the keywords have been established the query is conducted selecting one of the hook-keyword sets." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0001, 0002, 0004, 0015, Fig. 4. Wieser, p. 5, lines 8-12, p. 2 lines 16-19 Rhodes, 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein, p. 406 Horowitz, 2:42-51, 3:20-23 Fleming, Abstract, Fig. 2 Ford, 5:38, 5:30-39 Kraft, Abstract Apte, 6:60-67 Henkin, 2:42-49, 44:8-19 Pretschner, p. 1
defining an organized classification of document content with each class in the organized classification	HyPursuit, Fig. 3. HyPursuit, p. 184 col. 2: "The HyPursuit prototype is a scalable system that uses content-link hypertext clustering, based on document contents and link

'979 Patent	BLACK AND HYPURSUIT
of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	information, to structure the information space and to support the entire range of search activities. Content-link clustering automatically computes sets of related documents called clusters. HyPursuit admits multiple coexisting cluster hierarchies based on different principles of grouping documents, such as the Library of Congress catalog scheme and institutional structures. These hierarchies may be constructed automatically or manually" (emphasis included) HyPursuit, p. 184 col. 1: "For example, documents can be clustered based on institutional boundaries or based on Library of Congress catalog subjects." HyPursuit, p. 181, col. 1: "Clusters also provide convenient units for the partitioning of work and resource allocation among the distributed components of the system. For example, a separate information server on a separate host may represent each individual cluster, performing operations on its local data." *See also* Fig. 3. HyPursuit, p. 184: "Each content router users its abstraction functions to compute a content label that summarizes its associated cluster." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Mase, Abstract, Fig. 1, p. 377-378 Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55 Donaldson, 15:38-46, 17:17-19, 19:51-56, Fig. 8c Finkelstein, p. 410 Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40

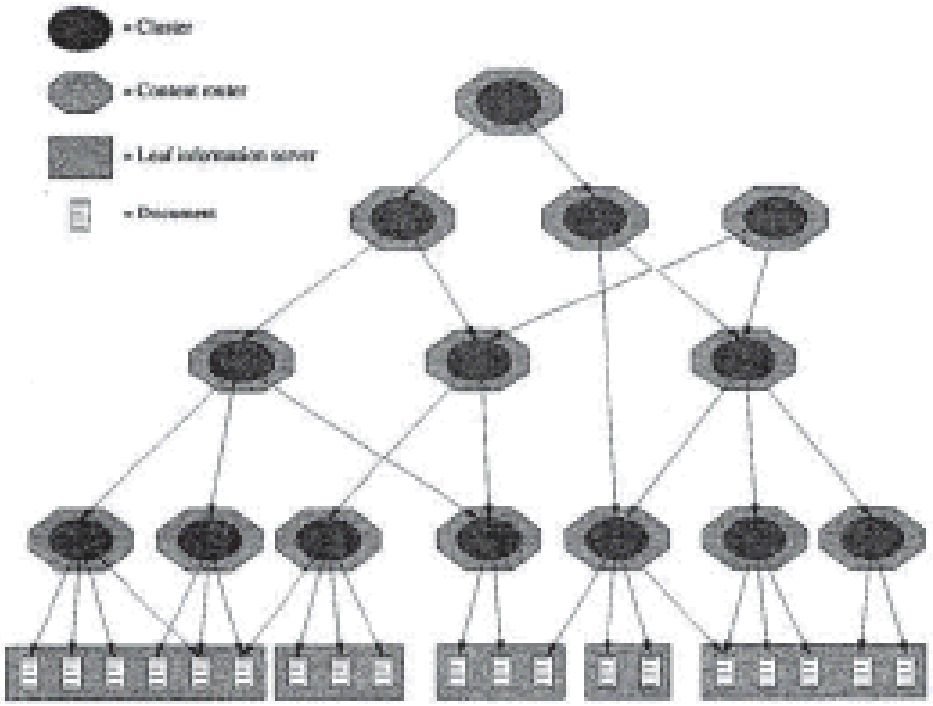
'979 Patent	BLACK AND HYPURSUIT
	Apte, 9:33-45 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19 Pretschner, pp. 1-2
automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	Black, 2:26-29: "Keywords are simply a collection of words, generated automatically … , that are deemed to be indicative of the topic matter or one of the topics for a given content selection." Black, 2:42-45: "A brilliant query requires a list of keywords that are generated by automatic ... statistical and empirical analysis of the body of content to be enhanced or a comparable body of content." Black, 3:29-35: "Automatic Generation of Keywords - A word frequency analysis is done on all of the text, with stopwords excluded, and the resulting words, by order of frequency are compared to a pre-selected keyword list. Those that match, based upon a desired frequency become keywords to be combined with the hook to form focused, optimal queries." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0015 Wieser, p. 14, line 19 to p. 15 line 5 Rhodes, 13:19-34, 1:56-2:6, 2:15-54 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7 Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Ford, 11:4-11, 12:52-63 Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B

'979 Patent	BLACK AND HYPURSUIT
	OracleText White Paper, p. 8, 9 Pretschner, p. 2
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	Black, 2:10-18: "The hook is the concept, primary subject matter or main topic for a body of text. The hook is used to define a query as narrowly as possible on a particular topic for a selected information source. To determine a "hook", a content layer must exist for which a context can be determined. There must be a perceivable structure to the information source and each content entry must have an associated context or place or places within the structure of the information source." Black, 2:57-61: "Automatic Generation of the Hook - One embodiment of the brilliant query to enable an automatic process for generating brilliant queries for a body of text, is to determine the hook by extracting the highest frequency proper names from the text body." HyPursuit, p. 182 col. 1: "To support a variety of query processing operations, HyPursuit uses query routing to identify relevant clusters, forward queries to the information servers for those clusters, and merge the results." HyPursuit, p. 186 col. 2: "HyPursuit uses query routing to support the search operations. Query routing uses the content labels stored in the content router to determine which of the child servers are likely to contain documents related to the user query. The query is then forwarded to these servers, and the results from each server are merged into a single result set. Documents returned by more than one child server are displayed only once." HyPursuit, p. 185 col. 1: "The abstraction function for query routing, on the other hand, computes a manageable set of terms that are used for identifying portions of the information space relevant to particular queries." HyPursuit, p. 181 col. 2: "To support operations like query processing in a scalable way, HyPursuit uses manageable summaries of cluster contents, called content labels, to approximate complete knowledge of the information space." HyPursuit, p. 191 col. 2: "To support scalable query processing, HyPursuit uses manageable summaries of cluster contents, called content labels, to approximate complete knowledge of the information space." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.:

'979 Patent	BLACK AND HYPURSUIT
	Reader, para. 0015 Mase, p. 377-379, Fig. 1 Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 5:12-28 Donaldson, Figs. 9b and 9c, 18:47-63, 20:3-18, 20:19-29 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17 Kraft, 11:16-40 Apte, 9:33-37 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	Black, 3:37-45: "Once the keywords have been selected and the hook for a body of text has been determined or automatically generated, the searches are created by generating a link for every keyword extracted from the body of text and combining it with the hook in a search that results in a result set that is the logical intersection of the results generated by the hook and the keyword. Basically, each entry in the list of search results must contain both the hook and the keyword and not just one or the other." Black, 3:55-57: "After the hook and the keywords have been established the query is conducted selecting one of the hook-keyword sets." HyPursuit, p. 182 col. 1:"To support a variety of query processing operations, HyPursuit uses query routing to identify relevant clusters, forward queries to the information servers for those clusters, and merge the results." HyPursuit, p. 186 col. 2: "HyPursuit uses query routing to support the search operations. Query routing uses the content labels stored in the

'979 Patent	BLACK AND HYPURSUIT
	content router to determine which of the child servers are likely to contain documents related to the user query. The query is then forwarded to these servers, and the results from each server are merged into a single result set. Documents returned by more than one child server are displayed only once." HyPursuit, p. 185 col. 1: "The abstraction function for query routing, on the other hand, computes a manageable set of terms that are used for identifying portions of the information space relevant to particular queries." HyPursuit, p. 184 col. 1: "To support operations like query processing in a scalable way, HyPursuit uses manageable summaries of cluster contents, called content labels, to approximate complete knowledge of the information space" HyPursuit, p. 181 col. 2: "To support scalable query processing, HyPursuit uses manageable summaries of cluster contents, called content labels, to approximate complete knowledge of the information space." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015, Fig. 4. Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Rhodes, 5:12-28 Donaldson, 18:47-54, Figs. 9b and 9c, 18:55-63, 20:3-29 Finkelstein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8 Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17 Kraft, 13:40-50, 13:64-67, 14:43-52 Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44

'979 Patent	BLACK AND HYPURSUIT
	Pretschner, pp. 1-2
Claim 2	
2. The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	Black, 1:64-2:9: "For example, an article on the SR-71 Blackbird Airplane might have the following brilliant queries: 1. Search for more information on BLACKBIRD and AVIATION 2. Search for more information on BLACKBIRD and ELINT 3. Search for more information on BLACKBIRD and RECONNAISSANCE 4. Search for more information on BLACKBIRD and TRANSPORT The book is BLACKBIRD and the keywords are AVIATION, ELINT (electronic intelligence) RECONNAISSANCE and TRANSPORT." HyPursuit, p. 186 col. 2: "*Query Refinement* HyPursuit uses term information about sub-clusters to dynamically compute recall- and precision-enhancing terms related to a user query. Figure 6 shows the interface of our system after an interaction with the search facilities to produce a result set and a subsequent query refinement operation. The region titled suggested terms in Figure 6 contains three scrollable lists of terms. A content router suggests query refinement terms using the subclusters in the content labels of its child servers. Collocated terms are the highest weighted terms from the subclusters that match the query. HyPursuit's term weights approximate conditional probabilities of term collocation. Term collocation in sub-clusters approximates term collocation in documents." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Rhodes, 10:42-51, 11:10-31, 12:41-49, 13:1-8, 13:15-34, 13:42-47 Donaldson, 18:47-63, 20:3-29, Fig. 9b, Fig. 9c Horowitz, 7:19-25, 7:46-49 Kraft, 10:32-36, 12:13-21, 13:2-50 Henkin, 27:65-28:5 Stibel, 2:42-47, 9:21-43, 10:30-36, 11:56-66, 12:13-21, Table 2

‘979 Patent	BLACK AND HYPURSUIT
	Pretschner, p. 2
Claim 5	
5. The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	HyPursuit, Fig. 3:  Figure 3: Content Routing as Cluster Hierarchies HyPursuit, p. 180 col. 2: “The HyPursuit prototype is a scalable system that uses *content-link* hypertext clustering, based on document contents and link information, to structure the information space and to support the entire range of search activities. Content-link clustering automatically computes sets of related documents called *clusters*. HyPursuit admits multiple coexisting cluster hierarchies based on different principles of grouping documents, such as the Library of Congress catalog scheme and institutional structures. These hierarchies may be constructed automatically or manually.” HyPursuit, p. 184 col. 1: “For example, documents can be clustered based on institutional boundaries or based on Library of Congress catalog subjects.” To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.:

'979 Patent	BLACK AND HYPURSUIT
	Reader, para. 0015 Wieser, p. 18, line 24 to p. 19, line 1; p. 22, lines 3-15; Fig. 15 Rhodes, 11:62-12:4, 12:41-49, 13:1-8, 13:15-19 Donaldson, 15:38-46, 17:17-19, Fig. 8c Horowitz, 5:65-6:4 Henkin, 3:1-3, 14:4-9, 14:25-36, 27:30-34, Fig. 3, Fig. 23 Finkelstein, p. 410 OracleText White Paper, p. 19 Pretschner, p. 2
Claim 10	
10. The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary.	HyPursuit, p. 185, col. 1: "The abstraction function for query routing, on the other hand, computes a manageable set of terms that are used for identifying portions of the information space relevant to particular queries. The abstraction function uses term and term frequency information in the children's content labels to compute term weights. The abstraction function then selects the most heavily weighted terms for generating the content router's content label. The abstraction function may also choose to add additional terms that characterize the information space but were not among the terms transmitted upon the hierarchy. For example, the abstraction function could add a term describing a poetry cluster as literature even though none of the poems mention literature explicitly." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Wieser, p. 18, line 24 to p. 19, line 1; p. 22, lines 3-15; Fig. 15 Rhodes, 4:45-55 Mase, p. 377, col. 2; p. 379, col. 1 Donaldson, 16:5-9

'979 Patent	BLACK AND HYPURSUIT
	Horowitz, 3:37-39, 5:50-55 Fleming, 6:45-60, 12:3-11, Fig. 9 OracleText White Paper, p. 19 Pretschner, p. 2
Claim 18	
18. An article of manufacture for use in a computer system, comprising: a memory;	*See* claim 1 above. Black, 4:58-67, 1:50-57, 4:30-43. HyPursuit, p. 189 col. 2. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 4, line 26 to p. 5 line 1 Rhodes, Fig. 1, 9:18-34 Donaldson, Fig. 4, 13:62-14:6, 6:53-7:10, 3:45-58 Finkelstein, p. 406 Horowitz, 5:25-30, Figs. 2, 9 Fleming, Fig. 1, 4:34-41 Ford, 3:34-39, Fig. 3 Kraft, 1:23-27, Fig. 1 Apte, Figs. 3, 4, claim 33 Henkin, 45:37-38, Fig. 22

'979 Patent	BLACK AND HYPURSUIT
	Pretschner, p. 7
instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	*See* claim 1 above. Black, 1:50-57, 4:58-67. HyPursuit, p. 189 col. 2. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 5, lines 8-12, p. 2 lines 16-19; p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22; p. 14, line 19 to p. 15 line 5 Rhodes, 9:18-34, 10:1-17 Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 Finkelstein, p. 406 Horowitz, 5:30-33, Fig. 2 at 103 Fleming, 4:34-41 Ford, 13:10-19 Kraft, 1:23-27 Henkin, 45:46-50, Fig. 22 Pretschner, pp. 1, 7
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category	*See* claim 1[a] above. HyPursuit, p. 189 col. 2. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015

'979 Patent	BLACK AND HYPURSUIT
of information in an information retrieval system;	Mase, p. 382 col. 1. Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55 Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 Finkelstein, p. 410 Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40 Apte, 6:60-67 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19 Pretschner, pp. 1-2
automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	*See* claim 1[b] above. Black, 1:50-57, 4:58-67. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Wieser, p. 14, line 19 to p. 15 line 5 Rhodes, 13:19-34, 1:56-2:6, 2:15-54 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7 Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B

'979 Patent	BLACK AND HYPURSUIT
	Ford, 11:4-11, 12:52-63 Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B OracleText White Paper, p. 8, 9 Pretschner, p. 2
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content;	*See* claim 1[c] above. HyPursuit, p. 189 col. 2. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014, 0015 Mase, para. 382 col. 1. Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 5:12-28 Donaldson, 7:22-32, 7:34-36, 13:62-14:6 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17 Kraft, 11:16-40 Apte, 9:33-37 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A

'979 Patent	BLACK AND HYPURSUIT
	OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	*See* claim 1[d] above. Black, 1:50-57, 4:58-67. HyPursuit, p. 189 col. 2. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014, 0015, Fig. 4 Mase, p. 382, col. 1 Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Rhodes, 5:12-28 Donaldson, 7:22-32, 7:34-36, 13:62-14:6. Finkestein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8 Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17 Kraft, 13:40-50, 13:64-67, 14:43-52 Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44 Pretschner, p. 1-2
Claim 19	
19. The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise	*See* claim 2 above. Black, 1:50-57, 4:58-67. HyPursuit, p. 189 col. 2.

'979 Patent	BLACK AND HYPURSUIT
limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Rhodes, 10:1-17 Donaldson 7:22-32, 13:62-14:6 Horowitz, 7:19-25, 7:46-49 Kraft, 10:32-36, 12:13-21, 13:2-50 Henkin, 3:1-3, 14:4-9, 14:25-36, 27:30-34, Fig. 3, Fig. 23 Stibel, 2:42-47, 9:21-43, 10:30-36, 11:56-66, 12:13-21, Table 2 Pretschner, p. 2

Second Supplemental Chart B-7

Claim chart of U.S. Patent No. 6,122,647 to Horowitz

as prior art to

U.S. Patent No. 6,778,979

This chart is based upon Xerox's apparent construction of the claims, and is not an admission that those constructions are correct or appropriate.

'979 Patent	**U.S. Patent No. 6,122,647**
Claim 1	
1. A method for automatically generating a query from selected document content, comprising:	"The present invention overcomes the limitations of conventional information retrieval systems generally, and conventional Internet and intranet search engines particularly, by providing a system and method that dynamically generates contextual hypertext links in a source document to other topically relevant documents in response to the content of the source document or user-selected portion thereof. These new links are contextual links because they are generated in specific response to the content of a selected portion of the source document." 2:42-51. "[T]he present invention can treat any document (or portion of a document) a user is viewing as an inquiry and create new links in the source document to other related documents." 3:20-23. See Fig. 3. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0001, 0002, 0004, 0015, Fig. 4. Black, Abstract, 1:64-67, 2:43, 1:47-48, Fig. 1, 4:18-20, 1:50-54, 2:10-14, 2:57-61, 2:42-45, 2:26-29, 3:29-35, 3:37-45, 3:55-57 Wieser, p. 5, lines 8-12, p. 2 lines 16-19

	Rhodes, 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein, p. 406 Fleming, Abstract, Fig. 2 Ford, 5:38, 5:30-39 Kraft, Abstract Apte, 6:60-67 Henkin, 2:42-49, 44:8-19 Pretschner, p. 1
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	"A knowledge base or other data repository stores information associating individual topics with sets of documents related to the topic, and with terms descriptive of the topic." 3:37-39. "The knowledge base 130 is a persistent data store that system 100 uses to store topic information. The knowledge base 130 comprises an arbitrary number of topics. Each topic is associated with one or more terms that are synonyms for each other. A term is a word or series of words (e.g., a noun phrase) that refer to a topic. A topic describes a possible subject annotation for documents in the document collection 140." 5:49-56. "Each topic in the knowledge base 130 may have a unique topic ID code for cross-referencing in other tables." 5:60-62. "[E]ach document (and document reference) has an association with at least one topic in the knowledge base 130, and preferably with many topics." 6:28-30. "Also, as a further embodiment, hierarchical topic menus may be created. Here, each menu item may be a topic in the knowledge base 130, with a submenu of related topics, and each related topic may have its own submenu of links to target documents. This hierarchical approach provides the user the ability to explore the entire document collection 140." 11:8-15. "The knowledge base 130 is a persistent data store that system 100 uses to store topic information. The knowledge base 130

	comprises an arbitrary number of topics. Each topic is associated with one or more terms that are synonyms for each other. A term is a word or series of words (e.g., a noun phrase) that refer to a topic. A topic describes a possible subject annotation for documents in the document collection 140." 5:49-56. See Fig. 5. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Mase, Abstract, Fig. 1, p. 377-378 Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55 Donaldson, 15:38-46, 17:17-19, 19:51-56, Fig. 8c HyPursuit, Fig. 3, p. 181, 184 Finkelstein, p. 410 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40 Apte, 9:33-45 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19 Pretschner, pp. 1-2
automatically identifying a set of entities in the selected document content for searching	"The tagging module 120 is responsible for analyzing a selected portion of a document, identifying a set of terms and topics that are relevant to the selected portion or about the

additional information related thereto using the information retrieval system;	selected portion, and generating tags in the document which associate the terms and topics. The tagging module 120 may apply a variety of linguistic analysis techniques to identify the relevant topics." 6:49-55. "From the selected portion 304 of the source document 300, a number of new contextual links to target documents 310 relevant to the selected portion 304 are generated 306. These new contextual links are associated with selected terms of the source document, typically, but not necessarily only, in the selected portion 304." 7:19-25. "The terms selected for the links may be any terms of the selected portion, or may be other terms not necessarily appearing in the selected portion 304, but associated with topics that are most relevant to the selected portion 304." 7:46-49. "Referring now to FIG. 7, there is shown a flowgraph of one embodiment of a process for selecting topics and generating tags to such topics, as preferably implemented by the tagging module 120. The input to the process is a selected portion of a source document." 8:40-43. "Referring to the figure, the tagging module 120 tokenizes 702 the selected portion, dividing the selected portion into words and/or word phrases, each of which constitutes a token." 8:50-53. "The tagging module 120 then parses 706 the tokens to recognize groups of grammatical noun phrases." 8:62-63. "The tagging module 120 then selects 712 terms from the set of unified terms, based on threshold parameters for the number of terms to be selected." 9:13-15. "For each selected term, the tagging module 120 looks up 716 the term in the knowledge base 130." 9:27-28. "The contextual links provide the user with access to target documents that are relevant to the selected portion. The contextual links are generated from a linguistic analysis of the selected portion which identifies particular terms or keywords that are relevant to or about the selected portion." 12:27-32. See Figs. 4c-4d, 5, 7. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.:

	Reader, para. 0013, 0015 Wieser, p. 14, line 19 to p. 15 line 5 Rhodes, 13:19-34, 1:56-2:6, 2:15-54 Black, 2:26-29, 2:42-45, 3:29-35 Finkestein, p. 410, 408 Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Ford, 11:4-11, 12:52-63 Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B OracleText White Paper, p. 8, 9 Pretschner, p. 2
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	"In one embodiment, the present invention provides a computer-implemented process in which a portion of text of a source document is analyzed and a number of topics are determined as being representative of what the selected portion is about. Topic analysis may be determined by various syntactic and semantic processes, such as identification and frequency analysis of terms of the selected portion. For each of the topics, a new tag is added to the source document. A tag includes a term, preferably from the text of the document, and a reference to the topic associated with the term. These tags are preferably stored with the source document." 3:25-35. "The user selects a portion 304 of the source document 300, which may be the entire source document 300, or any lesser portion of it, such as a selected set of words, a sentence, paragraph, or the like. The selected portion 304 is provided to the tagging module 120, which is coupled to the knowledge base 130." 7:59-65. "The tagging module 120 determines the topics in the knowledge base 130 that are about the selected portion 304. Preferably the tagging module 120 applies some type of linguistic analysis to the selected portion, including

	either syntactic or semantic analysis methods to determine the topics that are most representative or relevant to the selected portion 304." 7:66-8:5. See Figs. 6-8. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Mase, p. 377-379, Fig. 1 Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 5:12-28 Black, 2:10-18, 2:57-61 Donaldson, Figs. 9b and 9c, 18:47-63, 20:3-18, 20:19-29 HyPursuit, p. 181, 182, 185, 186, 191 Finkestein, p. 410, 408 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17 Kraft, 11:16-40 Apte, 9:33-37 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of	"The tagging module receives as input a user selected portion of a source document and determines the topics relevant to the user selected portion, and creates a set of tags, each tag associating a term of the selected portion to one or more topics in the knowledge base. The presentation module receives the

information in the information retrieval system identified by the assigned classification label.	document and its set of tags, determines from the knowledge base the topics associated with each tag, and the target documents associated with each topic. The presentation module adds links to these target documents to the source document, either as links in the document body where one or more of the keywords appears, or in a separate navigational component." 3:64-4:8. "The presentation module 150 is responsible for determining a set of target documents associated with the generated tags in a document. . . ." 6:56-58. "Preferably the tagging module 120 applies some type of linguistic analysis to the selected portion, including either syntactic or semantic analysis methods to determine the topics that are most representative or relevant to the selected portion 304. The tagging module 120 adds a tag to the source document 300 for each of these topics, the tag specifying the topic in the knowledge base 130 . . . ." 8:1-7. "The presentation module 150 receives a source document including the set of tags in the document, and creates new links in the source document to other target documents. The presentation module 150 uses the knowledge base 130 to access the topics in knowledge base 130 associated with the tags. The presentation module 150 uses the document collection 140 to obtain references to the target documents associated with these topics." 8:17-24. "The presentation module 150 receives the source document 300 and the set of tags created by the tagging module 120. For each tag (802), the presentation module 150 looks up 804 the topic(s) in the knowledge base 130 specified in the tag. For each such topic the presentation module 150 retrieves 806 the documents associated with the topic, or more particularly, retrieves the URL (or other specification of network location) for the document." 10:11-18. See Figs. 5-6, 8. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015, Fig. 4.

	Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Rhodes, 5:12-28 Black, 3:37-45, 3:55-57 Donaldson, 18:47-54, Figs. 9b and 9c, 18:55-63, 20:3-29 HyPursuit, p. 181, 182, 184, 185, 186 Finkelstein, p. 410 Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17 Kraft, 13:40-50, 13:64-67, 14:43-52 Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44 Pretschner, pp. 1-2
Claim 2	
2. The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	"From the selected portion 304 of the source document 300, a number of new contextual links to target documents 310 relevant to the selected portion 304 are generated 306. These new contextual links are associated with selected terms of the source document, typically, but not necessarily only, in the selected portion 304." 7:19-25. "The terms selected for the links may be any terms of the selected portion, or may be other terms not necessarily appearing in the selected portion 304, but associated with topics that are most relevant to the selected portion 304." 7:46-49. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Rhodes, 10:42-51, 11:10-31, 12:41-49, 13:1-8, 13:15-34,

	13:42-47 Black, 1:64-2:9, 3:37-45, 3:55-57 Donaldson, 18:47-63, 20:3-29, Fig. 9b, Fig. 9c HyPursuit, p. 186, col. 2 Kraft, 10:32-36, 12:13-21, 13:2-50 Henkin, 27:65-28:5 Stibel, 2:42-47, 9:21-43, 10:30-36, 11:56-66, 12:13-21, Table 2 Pretschner, p. 2
Claim 5	
5. The method according to claim I, wherein the organized classification of document content is defined using a hierarchical organization.	"The knowledge base 130 further contains a number of topics linked together in various hierarchical interconnected graphs by relations. A topic may have any number of relationships to other topics. Each relationship defines a semantic relationship between two topics, and has a predefined type. Relationship types include, for example, parent and child relationships (e.g., "is-a" or "type-of")." 5:65-6:4. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g. Reader, para. 0015 Wieser, p. 18, line 24 to p. 19, line 1; p. 22, lines 3-15; Fig. 15 Rhodes, 11:62-12:4, 12:41-49, 13:1-8, 13:15-19 Donaldson, 15:38-46, 17:17-19, Fig. 8c HyPursuit, p. 180, col. 2; p. 184, col. 1; Fig. 3 Henkin, 3:1-3, 14:4-9, 14:25-36, 27:30-34, Fig. 3, Fig. 23 Finkelstein, p. 410

	OracleText White Paper, p. 19 Pretschner, p. 2
Claim 10	
10. The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary	"A knowledge base or other data repository stores information associating individual topics with sets of documents related to the topic, and with terms descriptive of the topic." 3:37-39 "Each topic is associated with one or more terms that are synonyms for each other. A term is a word or series of words (e.g., a noun phrase) that refer to a topic. A topic describes a possible subject annotation for documents in the document collection 140." 5:50-55. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g. Reader, para. 0015 Wieser, p. 18, line 24 to p. 19, line 1; p. 22, lines 3-15; Fig. 15 Rhodes, 4:45-55 Mase, p. 377, col. 2; p. 379, col. 1 Donaldson, 16:5-9 HyPursuit, p. 185, col. 1 Fleming, 6:45-60, 12:3-11, Fig. 9 OracleText White Paper, p. 19 Pretschner, p. 2
Claim 18	
18. An article of manufacture for use in a computer system, comprising:	"Referring now FIG. 2 there is shown an illustration of the software architecture of an information retrieval system 100 in accordance with the present invention. The information retrieval system 100 includes various functional software modules and structures that execute on a conventional computer system." 5:25-30.

	See Figs. 2, 9.
a memory;	"The computer system includes a processor 105, addressable memory 103, operating system 107, display device 109, and user input device, such as a keyboard 111 or a mouse." 5:30-33. See Fig. 2 at 103. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 4, line 26 to p. 5 line 1 Rhodes, Fig. 1, 9:18-34 Black, 4:58-67, 1:50-57, 4:30-43. Donaldson, Fig. 4, 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 Finkelstein, p. 406 Fleming, Fig. 1, 4:34-41 Ford, 3:34-39, Fig. 3 Kraft, 1:23-27, Fig. 1 Apte, Figs. 3, 4, claim 33 Henkin, 45:37-38, Fig. 22 Pretschner, p. 7
instructions stored in the	See Chart for Claim 1 (above).

memory for operating a method for automatically generating a query from selected document content, comprising:	"Referring now FIG. 2 there is shown an illustration of the software architecture of an information retrieval system 100 in accordance with the present invention. The information retrieval system 100 includes various functional software modules and structures that execute on a conventional computer system." 5:25-30. See Fig. 2 at 110, 120, 150. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 5, lines 8-12, p. 2 lines 16-19; p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22; p. 14, line 19 to p. 15 line 5 Rhodes, 9:18-34, 10:1-17 Black, 4:58-67, 1L50-57 Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 col. 2. Finkelstein, p. 406 Fleming, 4:34-41 Ford, 13:10-19 Kraft, 1:23-27 Henkin, 45:46-50, Fig. 22 Pretschner, pp. 1, 7
defining an organized classification of document content with each class in the organized classification of	See Chart for Claim 1 (above).

document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	
automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	See Chart for Claim 1 (above).
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).
Claim 19	
19. The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content."	See Chart for Claim 2 (above).

Second Supplemental Chart B-8

Claim chart for U.S. Patent No. 6,473,752 to Fleming

as prior art to

U.S. Patent No. 6,778,979

This chart is based upon Xerox's apparent construction of the claims, and is not an admission that those constructions are correct or appropriate.

'979 Patent	**U.S. Patent No. 6,473,752**
Claim 1	
A method for automatically generating a query from selected document content, comprising:	"A system for locating computer documents or data of interest to a user without specification by the user of topics of interest. The system detects the selection of computer documents by the user of the system, and analyzes the contents of the selected computer documents to identify topics to which the contents are related. . . . The system then proceeds without user intervention, using the identified topics . . . to generate topics of interest to the user. . . . The system then uses the prioritized generated topics of user interest to locate documents whose contents are of interest to the user, and makes the located documents available to the user for selection. . . . For example, a computer document search engine can be used to locate additional documents by generating an appropriate search query." Abstract. See Fig. 2. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0001, 0002, 0004, 0015, Fig. 4. Black, Abstract, 1:64-67, 2:43, 1:47-48, Fig. 1, 4:18-20, 1:50-54, 2:10-14, 2:57-61, 2:42-45, 2:26-29, 3:29-35, 3:37-45, 3:55-57 Wieser, p. 5, lines 8-12, p. 2 lines 16-19 Rhodes, 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19,

	13:42-47 Finkelstein, p. 406 Horowitz, 2:42-51, 3:20-23 Ford, 5:38, 5:30-39 Kraft, Abstract Apte, 6:60-67 Henkin, 2:42-49, 44:8-19 Pretschner, p. 1
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	"[V]arious companies have developed information search engines which can automatically index and organize information that is accessible from a computer. This accessible information may be located on any networked computer or storage device that the computer can access, or may be located on the computer system itself. After the information is indexed or organized, these search engines can then search the indexed or organized information to locate particular information of interest." 1:44-53. "The search engine will analyze the contents of the documents, and create an index of some or all of the terms in the documents. The search engine may also attempt to identify one or more general topics to which the entire document relates. The search engine will next search the documents for references to other computer documents. Upon finding such references, the search engine will access those referenced documents and continue the same process. In this manner, the search engines can eventually traverse and index all computer documents that are interconnected with the first documents given to the search engine. After creating this comprehensive index, the search engine can locate documents by receiving a search query containing terms or topics of interest to a user, and by searching the index to locate documents with corresponding terms or topics." 2:12-26. "After usage information is recorded and relevant terms are extracted, the Document Access Monitor 131 then forwards the recorded usage information and the extracted document

	terms to the Topic Analyzer 132. The Topic Generator 221 first receives the information from the Document Access Monitor 131, end generates topics related to areas of interest to the computer user. The Topic Generator 221 begins by generating related terms for the extracted terms. . . . Those skilled in the art will appreciate that related terms can be generated in a variety of ways, including the use of a thesaurus or the use of empirical testing to determine how terms are actually used." 6:24-44. "The related terms . . . assist in generating topics that are related to the contents of accessed documents. For example, if an accessed document contained only the term 'bat,' it would be difficult to determine which meaning of 'bat' was of interest to the user. However, if other extracted terms from this or other accessed documents included related terms such as 'Count Dracula' or 'a flying mammal' but not terms related to baseball, then it is likely that the user is interested in topics related to the mammal but not to the sport. Therefore, after the related terms are generated, the extracted terms are grouped together and used to determine topics of interest to the user. Those skilled in the art will appreciate that this grouping and determination can be performed in a variety of ways." 6:45-60. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Mase, Abstract, Fig. 1, p. 377-378 Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55 Donaldson, 15:38-46, 17:17-19, 19:51-56, Fig. 8c HyPursuit, Fig. 3, p. 181, 184 Finkelstein, p. 410

	Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40 Apte, 9:33-45 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19 Pretschner, pp. 1-2
automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"The system also analyzes the contents of the selected computer documents to identify relevant terms in the contents of the documents, and more generally to identify topics to which the contents are related." 3:5-8. "The system then attempts to locate additional computer documents, on any computer or device that is accessible to the system, whose contents are related to these prioritized generated topics of user interest. One method that the system may use to locate these documents involves identifying a computer document search engine, generating an appropriate search query, and requesting the search engine to perform the search on the generated search query." 3:14-23. "The Document Term Extractor 215 receives the detected document access notifications and extracts relevant terms from the contents of the documents that are accessed. In one embodiment, the detected document access notifications indicate only that a document is being accessed by the user, and the Document Term Extractor 215 accesses the document to ascertain its contents. In another embodiment, the detected document access notifications not only indicate that a document is being accessed, but also include the document contents that are being accessed. Typically, only relevant terms are extracted from the document contents, with relevance measured by the degree of relation between a term and the contents." 5:62- 6:7. "Those skilled in the art will appreciate that relevant terms can be generated in a variety of ways, and can be extracted from the contents of the entire document or only from the

	content of the portions of the document with which the user interacts." 9:25-29. See Figs. 2, 5, 8, 10B. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0015 Wieser, p. 14, line 19 to p. 15 line 5 Rhodes, 13:19-34, 1:56-2:6, 2:15-54 Black, 2:26-29, 2:42-45, 3:29-35 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7 Ford, 11:4-11, 12:52-63 Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B OracleText White Paper, p. 8, 9 Pretschner, p. 2
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	"The present invention relates generally to locating computer documents and more particularly to determining topics of interest to a user and locating documents related to those topics." 1:6-9. "The system also analyzes the contents of the selected computer documents to identify relevant terms in the contents of the documents, and more generally to identify topics to which the contents are related." 3:5-8.

	"Those skilled in the art will appreciate that relevant terms can be generated in a variety of ways, and can be extracted from the contents of the entire document or only from the content of the portions of the document with which the user interacts." 9:25-29. "Those skilled in the art will appreciate that topics of user interest can be generated in a variety of ways, that the importance of a topic can be calculated in a variety of ways (i.e., using of a variety of importance measures), and that the topics can be prioritized in a variety of ways." 9:65-10:2. "Those skilled in the art will appreciate that other methods of generating topics related to document contents are possible. These may include methods that do not use extracted terms . . . or generating a new term that is broad enough to include multiple extracted and related terms." 12:18-24. See Figs. 2, 5, 6, 9. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Mase, p. 377-379, Fig. 1 Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 5:12-28 Black, 2:10-18, 2:57-61 Donaldson, Figs. 9b and 9c, 18:47-63, 20:3-18, 20:19-29 HyPursuit, p. 181, 182, 185, 186, 191 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Ford, 11:4-17

	Kraft, 11:16-40 Apte, 9:33-37 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	"The New Document Identifier receives the prioritized topics from the Topic Analyzer, generates a search query related to the prioritized topics, performs a search of accessible documents using the search query, and identifies documents that contain terms or topics related to the search query. These identified documents are then made available to the user for selection." 4:23-29. "Those skilled in the art will appreciate that a variety of search queries can be formulated to identify documents related to a given set of topics, and that these different search queries will often identify different groups of documents. For example, a search could be generated to identify only those documents whose contents contain every prioritized generated topic, or instead a search could be generated to identify those documents whose contents contain any prioritized generated topic. The generated search query is forwarded to the Search Engine 233, which uses the search query to perform a search on accessible computer documents. Those skilled in the art will appreciate that a variety of search engines are known in the art, including search engines from companies such as Infoseek, Excite, and Digital Equipment Corporation (DEC). Each search engine has its own rules and syntax for the search queries used by it. Thus, the particular search query that is generated will depend on the search engine to be used and on a determination of which available search query for that engine is most likely to return the desired documents." 7:51-8:4. "Those skilled in the art will appreciate that a search query can be generated in a variety of ways, including varying syntax to reflect a particular search engine and varying the topics in the search query in an attempt to identify different groups of documents." 10:34-38.

	"Generated search query 1050 illustrates that although one preferred embodiment prioritized topics based only on extracted terms, the generated search query includes both extracted terms and related terms." 15:19-23. See Figs. 2, 7, 10B. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015, Fig. 4. Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Rhodes, 5:12-28 Black, 3:37-45, 3:55-57 Donaldson, 18:47-54, Figs. 9b and 9c, 18:55-63, 20:3-29 HyPursuit, p. 181, 182, 184, 185, 186 Finkelstein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8 Ford, 12:2-7, 11:4-17 Kraft, 13:40-50, 13:64-67, 14:43-52 Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44 Pretschner, pp. 1-2
Claim 10	
10. The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic	"The related terms . . . assist in generating topics that are related to the contents of accessed documents. For example, if an accessed document contained only the term 'bat,' it would be difficult to determine which meaning of 'bat' was of interest to the user. However, if other extracted terms from

vocabulary.	this or other accessed documents included related terms such as ‘Count Dracula’ or ‘a flying mammal’ but not terms related to baseball, then it is likely that the user is interested in topics related to the mammal but not to the sport. Therefore, after the related terms are generated, the extracted terms are grouped together and used to determine topics of interest to the user. Those skilled in the art will appreciate that this grouping and determination can be performed in a variety of ways.” 6:45-60. “After the related terms are generated, the extracted and related terms are used to create the generated topics 950. Generated topics 950 comprises seven groups, with Group 1 including the extracted terms "President Clinton," "Hillary Clinton," "White House," and "U.S.," as well as related terms "Bill Clinton" and "Arkansas Governor." Group 2 includes the extracted terms of "educational system," "testing," "teachers," "standards," and "schools," as well as the related terms of "Secretary of Education," "universities," "tests," "exams," and "recess." Groups 3-7 contain similar sets of extracted and related terms.” 12:3-11 See Fig. 9. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Wieser, p. 18, line 24 to p. 19, line 1; p. 22, lines 3-15; Fig. 15 Rhodes, 4:45-55 Mase, p. 377, col. 2; p. 379, col. 1 Donaldson, 16:5-9 HyPursuit, p. 185, col. 1 Horowitz, 3:37-39, 5:50-55 OracleText White Paper, p. 19

	Pretschner, p. 2
Claim 18	
18. An article of manufacture for use in a computer system, comprising:	See Fig. 1 (depicting system implemented on computer system). "The computer system 110 includes the CPU 120, the memory 130, the input/output devices 140 and the bus 148. The input/output devices 140 include a storage device 141, a display 142, a keyboard 143 and a computer-readable media drive 144. The memory 130 includes the RDS system 135, which comprises the Document Access Monitor component 131, the Topic Analyzer component 132, and the New Document Identifier component 133." 4:34-41.
a memory;	See Fig. 1 at 130. "The computer system 110 includes the CPU 120, the memory 130, the input/output devices 140 and the bus 148. The input/output devices 140 include a storage device 141, a display 142, a keyboard 143 and a computer-readable media drive 144. The memory 130 includes the RDS system 135, which comprises the Document Access Monitor component 131, the Topic Analyzer component 132, and the New Document Identifier component 133." 4:34-41. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 4, line 26 to p. 5 line 1 Rhodes, Fig. 1, 9:18-34 Black, 4:58-67, 1:50-57, 4:30-43. Donaldson, Fig. 4, 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189

	Finkelstein, p. 406 Horowitz, 5:25-30, Figs. 2, 9 Ford, 3:34-39, Fig. 3 Kraft, 1:23-27, Fig. 1 Apte, Figs. 3, 4, claim 33 Henkin, 45:37-38, Fig. 22 Pretschner, p. 7
instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	See Chart for Claim 1 (above). Instructions for the system are inherently stored in the memory depicted in figure 1. "The computer system 110 includes the CPU 120, the memory 130, the input/output devices 140 and the bus 148. The input/output devices 140 include a storage device 141, a display 142, a keyboard 143 and a computer-readable media drive 144. The memory 130 includes the RDS system 135, which comprises the Document Access Monitor component 131, the Topic Analyzer component 132, and the New Document Identifier component 133." 4:34-41. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 5, lines 8-12, p. 2 lines 16-19; p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22; p. 14, line 19 to p. 15 line 5 Rhodes, 9:18-34, 10:1-17 Black, 4:58-67, 1L50-57 Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58

	HyPursuit, p. 189 col. 2. Finkelstein, p. 406 Horowitz, 5:30-33, Fig. 2 at 103 Ford, 13:10-19 Kraft, 1:23-27 Henkin, 45:46-50, Fig. 22 Pretschner, pp. 1, 7
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	See Chart for Claim 1 (above).
automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	See Chart for Claim 1 (above).
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).

Second Supplemental Chart B-9

Claim chart for U.S. Patent No. 6,606,644 to Ford

as prior art to

U.S. Patent No. 6,778,979

This chart is based upon Xerox's apparent construction of the claims, and is not an admission that those constructions are correct or appropriate.

'979 Patent	**U.S. Patent No. 6,606,644**
Claim 1	
A method for automatically generating a query from selected document content, comprising:	"The present invention provides a method and apparatus for automatically gathering, summarizing, and indexing real-time information derived from real-time communication on the Internet, such as Internet chat sessions, and to make that data readily available for immediate analysis and use such as targeted advertising." 5:3-8. "The present invention will give advertisers the ability to dynamically monitor the conversation being held in any given chat room, and be able to display advertising banners that match the theme of the conversation, thus, eliciting greater attention and interest from users. For example, a sudden occurrence of keywords such as 'car', 'automobile', 'drive', 'convertible', 'coupe', etc., may signal to an automobile manufacturer that now is an opportune time to display a banner advertising a special sale on convertible automobiles." 5:30-39. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0001, 0002, 0004, 0015, Fig. 4. Black, Abstract, 1:64-67, 2:43, 1:47-48, Fig. 1, 4:18-20, 1:50-54, 2:10-14, 2:57-61, 2:42-45, 2:26-29, 3:29-35, 3:37-45, 3:55-57 Wieser, p. 5, lines 8-12, p. 2 lines 16-19 Rhodes, 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein, p. 406 Horowitz, 2:42-51, 3:20-23

	Fleming, Abstract, Fig. 2 Kraft, Abstract Apte, 6:60-67 Henkin, 2:42-49, 44:8-19 Pretschner, p. 1
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	"As shown in FIG. 3, this aspect of the present invention [i.e., serving contextually-relevant advertisements] will extract and analyze the summary information generated by LISA 50 and then respond with an appropriate message or advertisement responsive to the subject matter of the summary information. Specifically, based upon the user's specifications, the present invention will look for the temporally contiguous occurrence of a particular keyword or set of keywords. The appearance of such keywords would be used to infer the general topic or subject matter of the communication or conversation. Once the general subject area of the conversation has been (probabilistically) deduced, the present invention may be used to present advertisements in the chat room that match the theme of the conversation." 11:4-17. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Mase, Abstract, Fig. 1, p. 377-378 Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55 Donaldson, 15:38-46, 17:17-19, 19:51-56, Fig. 8c HyPursuit, Fig. 3, p. 181, 184 Finkelstein, p. 410 Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Kraft, 5:55-6:10, 11:16-40

	Apte, 9:33-45 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19 Pretschner, pp. 1-2
automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"As shown in FIG. 3, this aspect of the present invention [i.e., serving contextually-relevant advertisements] will extract and analyze the summary information generated by LISA 50 and then respond with an appropriate message or advertisement responsive to the subject matter of the summary information. Specifically, based upon the user's specifications, the present invention will look for the temporally contiguous occurrence of a particular keyword or set of keywords." 11:4-11. "In particular, using the information it has received from database manager 34, client interface 36 can present the advertisement in extremely customized, personalized, and interactive manner. For example, 'Bob' is a participant in a conversation in which the key words 'linux', 'windows', 'operating systems', 'software', and 'application', have been detected by LISA 50. Client interface 36 may now present Bob with a personal message that contains an advertisement for 'red hat linux' products, and which may state something like 'Hi Bob, I felt you might be interested in this !" 12:52-63. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0015 Wieser, p. 14, line 19 to p. 15 line 5 Rhodes, 13:19-34, 1:56-2:6, 2:15-54 Black, 2:26-29, 2:42-45, 3:29-35 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7

	Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B OracleText White Paper, p. 8, 9 Pretschner, p. 2
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	"As shown in FIG. 3, this aspect of the present invention [i.e., serving contextually-relevant advertisements] will extract and analyze the summary information generated by LISA 50 and then respond with an appropriate message or advertisement responsive to the subject matter of the summary information. Specifically, based upon the user's specifications, the present invention will look for the temporally contiguous occurrence of a particular keyword or set of keywords. The appearance of such keywords would be used to infer the general topic or subject matter of the communication or conversation. Once the general subject area of the conversation has been (probabilistically) deduced, the present invention may be used to present advertisements in the chat room that match the theme of the conversation." 11:4-17. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Mase, p. 377-379, Fig. 1 Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 5:12-28 Black, 2:10-18, 2:57-61 Donaldson, Figs. 9b and 9c, 18:47-63, 20:3-18, 20:19-29 HyPursuit, p. 181, 182, 185, 186, 191 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8

	Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Kraft, 11:16-40 Apte, 9:33-37 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	"As shown in FIG. 3, this aspect of the present invention [i.e., serving contextually-relevant advertisements] will extract and analyze the summary information generated by LISA 50 and then respond with an appropriate message or advertisement responsive to the subject matter of the summary information. Specifically, based upon the user's specifications, the present invention will look for the temporally contiguous occurrence of a particular keyword or set of keywords. The appearance of such keywords would be used to infer the general topic or subject matter of the communication or conversation. Once the general subject area of the conversation has been (probabilistically) deduced, the present invention may be used to present advertisements in the chat room that match the theme of the conversation." 11:4-17. "In a targeted advertising aspect of the present invention, database manager 34 is adapted to be able to determine which advertisement(s) from the database 60 of advertisements should be displayed in response to a particular message being exchanged in real-time between two 'chatters' or clients 70." 12:2-7. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015, Fig. 4. Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Rhodes, 5:12-28 Black, 3:37-45, 3:55-57 Donaldson, 18:47-54, Figs. 9b and 9c, 18:55-63, 20:3-29

	HyPursuit, p. 181, 182, 184, 185, 186 Finkelstein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8 Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Kraft, 13:40-50, 13:64-67, 14:43-52 Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44 Pretschner, pp. 1-2
Claim 18	
An article of manufacture for use in a computer system, comprising:	"In another aspect, the present invention is directed to an article of manufacture comprising a computer usable medium having computer readable program code means embodied therein for causing a computer to automatically gather, summarize and index real-time communication on a computer network between at least a first and second user." 3:34-39. See also Fig. 3.
a memory;	"The program storage devices of the present invention may be devised, made and used as a component of a machine utilizing optics, magnetic properties and/or electronics to perform the method steps of the present invention. Program storage devices include, but are not limited to, magnetic media such as a diskette or computer hard drive, which is readable and executable by a computer(s), optical disks, Read Only Memory (ROM), floppy disks, and semiconductor chips." 13:11-19. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 4, line 26 to p. 5 line 1 Rhodes, Fig. 1, 9:18-34

	Black, 4:58-67, 1:50-57, 4:30-43. Donaldson, Fig. 4, 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 Finkelstein, p. 406 Horowitz, 5:25-30, Figs. 2, 9 Fleming, Fig. 1, 4:34-41 Kraft, 1:23-27, Fig. 1 Apte, Figs. 3, 4, claim 33 Henkin, 45:37-38, Fig. 22 Pretschner, p. 7
instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	See Chart for Claim 1 (above). "The program storage devices of the present invention may be devised, made and used as a component of a machine utilizing optics, magnetic properties and/or electronics to perform the method steps of the present invention. Program storage devices include, but are not limited to, magnetic media such as a diskette or computer hard drive, which is readable and executable by a computer(s), optical disks, Read Only Memory (ROM), floppy disks, and semiconductor chips." 13:11-19. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 5, lines 8-12, p. 2 lines 16-19; p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22; p. 14, line 19 to p. 15 line 5 Rhodes, 9:18-34, 10:1-17 Black, 4:58-67, 1L50-57

	Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 col. 2. Finkelstein, p. 406 Horowitz, 5:30-33, Fig. 2 at 103 Fleming, 4:34-41 Kraft, 1:23-27 Henkin, 45:46-50, Fig. 22 Pretschner, pp. 1, 7
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	See Chart for Claim 1 (above).
automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	See Chart for Claim 1 (above).
automatically formulating the query to restrict a search at the information retrieval system	See Chart for Claim 1 (above).

for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	

Second Supplemental Chart B-10

Claim chart of U.S. Patent No. 6,829,780 to Kraft

as prior art to

U.S. Patent No. 6,778,979

This chart is based upon Xerox's apparent construction of the claims, and is not an admission that those constructions are correct or appropriate.

'979 Patent	**U.S. Patent No. 6,829,780**
Claim 1	
1. A method for automatically generating a query from selected document content, comprising:	"Transparently to the user, the system continuously operates in the background to adapt banner advertisements based on the detection of competing ads. The system includes a page analyzer that translates the hosted ad's web page into a document that can be analyzed for the presence of competing ads. An ad comparison unit compares the output of the page analyzer with information stored in the competitor ad database to detect competing ads. . . . The ad summary evaluator identifies competing ads and devises a counter strategy for banner ad display." Abstract. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0001, 0002, 0004, 0015, Fig. 4. Black, Abstract, 1:64-67, 2:43, 1:47-48, Fig. 1, 4:18-20, 1:50-54, 2:10-14, 2:57-61, 2:42-45, 2:26-29, 3:29-35, 3:37-45, 3:55-57 Wieser, p. 5, lines 8-12, p. 2 lines 16-19 Rhodes, 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein, p. 406 Horowitz, 2:42-51, 3:20-23 Fleming, Abstract, Fig. 2 Ford, 5:38, 5:30-39

	Apte, 6:60-67 Henkin, 2:42-49, 44:8-19 Pretschner, p. 1
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	"In one embodiment, the system 10 operates with a banner display module 200 that displays the content of the adaptive banner ad, a web server 15, and an ad proxy router 205 that provides secure communication link between banner display module and the advertiser's server, and is generally comprised of: . . . a hosted ad database 220 that contains information about each banner ad hosted, including primary competitors; . . . a competitor ad database 235 that contains data about competing advertising;" 5:55-6:10. "The competitor ad database 235 contains data about competitors' advertisements that have [been] collected from the various advertisers and entered either manually or automatically. This data comprises a sample of the ad (e.g. an image) along with additional data useful for ad analysis. The system 10 performs database queries utilizing available image/multimedia comparison algorithms to locate a match. For the example of FIG. 4, the ad comparison unit 230 detects a match and the ad for the Mac Store is marked in the summary list as a <MATCHED_COMPETITOR>: <COMPETITOR id=4711> <MEDIA_URL>http://ad.doubleclicknet/ad/buy.prod.sm/homepage;cat=homepage _5;sz=100x60;tile=5;ord=16115127561</MEDIA_URL> <TARGET_URL>http://ad.doubleclick.net/jump/buy.prod.sm/homepage;cat=homep age_5;sz=100x60;ti1e=5;ord=16115127561</TARGET_URL> <title>BUYCOMP.COM - The Computer Superstore</title> <KEYWORDS>macintosh, mac, store</KEYWORDS> <MATCHED_COMPETITOR>Apple Inc.</MATCHED_COMPETITOR> <COMPETITOR>" 11:16-40. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill

	in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Mase, Abstract, Fig. 1, p. 377-378 Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55 Donaldson, 15:38-46, 17:17-19, 19:51-56, Fig. 8c HyPursuit, Fig. 3, p. 181, 184 Finkelstein, p. 410 Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Ford, 11:4-17 Apte, 9:33-45 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19 Pretschner, pp. 1-2
automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"In one embodiment, the system 10 operates with a banner display module 200 that displays the content of the adaptive banner ad, a web server 15, and an ad proxy router 205 that provides secure communication link between banner display module and the advertiser's server, and is generally comprised of: an ad identification manager 215 that performs competitive analysis of all the ads on the hosted ad's web page; . . . a page analyzer 225 that translates the hosted ad's web page into a document that can be analyzed for the presence of competing advertising; . . . an ad comparison unit 230 that compares the output of the page analyzer 225 with information stored in the competitor ad database 235 to detect competing advertising;" 5:55-6:14.

	"As illustrated in the foregoing example, the page analyzer 225 also extracts keywords and alternate representations from the web page and adds them to the page summary." 9:34-36. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0015 Wieser, p. 14, line 19 to p. 15 line 5 Rhodes, 13:19-34, 1:56-2:6, 2:15-54 Black, 2:26-29, 2:42-45, 3:29-35 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7 Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Ford, 11:4-11, 12:52-63 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B OracleText White Paper, p. 8, 9 Pretschner, p. 2
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	"The competitor ad database 235 contains data about competitors' advertisements that have [been] collected from the various advertisers and entered either manually or automatically. This data comprises a sample of the ad (e.g. an image) along with additional data useful for ad analysis. The system 10 performs database queries utilizing available image/multimedia comparison algorithms to locate a match. For the example of FIG. 4, the ad comparison unit 230 detects a match and the ad for the Mac Store is marked in the summary list as a <MATCHED_COMPETITOR>: <COMPETITOR id=4711>

	<MEDIA_URL>http://ad.doubleclicknet/ad/buy.prod.sm/homepage;cat=homepage _5;sz=100x60;tile=5;ord=16115127561</MEDIA_URL> <TARGET_URL>http://ad.doubleclick.net/jump/buy.prod.sm/homepage;cat=homep age_5;sz=100x60;ti1e=5;ord=16115127561</TARGET_URL> <title>BUYCOMP.COM - The Computer Superstore</title> <KEYWORDS>macintosh, mac, store</KEYWORDS> <MATCHED_COMPETITOR>Apple Inc.</MATCHED_COMPETITOR> <COMPETITOR>" 11:16-40. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Mase, p. 377-379, Fig. 1 Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 5:12-28 Black, 2:10-18, 2:57-61 Donaldson, Figs. 9b and 9c, 18:47-63, 20:3-18, 20:19-29 HyPursuit, p. 181, 182, 185, 186, 191 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17 Apte, 9:33-37 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2

automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	"The web page summary now contains all the information necessary . . . to define an appropriate ad strategy for the web page: Ad ID used to identify the hosted ad in the hosted ad database 220 and associated information; a list of all potential competitors on the web page including the media URL, target URL and additional metadata gained from document analysis or OCR shown highlighted in bold letters; and a list of competitors clearly identified by the ad comparison unit 230 as <MATCHED COMPETITOR>." 13:40-50. "Method 300 then determines from the hosted ad database 220, at decision step 350, if any of the ads on the web page are key competitors by having the page analyzer 225 pass the web page summary to the ad summary evaluator 255." 13:64-67. "The strategy or strategies used by the ad summary evaluator 255 are implemented as rules that are stored in the ad rules database 260. A rules engine is used to interpret these rules. For the example of FIG. 4, the ad summary evaluator 255 identifies the Mac Store as a key competitor. The ad summary evaluator 255 ignores the other ads because they offer services in different markets and thus are not competitors. From the ad rules database 260, the ad summary evaluator 255 obtains a rule which instructs the system 10 to use a specific Mac-targeted ad instead of the original ad." 14:43-52. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015, Fig. 4. Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Rhodes, 5:12-28 Black, 3:37-45, 3:55-57 Donaldson, 18:47-54, Figs. 9b and 9c, 18:55-63, 20:3-29 HyPursuit, p. 181, 182, 184, 185, 186 Finkelstein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8

	Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17 Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44 Pretschner, pp. 1-2
Claim 2	
2. The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	"At this point, the page analyzer 225 can also download the target URLs of each potential candidate and analyze these documents. Typically, these documents contain useful information (e.g. page title, abstracts etc.) that can be used to enrich the metadata of the summary document…" 10:32-36. The final XML representation created by the page analyzer 225 for the example of FIG. 4 is as follows: … <COMPETITOR id~4711> <MEDIA_URL>http://ad.doubleclicknet/adibuy.prod.sm/homepage;cat~homepage_ 5;sz~100x60;tile~5;ord~16115127561</MEDIA_URL> <TARGET_URL>http://ad.doubleclicknet/jump/buy.prod.sm/homepage;cat~homepage_ 5;sz~100x60;tile~5;ord~16115127561</TARGET_URL> <title>BUYCOMP.COM -The Computer Superstore</title> <KEYWORDS>macintosh, mac, store</KEYWORDS> <MATCHED_COMPETITOR>Apple [nc.</MATCHED_ COMPETITOR> <COMPETITOR> ... 13:2-40. "If, however, none of the identified candidates are competitors, other unidentified competitors may still reside on the web page. Consequently, a supplemental analysis might be required, and the page analyzer 225 invokes the OCR engine 240 at step 330 to convert the image data into text data. Performing OCR analysis on the web page retrieves additional information that can enhance the XML representation of potential candidates." 12:13-21. "The web page summary now contains all the information necessary . . . to define an appropriate ad strategy for the web page: Ad ID used to identify the hosted ad in the hosted ad database 220

	and associated information; a list of all potential competitors on the web page including the media URL, target URL and additional metadata gained from document analysis or OCR shown highlighted in bold letters; and a list of competitors clearly identified by the ad comparison unit 230 as <MATCHED COMPETITOR>." 13:40-50. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g Rhodes, 10:42-51, 11:10-31, 12:41-49, 13:1-8, 13:15-34, 13:42-47 Black, 1:64-2:9, 3:37-45, 3:55-57 Donaldson, 18:47-63, 20:3-29, Fig. 9b, Fig. 9c HyPursuit, p. 186, col. 2 Horowitz, 7:19-25, 7:46-49 Henkin, 27:65-28:5 Stibel, 2:42-47, 9:21-43, 10:30-36, 11:56-66, 12:13-21, Table 2 Pretschner, p. 2
Claim 18	
18. An article of manufacture for use in a computer system, comprising:	"[T]his invention pertains to a computer software product for dynamically adapting, enhancing, and optimizing the appearance and content of a banner advertisement based on the automatic detection of competing advertising within a document." 1:23-27. See also Fig. 1.
a memory;	"[T]his invention pertains to a computer software product for dynamically adapting, enhancing, and optimizing the appearance and content of a banner advertisement based on the automatic detection of competing advertising within a document." 1:23-27. A computer software product inherently requires the use of a memory. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.:

	Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 4, line 26 to p. 5 line 1 Rhodes, Fig. 1, 9:18-34 Black, 4:58-67, 1:50-57, 4:30-43. Donaldson, Fig. 4, 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 Finkelstein, p. 406 Horowitz, 5:25-30, Figs. 2, 9 Fleming, Fig. 1, 4:34-41 Ford, 3:34-39, Fig. 3 Apte, Figs. 3, 4, claim 33 Henkin, 45:37-38, Fig. 22 Pretschner, p. 7
instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	See Chart for Claim 1 (above). "[T]his invention pertains to a computer software product for dynamically adapting, enhancing, and optimizing the appearance and content of a banner advertisement based on the automatic detection of competing advertising within a document." 1:23-27. A computer software product inherently requires the use of instructions stored in memory for it to be executable. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1.

	Wieser, p. 5, lines 8-12, p. 2 lines 16-19; p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22; p. 14, line 19 to p. 15 line 5 Rhodes, 9:18-34, 10:1-17 Black, 4:58-67, 1L50-57 Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 col. 2. Finkelstein, p. 406 Horowitz, 5:30-33, Fig. 2 at 103 Fleming, 4:34-41 Ford, 13:10-19 Henkin, 45:46-50, Fig. 22 Pretschner, pp. 1, 7
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	See Chart for Claim 1 (above).
automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
automatically categorizing the selected document content using the organized classification of document content for assigning the	See Chart for Claim 1 (above).

selected document content a classification label from the organized classification of content; and	
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).
Claim 19	
19. The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	See Chart for Claim 2 (above).

Second Supplemental Chart B-11

Claim chart for U.S. Patent No. 7,225,142 to Apte

as prior art to

U.S. Patent No. 6,778,979

This chart is based upon Xerox's apparent construction of the claims, and is not an admission that those constructions are correct or appropriate.

'979 Patent	**U.S. Patent No. 7,225,142**
Claim 1	
A method for automatically generating a query from selected document content, comprising:	"Because the advertisements are streamed from a server rather than downloaded as a set and played to the user in a loop, the present invention can make choices about which advertisements to display to the user that are responsive to the user's current viewing habits. Thus, if a user is selecting and viewing pages in the browser area 31 concerning outdoor activities, the present invention can select advertisements for camping gear . . . ." 6:60-67. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0001, 0002, 0004, 0015, Fig. 4. Black, Abstract, 1:64-67, 2:43, 1:47-48, Fig. 1, 4:18-20, 1:50-54, 2:10-14, 2:57-61, 2:42-45, 2:26-29, 3:29-35, 3:37-45, 3:55-57 Wieser, p. 5, lines 8-12, p. 2 lines 16-19 Rhodes, 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein, p. 406 Horowitz, 2:42-51, 3:20-23 Fleming, Abstract, Fig. 2 Ford, 5:38, 5:30-39 Kraft, Abstract Henkin, 2:42-49, 44:8-19

	Pretschner, p. 1
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	"After extracting the keywords, the keywords are compared to a database index, which cross-references keywords with topic names. Thus, in the present example, the keyword 'surfing' matches topics 'outdoor adventure' and 'water sports.' 'Molokai' matches the topic 'Hawaii.' Each topic in the database is correlated with a series of URLs for advertisements that relate to the topic. Thus, the topic 'Hawaii' corresponds advertisements for the 'Airline Deals to Hawaii by TravelNow' and 'Luau Hawaiian Hotels,' which are now streamed to the user and displayed in the advertising area 37. In this way, the user's viewing habits are used to effectively target advertisements to the user that are pertinent to the user's interests." 9:33-45. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Mase, Abstract, Fig. 1, p. 377-378 Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55 Donaldson, 15:38-46, 17:17-19, 19:51-56, Fig. 8c HyPursuit, Fig. 3, p. 181, 184 Finkelstein, p. 410 Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19

	Pretschner, pp. 1-2
automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"In one embodiment, the present invention carries out this content-sensitive advertising by conducting a keyword search of a page requested to be displayed on the client computer by the user. Keywords are obtained by noting words that appear between TITLE headers in HTML documents. For example, a page that contains the code: <TITLE>Bill's Favorite Surfing Spots on Molokai</TITLE> the keywords 'surfing' and 'Molokai' would be extracted as keywords." 9:24-32. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0015 Wieser, p. 14, line 19 to p. 15 line 5 Rhodes, 13:19-34, 1:56-2:6, 2:15-54 Black, 2:26-29, 2:42-45, 3:29-35 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7 Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Ford, 11:4-11, 12:52-63 Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B OracleText White Paper, p. 8, 9 Pretschner, p. 2
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a	"After extracting the keywords, the keywords are compared to a database index, which cross-references keywords with topic names. Thus, in the present example, the keyword 'surfing' matches topics 'outdoor adventure' and 'water sports.' 'Molokai' matches the topic 'Hawaii.'" 9:33-37.

classification label from the organized classification of content; and	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Mase, p. 377-379, Fig. 1 Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 5:12-28 Black, 2:10-18, 2:57-61 Donaldson, Figs. 9b and 9c, 18:47-63, 20:3-18, 20:19-29 HyPursuit, p. 181, 182, 185, 186, 191 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17 Kraft, 11:16-40 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	"Advertisements may also be selected by deducing user areas of interest based upon the content of pages on the web selected by the user for viewing." 4:20-22. "Each topic in the database is correlated with a series of URLs for advertisements that relate to the topic. Thus, the topic 'Hawaii' corresponds advertisements for the 'Airline Deals to Hawaii by TravelNow' and 'Luau Hawaiian Hotels,' which are now streamed to the user and displayed in the advertising area 37. In this way, the user's viewing habits are used to effectively target advertisements to the user that are pertinent to the user's interests." 9:38-45.

	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015, Fig. 4. Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Rhodes, 5:12-28 Black, 3:37-45, 3:55-57 Donaldson, 18:47-54, Figs. 9b and 9c, 18:55-63, 20:3-29 HyPursuit, p. 181, 182, 184, 185, 186 Finkelstein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8 Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17 Kraft, 13:40-50, 13:64-67, 14:43-52 Henkin, 27:46-28:5, 28:30-44 Pretschner, pp. 1-2
Claim 18	
18. An article of manufacture for use in a computer system, comprising:	See Figs. 3, 4; Claim 33. "A client computer for presenting advertising to a user, comprising:. . . [A] memory that stores browser software adapted to be executed . . . ." Claim 33.
a memory;	See Figs. 3, 4. The computers, which are used to carry out the advertising selection processes disclosed in the patent, inherently include memory. "A client computer for presenting advertising to a user, comprising: . . . [A] memory that stores browser software adapted to be executed . . . ." Claim 33. To the extent this reference does not teach this claim element, this

	reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 4, line 26 to p. 5 line 1 Rhodes, Fig. 1, 9:18-34 Black, 4:58-67, 1:50-57, 4:30-43. Donaldson, Fig. 4, 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 Finkelstein, p. 406 Horowitz, 5:25-30, Figs. 2, 9 Fleming, Fig. 1, 4:34-41 Ford, 3:34-39, Fig. 3 Kraft, 1:23-27, Fig. 1 Henkin, 45:37-38, Fig. 22 Pretschner, p. 7
instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	See Chart for Claim 1 (above). See Figs. 3, 4. The advertising selection process inherently requires the storage of instructions in memory that are executed on the computer. "A client computer for presenting advertising to a user, comprising: . . . [A] memory that stores browser software adapted to be executed . . . ." Claim 33. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.:

	Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 5, lines 8-12, p. 2 lines 16-19; p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22; p. 14, line 19 to p. 15 line 5 Rhodes, 9:18-34, 10:1-17 Black, 4:58-67, 1L50-57 Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 col. 2. Finkelstein, p. 406 Horowitz, 5:30-33, Fig. 2 at 103 Fleming, 4:34-41 Ford, 13:10-19 Kraft, 1:23-27 Henkin, 45:46-50, Fig. 22 Pretschner, pp. 1, 7
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	See Chart for Claim 1 (above).
automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).

automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	See Chart for Claim 1 (above).
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).

Second Supplemental Chart B-12

Claim chart for U.S. Patent No. 7,451,099 to Henkin

as prior art to

U.S. Patent No. 6,778,979

This chart is based upon Xerox's apparent construction of the claims, and is not an admission that those constructions are correct or appropriate.

'979 Patent	U.S. Patent No. 7,451,099
Claim 1	
1. A method for automatically generating a query from selected document content, comprising:	"When a new document (e.g. a web page) is displayed on the client system to an end user, selected context associated from the document is analyzed for selected keywords. Specific context in the document may then be identified using the selected keyword information. Based upon the identified context in the document, a selected pop-up advertisement may be automatically displayed on the client system." 2:42-49. "It will be appreciated that the technique of the present invention enables businesses and advertisers to proactively interact with existing and potential on-line customers by marking up (e.g. underlining, highlighting, displaying additional text, graphics, and/or sound) selected keywords or phrases on any document, web page or web page which is currently being displayed on the user's computer system. In this way, static HTML pages may be converted at the user's computer system into customized, dynamic information which provides the ability for businesses and advertisers to proactively deliver dynamic, targeted and customized service to the end users via additional information." 44:8-19. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0001, 0002, 0004, 0015, Fig. 4. Black, Abstract, 1:64-67, 2:43, 1:47-48, Fig. 1, 4:18-20, 1:50-54, 2:10-14, 2:57-61, 2:42-45, 2:26-29, 3:29-35, 3:37-45, 3:55-57

	Wieser, p. 5, lines 8-12, p. 2 lines 16-19 Rhodes, 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein, p. 406 Horowitz, 2:42-51, 3:20-23 Fleming, Abstract, Fig. 2 Ford, 5:38, 5:30-39 Kraft, Abstract Apte, 6:60-67 Pretschner, p. 1
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	"Briefly, the contextual inventory is organized and categorized into Super Categories 302, Sub-Categories 304, and Keywords 306, as shown in FIG. 3. In accordance with one specific embodiment, this organizational tree is applied to organize the Keywords and/or phrases under their appropriate product Categories." 14:4-9. "The EZ Gateway 204 also performs category management tasks such as permitting the Ad Campaign Provider to enhance their ontology (the database of categories and keywords) on an ongoing basis, in real time." 9:20-25. "According to a specific embodiment, the contextual pop-ups media feature is based on the ability to identify keywords on the page, classify them into categories, and using the category assign a matching category to a given page. In order to illustrate this aspect of the present invention, an example will now be described in which it is assumed that a document (e.g. web page) is displayed on the user's computer system which includes the following text: truck, car, vehicle, SUV, sport car. In this particular example, the document may be classified as a page corresponding to the category name 'Auto'. Accordingly, in one implementation, it will be appropriate to display information from the 'Auto' category to the end user. In this way, the technique in the present invention provides a benefit of automatically displaying advertisements which match specific context of the page or documents displayed to the end user." 27:14-29.

	"According to a specific embodiment, the Category ID field 802 may be used to identify a specific category (e.g. 304 of FIG. 3) associated with specific keywords, key phrases, or titles. In one implementation, the Category ID value may be represented as a 4-byte integer." 29:25-29. See Figs. 3, 8, 23, 25C; 13:25-55. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Mase, Abstract, Fig. 1, p. 377-378 Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55 Donaldson, 15:38-46, 17:17-19, 19:51-56, Fig. 8c HyPursuit, Fig. 3, p. 181, 184 Finkelstein, p. 410 Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40 Apte, 9:33-45 OracleText White Paper, p. 11, 18, 19 Pretschner, pp. 1-2
automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"According to a specific embodiment, the search engine is designed to support different business requirements. It may operate in a variety of search modes, including an exact search mode and a fuzzy search mode. The search engine may search the document text, WEB PAGE, title, Meta tags, or any other property of the

	selected document for selected key words or phrases. In one embodiment, a search is conducted by analyzing words in the text of a selected document to see if it includes specified keywords or phrases." 4:56-64. "Accordingly, when one of the Clients 110 is surfing the Internet, regardless of what web page they are viewing, the Client Application scans the text of the web page, analyzes the context, and marks up keywords and/or phrases." 7:22-25. "According to a specific embodiment, the contextual pop-ups media feature is based on the ability to identify keywords on the page, classify them into categories, and using the category assign a matching category to a given page. In order to illustrate this aspect of the present invention, an example will now be described in which it is assumed that a document (e.g. web page) is displayed on the user's computer system which includes the following text: truck, car, vehicle, SUV, sport car. In this particular example, the document may be classified as a page corresponding to the category name 'Auto'. Accordingly, in one implementation, it will be appropriate to display information from the 'Auto' category to the end user. In this way, the technique in the present invention provides a benefit of automatically displaying advertisements which match specific context of the page or documents displayed to the end user." 27:14-29. "Further, according to one implementation, different types of context within the document (e.g. document title, Meta keywords, Meta information, document text, etc.) may be weighted differently to emphasize each type's particular relevance. If more than one advertisement is associated with a particular campaign, selection of the appropriate advertisement may be based upon different mechanisms such as, for example, assigned priority, round robin, relative age, etc." 27:65-28:5. See Figs. 16A, 16B. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0015 Wieser, p. 14, line 19 to p. 15 line 5

	Rhodes, 13:19-34, 1:56-2:6, 2:15-54 Black, 2:26-29, 2:42-45, 3:29-35 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7 Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Ford, 11:4-11, 12:52-63 Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 OracleText White Paper, p. 8, 9 Pretschner, p. 2
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	"In a specific embodiment, the MAIN application 520 may be configured to analyze a selected document for keywords, categories and/or super categories in order to find a match for an appropriate pop-up advertisement or window to be displayed." 27:30-34. "According to a specific embodiment, the contextual pop-ups media feature is based on the ability to identify keywords on the page, classify them into categories, and using the category assign a matching category to a given page. In order to illustrate this aspect of the present invention, an example will now be described in which it is assumed that a document (e.g. web page) is displayed on the user's computer system which includes the following text: truck, car, vehicle, SUV, sport car. In this particular example, the document may be classified as a page corresponding to the category name 'Auto'. Accordingly, in one implementation, it will be appropriate to display information from the 'Auto' category to the end user. In this way, the technique in the present invention provides a benefit of automatically displaying advertisements which match specific context of the page or documents displayed to the end user." 27:14-29. "According to a specific embodiment, one or more algorithms may be used for determining the most appropriate matching category for the selected document being analyzed. For example, in one algorithm, a variety of different parameters relating to the current

	document may be analyzed in order to determine the most appropriate matching category." 27:46-51. See Fig. 16A. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Mase, p. 377-379, Fig. 1 Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 5:12-28 Black, 2:10-18, 2:57-61 Donaldson, Figs. 9b and 9c, 18:47-63, 20:3-18, 20:19-29 HyPursuit, p. 181, 182, 185, 186, 191 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17 Kraft, 11:16-40 Apte, 9:33-37 OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information	"According to a specific embodiment, one or more algorithms may be used for determining the most appropriate matching category for the selected document being analyzed. For example, in one algorithm, a variety of different parameters relating to the current document may be analyzed in order to determine the most appropriate matching category. For example, the current document

retrieval system identified by the assigned classification label.	may be analyzed and assigned a specific context score (CS) that is then compared with specific campaign requirement included in the campaign update files. If the context score is greater than or equal to a predetermined threshold value TH, then a pop-up ad (or other media type ads) may be displayed. According to one implementation, keywords which are identified in different elements of the document may be scored appropriately. The cumulative score of all the keywords that are found may be used to determine the CS value. If the identified keywords match a specific category of an ad campaign, and the cumulative CS value is above the threshold for that campaign, then a pop-up advertisement for that campaign may be displayed. Further, according to one implementation, different types of context within the document (e.g. document title, Meta keywords, Meta information, document text, etc.) may be weighted differently to emphasize each type's particular relevance. If more than one advertisement is associated with a particular campaign, selection of the appropriate advertisement may be based upon different mechanisms such as, for example, assigned priority, round robin, relative age, etc." 27:46-28:5. "FIG. 7 shows a specific embodiment of a flow diagram illustrating how various information flows are passed between the client system and the server system of the present invention. Initially, at (30) it is assumed that the user has clicked or selected a particular portion of text which has been marked up in accordance with the technique of the present invention. According to at least one embodiment, when the user clicks on a particular portion of marked up text, a pop-up layer (e.g. dynamic browser control layer) may be displayed (31) to the user providing the user with additional information relating to the topic of the marked up text portion. An example of one type of pop-up layer is illustrated in FIG. 21 of the drawings. According to specific embodiments, the pop-up layer may include, for example, one or more links, audio information, video information, and/or textual information." 28:30-44. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015, Fig. 4. Wieser, p. 17, lines 15-32; p. 18, lines 14-22

	Rhodes, 5:12-28 Black, 3:37-45, 3:55-57 Donaldson, 18:47-54, Figs. 9b and 9c, 18:55-63, 20:3-29 HyPursuit, p. 181, 182, 184, 185, 186 Finkelstein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8 Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17 Kraft, 13:40-50, 13:64-67, 14:43-52 Apte, 4:20-22, 9:38-45 Pretschner, pp. 1-2
Claim 2	
2. The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	"Further, according to one implementation, different types of context within the document (e.g. document title, Meta keywords, Meta information, document text, etc.) may be weighted differently to emphasize each type's particular relevance. If more than one advertisement is associated with a particular campaign, selection of the appropriate advertisement may be based upon different mechanisms such as, for example, assigned priority, round robin, relative age, etc." 27:65-28:5. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Rhodes, 10:42-51, 11:10-31, 12:41-49, 13:1-8, 13:15-34, 13:42-47 Black, 1:64-2:9, 3:37-45, 3:55-57 Donaldson, 18:47-63, 20:3-29, Fig. 9b, Fig. 9c HyPursuit, p. 186, col. 2 Horowitz, 7:19-25, 7:46-49

	Kraft, 10:32-36, 12:13-21, 13:2-50 Stibel, 2:42-47, 9:21-43, 10:30-36, 11:56-66, 12:13-21, Table 2 Pretschner, p. 2
Claim 5	
5. The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	"FIG. 3 is a schematic diagram of the context hierarchy in accordance with a specific embodinment of the present invention." 3:1-3. "Briefly, the contextual inventory is organized and categorized into Super Categories 302, Sub-Categories 304, and Keywords 306, as shown in FIG. 3. In accordance with one specific embodiment, this organizational tree is applied to organize the Keywords and/or phrases under their appropriate product Categories." 14:4-9. "In one implementation, the hierarchy of the Super-Category 302 is designed to provide Keywords that can apply to multiple Categories at a highest level and for very specific Keywords at a lower level. For example, a Credit Card company may be offered Keywords in the "Personal Finance" Super-Category, such as Keyword "credit", and then be offered Keywords in the Sub-Categories "Personal Finance-Credit Cards" and "Personal Finance-Credit Cards-Low Rate", such as the Keywords "credit card" and "low rate credit card", respectively. This is but one organizational example, and it will be appreciated that the such categorization be adjusted according to need. 14:25-36 "In a specific embodiment, the MAIN application 520 may be configured to analyze a selected document for keywords, categories and/or super categories in order to find a match for an appropriate pop-up advertisement or window to be displayed." 27:30-34. See Figs. 3, 23. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Wieser, p. 18, line 24 to p. 19, line 1; p. 22, lines 3-15; Fig. 15

	Rhodes, 11:62-12:4, 12:41-49, 13:1-8, 13:15-19 Donaldson, 15:38-46, 17:17-19, Fig. 8c HyPursuit, p. 180, col. 2; p. 184, col. 1; Fig. 3 Horowitz, 5:65-6:4 Finkelstein, p. 410 OracleText White Paper, p. 19 Pretschner, p. 2
Claim 18	
18. An article of manufacture for use in a computer system, comprising:	"Generally, the various techniques of the present invention may be implemented on software and/or hardware." 45:37-38. See Fig. 22.
a memory;	"A software or software/hardware hybrid implementation of the various technique of this invention may be implemented on a general-purpose programmable machine selectively activated or reconfigured by a computer program stored in memory." 45:46-50. See Fig. 22. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 4, line 26 to p. 5 line 1 Rhodes, Fig. 1, 9:18-34 Black, 4:58-67, 1:50-57, 4:30-43. Donaldson, Fig. 4, 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189

	Finkelstein, p. 406 Horowitz, 5:25-30, Figs. 2, 9 Fleming, Fig. 1, 4:34-41 Ford, 3:34-39, Fig. 3 Kraft, 1:23-27, Fig. 1 Apte, Figs. 3, 4, claim 33 Pretschner, p. 7
instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	See Chart for Claim 1 (above). "A software or software/hardware hybrid implementation of the various technique of this invention may be implemented on a general-purpose programmable machine selectively activated or reconfigured by a computer program stored in memory." 45:46-50. See Fig. 22. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 5, lines 8-12, p. 2 lines 16-19; p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22; p. 14, line 19 to p. 15 line 5 Rhodes, 9:18-34, 10:1-17 Black, 4:58-67, 1L50-57 Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 col. 2. Finkelstein, p. 406 Horowitz, 5:30-33, Fig. 2 at 103

	Fleming, 4:34-41 Ford, 13:10-19 Kraft, 1:23-27 Pretschner, pp. 1, 7
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	See Chart for Claim 1 (above).
automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	See Chart for Claim 1 (above).
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).
Claim 19	
19. The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in	See Chart for Claim 2 (above).

the selected document content.	

Second Supplemental Chart B-13

Claim chart for Finkelstein et al., *Placing Search in Context: The Concept Revisited*, Proc. of the 10th International World Wide Web Conference (May 1-5, 2001)

as prior art to

U.S. Patent No. 6,778,979

This chart is based upon Xerox's apparent construction of the claims, and is not an admission that those constructions are correct or appropriate.

'979 Patent	**Finkelstein et al., *Placing Search in Context: The Concept Revisited*, Proc. of the 10th International World Wide Web Conference (May 1-5, 2001)**
Claim 1	
A method for automatically generating a query from selected document content, comprising:	"In the IntelliZap system we developed, search is initiated from a text query marked by the user in a document she views, and is guided by the text surrounding the marked query in that document ('the context'). The context-guided information retrieval process involves semantic keyword extraction and clustering to automatically generate new, augmented queries. The latter are submitted to a host of general and domain-specific search engines." Page 406. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0001, 0002, 0004, 0015, Fig. 4. Black, Abstract, 1:64-67, 2:43, 1:47-48, Fig. 1, 4:18-20, 1:50-54, 2:10-14, 2:57-61, 2:42-45, 2:26-29, 3:29-35, 3:37-45, 3:55-57 Wieser, p. 5, lines 8-12, p. 2 lines 16-19 Rhodes, 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Horowitz, 2:42-51, 3:20-23 Fleming, Abstract, Fig. 2 Ford, 5:38, 5:30-39

	Kraft, Abstract Apte, 6:60-67 Henkin, 2:42-49, 44:8-19 Pretschner, p. 1
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	"The classification algorithm classifies the context to a limited number of high-level domains[11] (e.g., medicine or law). A probabilistic analysis determines the amount of similarity between the domain signatures and the query context. The *a priori* assignment of search engines to domains is performed offline. Some of the search engines (such as AltaVista) allow limiting the search to a specific category. In such cases, categorizing the query in order to further constrain the search usually yields superior results." Page 410. FN 11: "Currently, nine domains are defined, each of which is mapped to two or three search engines." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Mase, Abstract, Fig. 1, p. 377-378 Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55 Donaldson, 15:38-46, 17:17-19, 19:51-56, Fig. 8c HyPursuit, Fig. 3, p. 181, 184 Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40 Apte, 9:33-45

	Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19 Pretschner, pp. 1-2
automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"The algorithm utilizes the semantic network to extract keywords from the context surrounding the user-selected text. These keywords are added to the text to form an augmented query. . . ." Page 410 "The IntelliZap system has three main components based on the semantic network: 1. Extracting keywords from the captured text and context. …" Page 410. "The context may include the sentence containing the query word or phrase, a few sentences surrounding the query term, the paragraph in which it resides, or even the whole document." Page 408. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0015 Wieser, p. 14, line 19 to p. 15 line 5 Rhodes, 13:19-34, 1:56-2:6, 2:15-54 Black, 2:26-29, 2:42-45, 3:29-35 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7 Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Ford, 11:4-11, 12:52-63 Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B OracleText White Paper, p. 8, 9

	Pretschner, p. 2
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	"[W]e attempt to classify the captured context in order to select domain-specific search engines that stand a good chance of providing more specialized results. The classification algorithm classifies the context to a limited number of high-level domains (e.g., medicine or law). A probabilistic analysis determines the amount of similarity between the domain signatures and the query context." Page 410 (footnote omitted). "The IntelliZap system has three main components based on the semantic network: … 2. High-level classification of the query to a small set of predefined domains. …" Page 410. "The context may include the sentence containing the query word or phrase, a few sentences surrounding the query term, the paragraph in which it resides, or even the whole document." Page 408. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Mase, p. 377-379, Fig. 1 Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 5:12-28 Black, 2:10-18, 2:57-61 Donaldson, Figs. 9b and 9c, 18:47-63, 20:3-18, 20:19-29 HyPursuit, p. 181, 182, 185, 186, 191 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17

	Kraft, 11:16-40 Apte, 9:33-37 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A OracleText White Paper, p. 11, 18, 19 Pretschner, p. 2
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	"The algorithm utilizes the semantic network to extract keywords from the context surrounding the user-selected text. These keywords are added to the text to form an augmented query. . . ." Page 410 "[W]e attempt to classify the captured context in order to select domain-specific search engines that stand a good chance of providing more specialized results. . . . Some of the search engines (such as AltaVista) allow limiting the search to a specific category. In such cases, categorizing the query in order to further constrain the search usually yields superior results." Page 410. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015, Fig. 4. Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Rhodes, 5:12-28 Black, 3:37-45, 3:55-57 Donaldson, 18:47-54, Figs. 9b and 9c, 18:55-63, 20:3-29 HyPursuit, p. 181, 182, 184, 185, 186 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8 Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17 Kraft, 13:40-50, 13:64-67, 14:43-52

	Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44 Pretschner, pp. 1-2
Claim 5	
5. The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	"Some of the search engines (such as AltaVista) allow limiting the search to a specific category. In such cases, categorizing the query in order to further constrain the search usually yields superior results." Page 410. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Wieser, p. 18, line 24 to p. 19, line 1; p. 22, lines 3-15; Fig. 15 Rhodes, 11:62-12:4, 12:41-49, 13:1-8, 13:15-19 Donaldson, 15:38-46, 17:17-19, Fig. 8c HyPursuit, p. 180, col. 2; p. 184, col. 1; Fig. 3 Horowitz, 5:65-6:4 Henkin, 3:1-3, 14:4-9, 14:25-36, 27:30-34, Fig. 3, Fig. 23 OracleText White Paper, p. 19 Pretschner, p. 2
Claim 18	
18. An article of manufacture for use in a computer system, comprising:	"Our system (named IntelliZap) is based on the client-server paradigm, where a client application running on user's computer captures the context around the text highlighted by the user. The server-based algorithms analyze the context, selecting most important context words and performing word sense disambiguation, and then prepare a set of augmented queries for subsequent search." Page 406.
a memory;	"Our system (named IntelliZap) is based on the client-server paradigm, where a client application running on user's computer captures the context around the text highlighted by the user. The server-based

	algorithms analyze the context, selecting most important context words and performing word sense disambiguation, and then prepare a set of augmented queries for subsequent search." Page 406. The client and server computers inherently include a memory. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 4, line 26 to p. 5 line 1 Rhodes, Fig. 1, 9:18-34 Black, 4:58-67, 1:50-57, 4:30-43. Donaldson, Fig. 4, 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 Horowitz, 5:25-30, Figs. 2, 9 Fleming, Fig. 1, 4:34-41 Ford, 3:34-39, Fig. 3 Kraft, 1:23-27, Fig. 1 Apte, Figs. 3, 4, claim 33 Henkin, 45:37-38, Fig. 22 Pretschner, p. 7
instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	See Chart for Claim 1 (above). "Our system (named IntelliZap) is based on the client-server paradigm, where a client application running on user's computer captures the context around the text highlighted by the user. The server-based algorithms analyze the context, selecting most important context words and performing word sense disambiguation, and then prepare a set of augmented queries for subsequent search." Page 406.

	The client and server computers inherently store the IntelliZap software in memory in order to make it executable. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 5, lines 8-12, p. 2 lines 16-19; p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22; p. 14, line 19 to p. 15 line 5 Rhodes, 9:18-34, 10:1-17 Black, 4:58-67, 1L50-57 Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 col. 2. Horowitz, 5:30-33, Fig. 2 at 103 Fleming, 4:34-41 Ford, 13:10-19 Kraft, 1:23-27 Henkin, 45:46-50, Fig. 22 Pretschner, pp. 1, 7
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	See Chart for Claim 1 (above).

automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	See Chart for Claim 1 (above).
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).

Second Supplemental Chart B-14

Claim chart of PCT Application Pub. No. WO 01/44992 to Weiser et al. ("Context Matching System and Method") taken in view of

U.S. Patent No. 7,089,236 to Stibel ("Search Engine Interface")

as prior art to

U.S. Patent No. 6,778,979

This chart is based upon Xerox's apparent construction of the claims, and is not an admission that those constructions are correct or appropriate.

'979 Patent	WIESER AND STIBEL
Claim 2	
2. The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	WIESER, p. 5, lines 8-12 (the client 12 generates a query composed of all or a portion of a document (e.g., a web page) and sends the query to a match server 14). WIESER, p. 2, lines 16-19: "Another object of the present invention is to provide a system and method which automatically and contextually matches products, advertisements or other content (hereinafter referred to as 'offers') to the content on a web page that a user has selected in real-time" Stibel, Table 2: TABLE 2 Search Terms Generated by the Relational Knowledgebase for User Query "java" java (required), coffee (meaning), cafe au lait, cafe noir, demitasse, decaffeinated coffee, decaf, espresso, capuccino, coffee capuccino, iced coffee, ice coffee, mocha, Turkish coffee, cafe royale, beverage, coffee royal, Irish coffee Stibel, 2:42-47: "The systems and methods described herein include systems that, inter alia, operate as a front end to a database search engine or engines, and act to process a user query to generate a new search request that will more effectively retrieve information from the database that is relevant to the query of the user. To this end, in one embodiment the systems can be realized as computer programs

'979 Patent	WIESER AND STIBEL
	that act as front ends to databases. The front ends may include a user interface that is presented to a user and which may prompt the user to enter one or more key phrases that are representative of a user search request. The user interface may collect the key phrases provided by the user and may analyze these key phrases to identify at least one meaning that may be associated with this user search request. The systems may then process the user search request and the identified meaning to generate an expanded search request that may be represented as a compound search string, such as a Boolean search string, or other logical string. This compound search string may then be processed to create one or more expanded user queries that may be presented to a search engine to collect from a search engine information that is relevant to the interest of the user." Stibel, 10:30-36: "Referring to FIG. 1 it may be seen that the query engine 14 may employ the knowledge base 16 to identify meanings that may be associated with the user query provided by the user interface 12. To this end, the query engine may employ words and phrases from the user query to query the knowledge base 15 and collect therefrom one or more meanings that may be associated with the user query." Stibel, 9:21-43: "In one particular embodiment, the knowledge base 16 comprises two types of entities: Senses and Words. The knowledge base 16 can store a Sense as a data structure that has associated with it five items. … As shown in FIG. 3, these items can include a generalization member that is representative of the Sense that is the closest term that represents a generalization for the associated Sense; a specific meaning that can be an optional string representative of the meaning to be displayed to the user for the Sense; related terms each being representative of a word form or Sense or Word meaning that is likely to appear on a web page when a user is querying for information associated with this Sense, Sense words which may be representative of a list of words from the knowledge base 16 that have this Sense as one of their meanings; and a flag member that may contain flags associated with this Sense." Stibel, 11:56-66: "In either embodiment, after the user has selected a meaning to be associated with the user query the Query engine 14 may then build an expanded search query by employing the related terms, depicted in FIG. 3, to amend the user query to one more suited for identifying documents associated with the interest of the client. For example, if the user enters "java" as the user query and selects "coffee" as its meaning, the knowledge base 16 will retrieve

'979 Patent	WIESER AND STIBEL
	a number of related words (e.g., mocha, espresso) and append these words, along with the chosen meaning, to the original user query." Stibel, 12:13-21: "These related words provide additional keyphrases that may be added, such as by Boolean logic operators or by other logical operators, to the keyphrases of the user query. For example, the original user query "java," can now be expanded to java+coffee+espresso(W1)+beverage(W2)+ (NOT)programming. As expanded, the query now includes terms that are selected to increase the likelihood that an Internet search engine will return a meaningful hit list." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Rhodes, 10:42-51, 11:10-31, 12:41-49, 13:1-8, 13:15-34, 13:42-47 Black, 1:64-2:9, 3:37-45, 3:55-57 Donaldson, 18:47-63, 20:3-29, Fig. 9b, Fig. 9c HyPursuit, p. 186, col. 2 Horowitz, 7:19-25, 7:46-49 Kraft, 10:32-36, 12:13-21, 13:2-50 Henkin, 27:65-28:5 Pretschner, p. 2
Claim 19	
19. The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	*See* claim 2 above.

Second Supplemental Chart B-15

Claim chart of "Syskill & Webert: Identifying interesting web sites" (AAAI-96 Proceedings, Copyright 1996)

as prior art to

U.S. Patent No. 6,778,979

This chart is based upon Xerox's apparent construction of the claims, and is not an admission that those constructions are correct or appropriate.

'979 Patent	SYSKILL
Claim 3	
3. The method according to claim 2, wherein the number of terms added is limited to a predefined number.	SYSKILL, p. 56, col. 2:3-6: "Since LYCOS cannot accept very long queries, we use the 7 most discriminating words that are found in a higher proportion of hot pages than all pages and the 7 most commonly occurring words as a query." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Rhodes, 10:42-51, 11:10-31, 12:41-49, 13:1-8, 13:15-34, 13:42-47 Black, 1:64-2:9, 3:37-45, 3:55-57 Donaldson, 18:47-63, 20:3-29, Fig. 9b, Fig. 9c HyPursuit, p. 186, col. 2 Horowitz, 7:19-25, 7:46-49 Kraft, 10:32-36, 12:13-21, 13:2-50 Henkin, 27:65-28:5 Stibel, 2:42-47, 9:21-43, 10:30-36, 11:56-66, 12:13-21, Table 2 Pretschner, p. 2

Second Supplemental Chart B-16

Claim chart of U.S. Patent No. 6,546,386 to Black et al. ("Brilliant Query System") taken in view of U.S. Patent No. 7,225,180 to Donaldson et al. ("Filtering Search Results") and U.S. Patent No. 7,089,236 to Stibel ("Search Engine Interface")

as prior art to

U.S. Patent No. 6,778,979

This chart is based upon Xerox's apparent construction of the claims, and is not an admission that those constructions are correct or appropriate.

'979 Patent	BLACK, DONALDSON, AND STIBEL
Claim 2	
2. The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	Black, 3:37-45: "Once the keywords have been selected and the hook for a body of text has been determined or automatically generated, the searches are created by generating a link for every keyword extracted from the body of text and combining it with the hook in a search that results in a result set that is the logical intersection of the results generated by the hook and the keyword. Basically, each entry in the list of search results must contain both the hook and the keyword and not just one or the other." Black, 3:55-57: "After the hook and the keywords have been established the query is conducted selecting one of the hook-keyword sets." Donaldson, 18:47-54: "Classifying the search term (922) generally includes classifying the received search term among one or more categories, with a first category and a second category being described and shown for illustrative purposes. If several search terms are grouped as a single string, the search terms may be collectively classified as a single string based on the grouping of search terms, or they may be classified individually based on each individual search term." Donaldson, Figs. 9b and 9c, 18:55-63: "Comparing the search terms (step 924) generally includes comparing the search term to first electronic information within a first electronic information store when the search term is classified within the first category. By contrast, comparing the search term (step 926) generally includes comparing the search term to the second electronic information within the second electronic information store to determine whether matches exist when the search term is classified within the second category." Donaldson, 20:3-18: "The following describes an example applying

'979 Patent	BLACK, DONALDSON, AND STIBEL
	the described search methods of FIG. 9b to this implementation. A user of a client system enters a search term (step 910). The search term is classified as either being offensive or non-offensive (step 922). If the term is classified as being non-offensive, then only the contents of the first electronic information store are searched (924) and results from the search are communicated for display to the user (step 930). In this example, the first electronic information store only contains contents that previously have been classified as non-offensive. If the search term entered by the user is classified as being offensive, the contents of either the second electronic information store or both the first and second electronic information stores are searched (step 926) and the results are communicated for display to the user (step 930)." Donaldson, 20:19-29: "The described filtering of results between offensive content and non-offensive content based on the classification of the search term may allow a web host to implement a parental type of control in determining what search results are displayed to the user. Because the offensive and non-offensive contents are stored in different electronic information stores, the ability to restrict access is enhanced. For instance, parental control can be exercised by blocking the access of a user to one or more electronic information stores. Other forms of data filtering also are enabled through this process and related techniques." Stibel, Table 2: TABLE 2 Search Terms Generated by the Relational Knowledgebase for User Query "java" java (required) coffee (meaning) cafe au lait cafe noir demitasse decaffeinated coffee decaf espresso capuccino coffee capuccino iced coffee ice coffee mocha Turkish coffee cafe royale beverage coffee royal Irish coffee Stibel, 2:42-47: "The systems and methods described herein include systems that, inter alia, operate as a front end to a database search engine or engines, and act to process a user query to generate a new search request that will more effectively retrieve information from the database that is relevant to the query of the user. To this end, in one embodiment the systems can be realized as computer programs

'979 Patent	BLACK, DONALDSON, AND STIBEL
	that act as front ends to databases. The front ends may include a user interface that is presented to a user and which may prompt the user to enter one or more key phrases that are representative of a user search request. The user interface may collect the key phrases provided by the user and may analyze these key phrases to identify at least one meaning that may be associated with this user search request. The systems may then process the user search request and the identified meaning to generate an expanded search request that may be represented as a compound search string, such as a Boolean search string, or other logical string. This compound search string may then be processed to create one or more expanded user queries that may be presented to a search engine to collect from a search engine information that is relevant to the interest of the user." Stibel, 10:30-36: "Referring to FIG. 1 it may be seen that the query engine 14 may employ the knowledge base 16 to identify meanings that may be associated with the user query provided by the user interface 12. To this end, the query engine may employ words and phrases from the user query to query the knowledge base 15 and collect therefrom one or more meanings that may be associated with the user query." Stibel, 9:21-43: "In one particular embodiment, the knowledge base 16 comprises two types of entities: Senses and Words. The knowledge base 16 can store a Sense as a data structure that has associated with it five items. … As shown in FIG. 3, these items can include a generalization member that is representative of the Sense that is the closest term that represents a generalization for the associated Sense; a specific meaning that can be an optional string representative of the meaning to be displayed to the user for the Sense; related terms each being representative of a word form or Sense or Word meaning that is likely to appear on a web page when a user is querying for information associated with this Sense, Sense words which may be representative of a list of words from the knowledge base 16 that have this Sense as one of their meanings; and a flag member that may contain flags associated with this Sense." Stibel, 11:56-66: "In either embodiment, after the user has selected a meaning to be associated with the user query the Query engine 14 may then build an expanded search query by employing the related terms, depicted in FIG. 3, to amend the user query to one more suited for identifying documents associated with the interest of the client. For example, if the user enters "java" as the user query and selects "coffee" as its meaning, the knowledge base 16 will retrieve

'979 Patent	BLACK, DONALDSON, AND STIBEL
	a number of related words (e.g., mocha, espresso) and append these words, along with the chosen meaning, to the original user query." Stibel, 12:13-21: "These related words provide additional keyphrases that may be added, such as by Boolean logic operators or by other logical operators, to the keyphrases of the user query. For example, the original user query "java," can now be expanded to java+coffee+espresso(W1)+beverage(W2)+ (NOT)programming. As expanded, the query now includes terms that are selected to increase the likelihood that an Internet search engine will return a meaningful hit list." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Rhodes, 10:42-51, 11:10-31, 12:41-49, 13:1-8, 13:15-34, 13:42-47 HyPursuit, p. 186, col. 2 Horowitz, 7:19-25, 7:46-49 Kraft, 10:32-36, 12:13-21, 13:2-50 Henkin, 27:65-28:5 Pretschner, p. 2
Claim 19	
19. The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	*See* claim 2 above.

Second Supplemental Chart B-17

Claim chart of "Ontology Based Personalized Search" by Pretschner and Gauch (Proceedings of the 11th IEEE Int'l Conf. on Tools with Artificial Intelligence, pp. 391-98, Nov. 1999)

as prior art to

U.S. Patent No. 6,778,979

This chart is based upon Xerox's apparent construction of the claims, and is not an admission that those constructions are correct or appropriate.

'979 Patent	PRETSCHNER
Claim 1	
1. A method for automatically generating a query from selected document content, comprising:	Pretschner, p. 1: "Or, the system could navigate through the Web on its own and notify the user if it found a page or site of presumed interest." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0001, 0002, 0004, 0015, Fig. 4. Black, Abstract, 1:64-67, 2:43, 1:47-48, Fig. 1, 4:18-20, 1:50-54, 2:10-14, 2:57-61, 2:42-45, 2:26-29, 3:29-35, 3:37-45, 3:55-57 Wieser, p. 5, lines 8-12, p. 2 lines 16-19 Rhodes, 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein, p. 406 Horowitz, 2:42-51, 3:20-23 Fleming, Abstract, Fig. 2 Ford, 5:38, 5:30-39 Kraft, Abstract Apte, 6:60-67 Henkin, 2:42-49, 44:8-19
defining an organized	Pretschner, p. 1: "Browsing is usually done by clicking through a

'979 Patent	PRETSCHNER
classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	hierarchy of subjects until the area of interest has been reached. The corresponding node then provides the user with links to related websites." Pretschner, p. 2: "User interests are inferred by analyzing the web pages the user visits. For this purpose, it is necessary to determine the content, or characterize, these surfed pages. This is done by using a hierarchy of concepts, or rather ontology. This ontology is based on a publicly accessible browsing hierarchy. For this paper, the Magellan hierarchy, which is comprised of approximately 4,400 nodes, has been mirrored (magellan.excite.com). The nodes of the ontology are labelled with the names of the nodes in the browsing hierarchy. The semantics of the edges of this hierarchy is not formally specified; in most cases, they correspond to a specialization relation (super-/subconcept)." Pretschner, p. 2: "For each of the surfed pages a keyword vector is calculated. This page vector is compared with the keyword vectors associated with every node to calculate similarities. The nodes with the top matching vectors are assumed to be most related to the content of the surfed page. The accuracy of this text categorization algorithm was validated in [30]." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Mase, Abstract, Fig. 1, p. 377-378 Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55 Donaldson, 15:38-46, 17:17-19, 19:51-56, Fig. 8c HyPursuit, Fig. 3, p. 181, 184 Finkelstein, p. 410 Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60

'979 Patent	PRETSCHNER
	Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40 Apte, 9:33-45 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19
automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	Pretschner, p. 2: "Documents as well as superdocuments are represented as weighted keyword vectors using the vector space model [23]. The weights are based on term frequencies and inverted document frequencies: It is assumed that multiple occurrences of a word indicate that its meaning contributes to the content of the document more than less frequent terms. However, words that occur with a very high overall frequency (i.e., in the collection of documents in question) do not discriminate between documents within this collection." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0015 Wieser, p. 14, line 19 to p. 15 line 5 Rhodes, 13:19-34, 1:56-2:6, 2:15-54 Black, 2:26-29, 2:42-45, 3:29-35 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7 Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Ford, 11:4-11, 12:52-63 Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B

'979 Patent	PRETSCHNER
	OracleText White Paper, p. 8, 9
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	Pretschner, p.2: "The goal of OBIWAN is to investigate a novel content-based approach to distributed information retrieval. Websites are clustered into regions. Examples for clustering criteria include but are not restricted to content, geographic location, and association with a specific company. Regions are clustered into super regions, super regions into hyper regions, etc., thus forming a hierarchy of regions. A node of this hierarchy can be browsed by simultaneously browsing its child nodes." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Mase, p. 377-379, Fig. 1 Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 5:12-28 Black, 2:10-18, 2:57-61 Donaldson, Figs. 9b and 9c, 18:47-63, 20:3-18, 20:19-29 HyPursuit, p. 181, 182, 185, 186, 191 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17 Kraft, 11:16-40 Apte, 9:33-37 Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A OracleText White Paper, p. 11, 18, 19
automatically formulating the	Pretschner, p. 1: "Or, the system could navigate through the Web on

'979 Patent	PRETSCHNER
query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	its own and notify the user if it found a page or site of presumed interest." Pretschner, p. 2: "In terms of searching, queries are brokered within one node by deciding which child nodes are the most promising candidates for the retrieval process. This is done by determining the content of the query and using a sitemap containing information about the content of every node in the (sub)hierarchy: the query is brokered to those nodes with a content that best matches the content of the query. The results of the child nodes are then merged and returned to the parent node or, eventually, to the initiator of the query." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015, Fig. 4. Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Rhodes, 5:12-28 Black, 3:37-45, 3:55-57 Donaldson, 18:47-54, Figs. 9b and 9c, 18:55-63, 20:3-29 HyPursuit, p. 181, 182, 184, 185, 186 Finkelstein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8 Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17 Kraft, 13:40-50, 13:64-67, 14:43-52 Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44
Claim 2	
2. The method according to claim	Pretschner, p. 2: "In terms of searching, queries are brokered within

'979 Patent	PRETSCHNER
1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	one node by deciding which child nodes are the most promising candidates for the retrieval process. This is done by determining the content of the query and using a sitemap containing information about the content of every node in the (sub)hierarchy: the query is brokered to those nodes with a content that best matches the content of the query. The results of the child nodes are then merged and returned to the parent node or, eventually, to the initiator of the query." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Rhodes, 10:42-51, 11:10-31, 12:41-49, 13:1-8, 13:15-34, 13:42-47 Black, 1:64-2:9, 3:37-45, 3:55-57 Donaldson, 18:47-63, 20:3-29, Fig. 9b, Fig. 9c HyPursuit, p. 186, col. 2 Horowitz, 7:19-25, 7:46-49 Kraft, 10:32-36, 12:13-21, 13:2-50 Henkin, 27:65-28:5 Stibel, 2:42-47, 9:21-43, 10:30-36, 11:56-66, 12:13-21, Table 2
Claim 5	
5. The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	Pretschner, p.2: "The goal of OBIWAN is to investigate a novel content-based approach to distributed information retrieval. Websites are clustered into regions. Examples for clustering criteria include but are not restricted to content, geographic location, and association with a specific company. Regions are clustered into super regions, super regions into hyper regions, etc., thus forming a hierarchy of regions. A node of this hierarchy can be browsed by simultaneously browsing its child nodes." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015

'979 Patent	PRETSCHNER
	Wieser, p. 18, line 24 to p. 19, line 1; p. 22, lines 3-15; Fig. 15 Rhodes, 11:62-12:4, 12:41-49, 13:1-8, 13:15-19 Donaldson, 15:38-46, 17:17-19, Fig. 8c HyPursuit, p. 180, col. 2; p. 184, col. 1; Fig. 3 Horowitz, 5:65-6:4 Henkin, 3:1-3, 14:4-9, 14:25-36, 27:30-34, Fig. 3, Fig. 23 Finkelstein, p. 410 OracleText White Paper, p. 19
Claim 10	
10. The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary.	Pretschner, p. 2: "Documents as well as superdocuments are represented as weighted keyword vectors using the vector space model [23]. The weights are based on term frequencies and inverted document frequencies: It is assumed that multiple occurrences of a word indicate that its meaning contributes to the content of the document more than less frequent terms. However, words that occur with a very high overall frequency (i.e., in the collection of documents in question) do not discriminate between documents within this collection. "For each of the surfed pages a keyword vector is calculated. This page vector is compared with the keyword vectors associated with every node to calculate similarities. The nodes with the top matching vectors are assumed to be most related to the content of the surfed page. The accuracy of this text categorization algorithm was validated in [30]." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0015 Wieser, p. 18, line 24 to p. 19, line 1; p. 22, lines 3-15; Fig. 15 Rhodes, 4:45-55 Mase, p. 377, col. 2; p. 379, col. 1

'979 Patent	PRETSCHNER
	Donaldson, 16:5-9 HyPursuit, p. 185, col. 1 Horowitz, 3:37-39, 5:50-55 Fleming, 6:45-60, 12:3-11, Fig. 9 OracleText White Paper, p. 19
Claim 18	
18. An article of manufacture for use in a computer system, comprising: a memory;	Pretschner, p. 7: "Future work includes the integration of the system into a web browser (right now, cache folders are analyzed) which will allow for more accurate interest detection if other interactions such as scrolling behavior are monitored." Pretschner, p. 7: "In terms of privacy, the existing system stores the profile on the user's machine." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 4, line 26 to p. 5 line 1 Rhodes, Fig. 1, 9:18-34 Black, 4:58-67, 1:50-57, 4:30-43. Donaldson, Fig. 4, 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 Finkelstein, p. 406 Horowitz, 5:25-30, Figs. 2, 9 Fleming, Fig. 1, 4:34-41 Ford, 3:34-39, Fig. 3

'979 Patent	PRETSCHNER
	Kraft, 1:23-27, Fig. 1 Apte, Figs. 3, 4, claim 33 Henkin, 45:37-38, Fig. 22
instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	Pretschner, p. 1: "Or, the system could navigate through the Web on its own and notify the user if it found a page or site of presumed interest." Pretschner, p. 7: "Future work includes the integration of the system into a web browser (right now, cache folders are analyzed) which will allow for more accurate interest detection if other interactions such as scrolling behavior are monitored." Pretschner, p. 7: "In terms of privacy, the existing system stores the profile on the user's machine." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014. Mase, p. 382 col. 1. Wieser, p. 5, lines 8-12, p. 2 lines 16-19; p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22; p. 14, line 19 to p. 15 line 5 Rhodes, 9:18-34, 10:1-17 Black, 4:58-67, 1L50-57 Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 col. 2. Finkelstein, p. 406 Horowitz, 5:30-33, Fig. 2 at 103 Fleming, 4:34-41 Ford, 13:10-19

'979 Patent	PRETSCHNER
	Kraft, 1:23-27 Henkin, 45:46-50, Fig. 22
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	Pretschner, p. 1: "Browsing is usually done by clicking through a hierarchy of subjects until the area of interest has been reached. The corresponding node then provides the user with links to related websites." Pretschner, p. 2: "User interests are inferred by analyzing the web pages the user visits. For this purpose, it is necessary to determine the content, or characterize, these surfed pages. This is done by using a hierarchy of concepts, or rather ontology. This ontology is based on a publicly accessible browsing hierarchy. For this paper, the Magellan hierarchy, which is comprised of approximately 4,400 nodes, has been mirrored (magellan.excite.com). The nodes of the ontology are labelled with the names of the nodes in the browsing hierarchy. The semantics of the edges of this hierarchy is not formally specified; in most cases, they correspond to a specialization relation (super-/subconcept)." Pretschner, p. 2: "For each of the surfed pages a keyword vector is calculated. This page vector is compared with the keyword vectors associated with every node to calculate similarities. The nodes with the top matching vectors are assumed to be most related to the content of the surfed page. The accuracy of this text categorization algorithm was validated in [30]." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Mase, p. 382 col. 1. Wieser, p. 15, lines 8-11; p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 2:15-24, 4:20-27, 4:45-55 Donaldson, Fig. 4 and 13:62-14:6, 6:53-7:10, 3:45-58 HyPursuit, p. 189 Finkelstein, p. 410

'979 Patent	PRETSCHNER
	Horowitz, 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56, Fig. 5 Fleming, 1:44-53, 2:12-26, 6:24-44, 6:45-60 Ford, 11:4-17 Kraft, 5:55-6:10, 11:16-40 Apte, 6:60-67 Henkin, 14:4-9, 9:20-25, 27:14-29, 29:25-29, Figs 3, 8, 23, 25C 13:25-55 OracleText White Paper, p. 11, 18, 19
automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	Pretschner, p. 2: "Documents as well as superdocuments are represented as weighted keyword vectors using the vector space model [23]. The weights are based on term frequencies and inverted document frequencies: It is assumed that multiple occurrences of a word indicate that its meaning contributes to the content of the document more than less frequent terms. However, words that occur with a very high overall frequency (i.e., in the collection of documents in question) do not discriminate between documents within this collection." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0013, 0014, 0015 Wieser, p. 14, line 19 to p. 15 line 5 Rhodes, 13:19-34, 1:56-2:6, 2:15-54 Black, 1:50-57 and 4:58-67 Finkestein, p. 410, 408 Horowitz, 6:49-55, 7:46-49, 9:27-28, 12:27-32, Figs. 4c-4d, 5, 7 Fleming, 3:5-8, 3:14-23, 5:62-6:7, 9:26-29, Figs. 2, 5, 8, 10B Ford, 11:4-11, 12:52-63

'979 Patent	PRETSCHNER
	Kraft, 5:55-6:14, 9:34-36 Apte, 9:24-32 Henkin, 4:56-64, 7:22-25, 27:14-29, 27:65-28:5, Figs. 16A, 16B OracleText White Paper, p. 8, 9
automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	Pretschner, p.2: "The goal of OBIWAN is to investigate a novel content-based approach to distributed information retrieval. Websites are clustered into regions. Examples for clustering criteria include but are not restricted to content, geographic location, and association with a specific company. Regions are clustered into super regions, super regions into hyper regions, etc., thus forming a hierarchy of regions. A node of this hierarchy can be browsed by simultaneously browsing its child nodes." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014, 0015 Mase, para. 382 col. 1. Wieser, p. 17, lines 15-32; p. 18, lines 16-22 Rhodes, Table 2, 5:12-28 Donaldson, 7:22-32, 7:34-36, 13:62-14:6 HyPursuit, p. 189 Finkestein, p. 410, 408 Horowitz, 3:25-35, 7:66-8:5, Figs. 6-8 Fleming, 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 Ford, 11:4-17 Kraft, 11:16-40 Apte, 9:33-37

'979 Patent	PRETSCHNER
	Henkin, 27:30-34, 27:14-29, 27:46-51, Fig 16A OracleText White Paper, p. 11, 18, 19
automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	Pretschner, p. 1: "Or, the system could navigate through the Web on its own and notify the user if it found a page or site of presumed interest." Pretschner, p. 2: "In terms of searching, queries are brokered within one node by deciding which child nodes are the most promising candidates for the retrieval process. This is done by determining the content of the query and using a sitemap containing information about the content of every node in the (sub)hierarchy: the query is brokered to those nodes with a content that best matches the content of the query. The results of the child nodes are then merged and returned to the parent node or, eventually, to the initiator of the query." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Reader, para. 0014, 0015, Fig. 4 Mase, p. 382, col. 1 Wieser, p. 17, lines 15-32; p. 18, lines 14-22 Rhodes, 5:12-28 Donaldson, 7:22-32, 7:34-36, 13:62-14:6. HyPursuit, p. 189 Finkestein, p. 410 Horowitz, 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18, Figs. 5-6, 8 Fleming, 4:23-29, 7:51-8:4, 10:34-38, 15:19-23, Figs. 2, 7, 10B Ford, 12:2-7, 11:4-17 Kraft, 13:40-50, 13:64-67, 14:43-52

'979 Patent	PRETSCHNER
	Apte, 4:20-22, 9:38-45 Henkin, 27:46-28:5, 28:30-44
Claim 19	
19. The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	Pretschner, p. 2: "In terms of searching, queries are brokered within one node by deciding which child nodes are the most promising candidates for the retrieval process. This is done by determining the content of the query and using a sitemap containing information about the content of every node in the (sub)hierarchy: the query is brokered to those nodes with a content that best matches the content of the query. The results of the child nodes are then merged and returned to the parent node or, eventually, to the initiator of the query." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Rhodes, 10:1-17 Black 1:50-57, 4:58-67 Donaldson 7:22-32, 13:62-14:6 HyPursuit, p. 189, col. 2 Horowitz, 7:19-25, 7:46-49 Kraft, 10:32-36, 12:13-21, 13:2-50 Henkin, 3:1-3, 14:4-9, 14:25-36, 27:30-34, Fig. 3, Fig. 23 Stibel, 2:42-47, 9:21-43, 10:30-36, 11:56-66, 12:13-21, Table 2

EXHIBIT 3

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

XEROX CORPORATION, Plaintiff, v. GOOGLE, INC., YAHOO! INC., RIGHT MEDIA INC., RIGHT MEDIA LLC, YOUTUBE, INC., and YOUTUBE, LLC, Defendants.	)	C.A. No. 10-136 (LPS) **HIGHLY CONFIDENTIAL – OUTSIDE COUNSEL ONLY**

YAHOO! INC. AND RIGHT MEDIA LLC'S FOURTH SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 7 OF XEROX'S FIRST SET OF INTERROGATORIES

Defendants and Counter-Claim Plaintiffs Yahoo! Inc. and Right Media LLC (collectively, "Yahoo!") provide the following fourth supplemental response to Interrogatory No. 7 of Plaintiff Xerox Corporation's ("Xerox") April 23, 2010 First Set of Interrogatories.[1] These supplemental responses are made in light of ongoing discovery and are based on information presently known to Yahoo!, which reserves the right to supplement or modify this supplemental response based on the discovery of additional or different information and/or in light of expert opinion and/or the Court's claim construction. This supplemental response is provided without the benefit of the Court's claim construction and with the understanding that a range of claim construction positions may potentially be advanced by the parties and/or adopted by the Court. This supplemental response therefore should not be deemed to admit the correctness or incorrectness of any construction of any limitation of any asserted claim.

[1] Right Media LLC responds on its own behalf and as the successor in interest to Right Media Inc., which no longer exists.

SUPPLEMENTAL RESPONSES TO INTERROGATORIES

INTERROGATORY NO. 7:

If you contend that any claim of the Patents in Suit is invalid and/or unenforceable, specify each claim that you contend is invalid and/or unenforceable and describe in full for each such claim the basis for your contention, identifying all prior art, all documents and all facts that you believe support your contention.

FIRST SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 7:

Yahoo maintains and fully incorporates herein each of the general objections and specific objections to this interrogatory listed in Yahoo's May 27, 2010 Objections and Responses to Xerox's First Set of Interrogatories. Subject to and without waiving these objections, Yahoo responds that the asserted claims of the Patents in Suit are invalid for at least the following reasons.

'979 Patent

Invalidity Under 35 U.S.C. § 101

Claims 1 and 18 of the '979 Patent are invalid under 35 U.S.C. § 101 because they claim unpatentable abstract ideas. Moreover, both claims fail the "machine-or-transformation" test indicative of § 101 patent eligible subject matter. Under the machine-or-transformation test, a claimed method is not patentable unless it (1) is tied to a particular machine or apparatus, or (2) transforms a particular article into a different state or thing. *See Bilski v. Kappos*, 561 U.S. __, slip op. at 3 (2010). The method of Claim 1, which analyzes document content to generate abstract queries, recites only general purpose computing equipment and does not meet the statutory requirements for patentable subject matter. Claim 18 is also invalid under 35 U.S.C. § 101 because the mere recitation of general purpose computer and software components does not

transform unpatentable method steps into patent-eligible subject matter and does not constitute recitation of a "particular machine."

Invalidity Under 35 U.S.C. §§ 102 and 103

Claims 1 and 18 of the '979 Patent are invalid under 35 U.S.C. §§ 102 and/or 103 in view of the prior art, including that identified below.

The following patents and patent applications are prior art under at least 35 U.S.C § 102(e): U.S. Patent Application 2002/0147738; U.S. Patent 6,122,647; U.S. Patent 6,473,752; U.S. Patent 6,606,644; U.S. Patent 6,829,780; U.S. Patent 7,076,443; U.S. Patent 7,225,142; U.S. Patent 7,418,657; and U.S. Patent 7,451,099 (collectively the "'979 Prior Art Patents").

The following systems were in public use prior to the invention date of the '979 Patent and are prior art under 35 U.S.C. §§ 102(a) and/or 102(b), and are 102(g)(2) prior inventions: eZula, YellowBrix IntelliClix, WebACE, IntelliZap / Zapper, and SemioMap Discovery Search (collectively the "'979 Prior Art Systems").

The following publications are prior art under 35 U.S.C. §§ 102(a) and/or 102(b): Finkelstein et al., *Placing Search in Context: The Concept Revisited*, Proc. of the 10th International World Wide Web Conference (May 1-5, 2001); Han et al., *WebACE: A Web Agent for Document Categorization and Exploration*, Proc. of the 2nd International Conference on Autonomous Agents (May 1998) (the "'979 Prior Art Publications"); and Wiesner et al., Context Matching System and Method, WO/2001/044992 (June 21, 2001).[2]

The '979 Prior Art Patents, the '979 Prior Art Systems and the '979 Prior Art Publications are collectively referred to as the "'979 Prior Art." The status of certain pieces of the '979 Prior Art may be affected by the Court's claim construction. In addition, some items of

[2] The correct citation for this reference is Wieser et al., Context Matching System and Method, WO/2001/044992 (June 21, 2001). Wieser is hereby included in the term "'979 Prior Art Patents."

art are presently believed to disclose certain elements of the asserted claims inherently. To the extent it is found that such elements are not inherently disclosed, it may be that the relevant claims are alternatively rendered obvious by the asserted reference and the knowledge of a person of ordinary skill in the art alone, or by various other art in combination with the asserted reference. A person of ordinary skill in the art would have found it obvious to combine each '979 Prior Art reference with one or more of the other '979 Prior Art references at least because each such reference relates to analyzing document content, and primarily the content of webpages, to find related information and/or relevant advertisements.

Yahoo incorporates by reference herein the identification by other defendants of any Prior Art as invalidating claims 1 and/or 18 of the '979 Patent under sections 102 and/or 103, to the extent such Prior Art is not specifically identified above. Yahoo reserves the right to use any of the identified references in support of an argument based on a disclosed system in prior use.

Invalidity Under 35 U.S.C. § 102(f) and 116

Pending further investigation, claims 1 and/or 18 '979 Patent may be invalid under 35 U.S.C. §§ 102(f) and 116 for failing to include all inventors of the claimed subject matter.

'994 Patent

Invalidity Under 35 U.S.C. § 101

Claim 9 of the '994 Patent is invalid under 35 U.S.C. § 101 because it claims an unpatentable abstract idea. Moreover, it fails the "machine-or-transformation" test indicative of § 101 patent eligible subject matter. The method of Claim 9, which relates to generic approaches to integrating abstract data and results of analyses thereof with abstract electronic documents, recites only general purpose computing and database equipment and does not meet the statutory requirements for patentable subject matter.

Invalidity Under 35 U.S.C. §§ 102 and 103

Claim 9 of the '994 Patent is invalid under 35 U.S.C. §§ 102 and/or 103 in view of the prior art, including that identified below.

The following patents are prior art under at least 35 U.S.C § 102(e): U.S. Patent 5,564,044; U.S. Patent 5,630,126; U.S. Patent 5,694,192; U.S. Patent 5,659,676; U.S. Patent 5,913,032; U.S. Patent 6,094,684 (collectively the "'994 Prior Art Patents").

The following systems were in public use prior to the invention date of the '994 Patent and are prior art under 35 U.S.C. §§ 102(a) and/or 102(b), and are 102(g)(2) prior inventions: Amazon.com Product Listings, the Internet Movie Database (IMDB), and Crystal Reports (collectively the "'994 Prior Art Systems").

The following publications are prior art under 35 U.S.C. §§ 102(a) and/or 102(b): Silvano Pozzi, et al., *ALIVE: A Distributed Live-link Documentation System*, Electronic Publishing, Vol. 5(3) (Sept. 1992), 131-142 and Premysl Brada, et al., *Dynamic Information Access Using WWW*, Proceedings of the 4th Conference on Interdisciplinary Information Management (1996), 97-102 (the "'994 Prior Art Publications").

The '994 Prior Art Patents, the '994 Prior Art Systems and the '994 Prior Art Publications are collectively referred to as the "'994 Prior Art." The status of certain pieces of the '994 Prior Art may be affected by the Court's claim construction. In addition, some items of art are presently believed to disclose certain elements of the asserted claims inherently. To the extent it is found that such elements are not inherently disclosed, it may be that the relevant claims are alternatively rendered obvious by the asserted reference and the knowledge of a person of ordinary skill in the art alone, or by various other art in combination with the asserted reference. A person of ordinary skill in the art would have found it obvious to combine each

‘994 Prior Art reference with one or more of the other ‘994 Prior Art references at least because each such reference relates to managing relationships between documents and data and analysis results.

Yahoo incorporates by reference herein the identification by other defendants of Prior Art as invalidating claim 9 of the ‘994 Patents under sections 102 and/or 103, to the extent such art is not specifically identified above. Yahoo reserves the right to use any of the identified references in support of an argument based on a disclosed system in prior use.

Invalidity Under 35 U.S.C. § 112

Claim 9 of the ‘994 Patent is invalid under at least plaintiff’s apparent construction (to the extent discernable, if at all, from its response to plaintiff’s interrogatory responses) for at least the following reasons: The claim limitations “storing knowledge,” “validating the accuracy of the knowledge”, “managing the flow of information between the first database and the document database to enable the integration of the data and analysis results with the documents and to automatically update the documents upon the occurrence of a change in the data or analysis results” and “to generate data and analysis results” do not meet the written description and enablement requirements of 35 U.S.C. § 112. To the extent that the term “knowledge” and the phrases “validating the accuracy of the knowledge”, “managing the flow of information”, “data and analysis results” and “data or analysis results” are insolubly ambiguous, claim 9 is indefinite.

Yahoo reserves the right to supplement, revise or render more specific its response to Interrogatory No. 7.

SECOND SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 7:

Yahoo! maintains and fully incorporates herein each of the general objections, specific objections and responses to this interrogatory listed in Yahoo!’s May 27, 2010 Objections and

Responses to Xerox's First Set of Interrogatories and Yahoo!'s July 9, 2010 Supplemental Response to Interrogatory No. 7. Subject to and without waiving these objections, Yahoo! hereby supplements its Response to Interrogatory No. 7. This Supplemental Response addresses only claims 1 and 18 of the '979 Patent. Just last week, more than eight months after serving its purported infringement contentions, Xerox alleged for the first time that certain Yahoo! products infringe claims 2, 3, 5, 10, and 19 of the '979 Patent. Yahoo! has objected to the assertion of these additional claims. Yahoo! will supplement this Response in due course to address these additional claims if and to the extent that the Court allows Xerox to expand its infringement case to encompass them. Xerox has indicated that it is dropping all claims based on the '994 Patent, and, accordingly, this Supplemental Response does not address the '994 Patent.

The following patents and patent applications are prior art under at least 35 U.S.C § 102(e): U.S. Patent 6,363,378; U.S. Patent 6,947,920; U.S. Patent 7,047,242; U.S. Patent 7,089,236; U.S. Patent 5,488,725; U.S. Patent 5,748,954; U.S. Patent 5,963,940; U.S. Patent 6,038,561; U.S. Patent 6,161,084; U.S. Patent 6,519,586; U.S. Patent 5,321,833; U.S. Patent Application 2003/0014405; and U.S. Patent Application 2002/052898. These patents and patent applications are hereby expressly and specifically added to the collection of "'979 Prior Art Patents" as that term is defined and used in Yahoo!'s First Supplemental Response to Interrogatory No. 7.

The following publications are prior art under 35 U.S.C. §§ 102(a) and/or 102(b): Mase et al., *Experimental Simulation for Automatic Patent Categorization*, Advances in Production management Systems, Kyoto, Japan, Nov. 1996, pages 377-382; Weiss et al., *HyPursuit: A Heirarchical Network Search Engine that Exploits Content-Link Hypertext Clustering*, Proc. of the 7th ACM Conference on Hypertext (March 1996), pages 180-193; Pazzani, et al., *Syskill &*

Webert: Identifying interesting web sites, AAAI-96 Proceedings (1996); and Salton, *Another Look at Automatic Text-Retrieval Systems*, Comm. of ACM (1986). These publications are hereby expressly and specifically added to the collection of "'979 Prior Art Publications" as that term is defined and used in Yahoo!'s First Supplemental Response to Interrogatory No. 7.

Yahoo! incorporates herein the discussion of prior art references and the invalidity arguments under 35 U.S.C. sections 102 and/or 103 set forth in Google's Corrected Request for Inter Partes Reexamination of the '979 Patent filed with the United States Patent and Trademark Office on or about September 8, 2010. Yahoo! further directs Xerox to all subsequent proceedings in connection with the reexamination of the '979 Patent. Yahoo! reserves the right to supplement, revise or render more specific its response to Interrogatory No. 7.

Attached hereto as Exhibits A though M are exemplary invalidity charts describing where individual elements of claims 1 and 18 of the '979 Patent may be found in certain prior art references that Yahoo! contends are invalidating under 35 U.S.C. § 102 and/or (alone or in combination) under 35 U.S.C. § 103. By identifying in the charts where in a given prior art reference a limitation of an asserted claim is found, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case. The charts include representative examples of where each element may be found in a given prior art reference and are not intended to be an exhaustive description of each instance of an element in a give prior art reference. These charts are not, and should not be construed to be, an admission by Yahoo! that the claims asserted by Xerox should be construed in the manner asserted by Xerox, or in any particular

manner so as to bring disclosed prior art within the scope of any particular claim limitation. Given that discovery is ongoing and that the Court has not yet issued an order construing the claims, Yahoo! reserves its rights to supplement or revise these charts.

THIRD SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 7:

Yahoo! maintains and fully incorporates herein each of the general objections, specific objections and responses to this interrogatory listed in Yahoo!'s May 27, 2010 Objections and Responses to Xerox's First Set of Interrogatories, Yahoo!'s July 9, 2010 Supplemental Response to Interrogatory No. 7, and Yahoo!'s February 17, 2011 Second Supplemental Response to Interrogatory No. 7. Subject to and without waiving these objections, Yahoo! hereby supplements its Response to Interrogatory No. 7.

Invalidity Under 35 U.S.C. §§ 102 and 103

Yahoo! attaches hereto Supplemental Appendices A through P, which are exemplary invalidity charts describing where individual elements of claims 1, 2, 3, 5, 10, 18, and 19 of the '979 Patent may be found in certain prior art references that Yahoo! contends are invalidating under 35 U.S.C. § 102 and/or (alone or in combination) under 35 U.S.C. §103. Yahoo! also incorporates herein the contentions and invalidity charts set forth in Defendant Google Inc.'s Third Supplemental Objections and Responses to Xerox's First Set of Interrogatories to Defendants (No. 7). By identifying in the charts where in a given prior art reference a limitation of an asserted claim is found, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case. The charts include representative examples of

where each element may be found in a given prior art reference and are not intended to be an exhaustive description of each instance of an element in a given prior art reference. These charts are not, and should not be construed to be, an admission by Yahoo! that the claims asserted by Xerox should be construed in the manner asserted by Xerox, or in any particular manner so as to bring disclosed prior art within the scope of any particular claim limitation. Given that discovery is ongoing and that the Court has not yet issued an order construing the claims, Yahoo! reserves its rights to supplement or revise these charts.

Yahoo! further identifies the following system in public use prior to the invention date of the '979 Patent, which is prior art under 35 U.S.C. §§ 102(a) and/or 102(b), and is a § 102(g)(2) prior invention: Oracle Text (also known as interMedia Text). This system is hereby expressly and specifically added to the collection of "'979 Prior Art Systems" as that term is defined and used in Yahoo!'s First Supplemental Response to Interrogatory No. 7. In addition, the following publications are prior art under 35 U.S.C. §§ 102(a) and/or 102(b):

Oracle Corp., *Oracle Text, An Oracle Technical White Paper* (May 2001) (available at http://ugweb.cs.ualberta.ca/~c391/W08/resources/oracle_text.pdf);

Oracle Corp., *Oracle8i interMedia Text Reference, Release 2* (8.1.6) (December 1999) (available at http://download.oracle.com/docs/cd/A87860_01/doc/inter.817/a77063.pdf);

Oracle Corp., *Oracle8i interMedia Text Migration, Release 2* (8.1.6) (December 1999) (available at http://download.oracle.com/docs/cd/A87860_01/doc/inter.817/a77061.pdf);

Oracle Corp., *Oracle8i interMedia Text 8.1.5 - Technical Overview* (1999) (available at http://www.oracle.com/technetwork/database/enterprise-edition/imt-815-083189.html).

These publications are hereby expressly and specifically added to the collection of "'979 Prior Art Publications" as that term is defined and used in Yahoo!'s First Supplemental Response to Interrogatory No. 7.

Yahoo! will rely on admissions of the inventors and Xerox as well as statements made in the '979 Patent concerning the scope of the prior art. Such admissions include: (1) the existence of known hierarchical ontologies such as DMOZ, Yahoo! and Northern Lights that could be used to classify document content (*see, e.g.,* '979 Patent at 41:5-9, 41:59-60; [redacted], (2) the existence of known techniques and products for classifying document content, including using a probabilistic model, fuzzy model, latent semantic indexing, and a vector space model (*see, e.g.,* '979 Patent at 41:52-67; [redacted] (3) the existence of known techniques to recognize and extract entities (*see, e.g.,* '979 Patent at 10:52-55; [redacted] (4) the existence of known techniques and products to automatically generate contextualized queries [redacted] and (5) the existence of information retrieval systems to which contextualized queries could be submitted as well as the ability to restrict a search to a certain category of information (*see, e.g.,* '979 Patent at 49:49-54; [redacted]

Yahoo! further contends that the "[c]ombination of known technologies" in the '979 Patent was obvious. [redacted] The motivation to combine the teachings of the prior art references is found in the references themselves and/or: (1) the nature of the problem being

solved, (2) the express, implied and inherent teachings of the prior art, (3) the knowledge of persons of ordinary skill in the art, (4) the fact that the prior art is generally directed towards providing personalized information services to a user, and/or (5) the predictable results obtained in combining the different elements of the prior art.

Prior to any alleged priority date relevant to the '979 Patent, numerous techniques for preventing the provision of too much information in response to a search request, which was the problem that the '979 Patent was purportedly designed to solve, were well known in the art. [redacted] These included simply adding search criteria such as keywords to formulate a more precise query. [redacted] Similarly, automating known search-narrowing techniques was well known in the art and the '979 Patent itself discloses numerous prior art systems that used document content to generate queries. (*See, e.g.,* '979 Patent at 2:10-33).

In addition, [redacted] Finally, the prior art references identified by Yahoo! also provide evidence of simultaneous invention, which further indicates that the '979 Patent was obvious.

Invalidity Under 35 U.S.C. §§ 102(f) and 116

[redacted]

Invalidity Under 35 U.S.C. § 112

Yahoo! contends that claims 1 and 18 of the '979 Patent as well as their dependent claims are invalid under 35 U.S.C. § 112. First, the claims are invalid because the phrase "selected document content" in claims 1 and 18 is indefinite. The phrase lacks any antecedent basis and is insolubly ambiguous because the specification fails to indicate how any document content is selected, by whom such a selection is made, or any other information regarding the selection of document content. Second, the claims are invalid for failure to disclose the best mode contemplated by the inventors of carrying out their invention. [redacted]

Given that discovery is ongoing and that the Court has not yet issued an order construing the claims, Yahoo! reserves its rights to supplement or revise its response concerning the bases of invalidity under 35 U.S.C. § 112 in response to positions taken by Xerox or the construction of claim terms by the Court. For example, to the extent that Xerox asserts that the "defining an organized classification of document content" element of claims 1 and 18 of the '979 Patent includes generating a classification scheme, Yahoo! contends that the element is invalid as insufficiently described and not enabled.

FOURTH SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 7:

Yahoo! maintains and fully incorporates herein each of the general objections, specific objections and responses to this interrogatory listed in Yahoo!'s May 27, 2010 Objections and Responses to Xerox's First Set of Interrogatories, Yahoo!'s July 9, 2010 Supplemental Response to Interrogatory No. 7, Yahoo!'s February 17, 2011 Second Supplemental Response to Interrogatory No. 7, and Yahoo!'s April 15, 2011 Third Supplemental Response to Interrogatory No. 7. Subject to and without waiving these objections, Yahoo! hereby supplements its Response to Interrogatory No. 7.

Yahoo! identifies the following publication as prior art under 35 U.S.C. §§ 102(a) and/or 102(b): Alexander Pretschner and Susan Gauch, *Ontology Based Personalized Search*, Proc. of the 11th IEEE Int'l Conf. on Tools with Artificial Intelligence, pp. 391-98 (Nov. 1999) ("Pretschner"). This publication is hereby expressly and specifically added to the collection of "'979 Prior Art Publications" as that term is defined and used in Yahoo!'s First Supplemental Response to Interrogatory No. 7.

Yahoo! attaches hereto Supplemental Appendices A through Q, which are exemplary invalidity charts describing where individual elements of claims 1, 2, 3, 5, 10, 18, and 19 of the '979 Patent may be found in certain prior art references that Yahoo! contends are invalidating under 35 U.S.C. § 102 and/or (alone or in combination) under 35 U.S.C. §103. Yahoo! also

incorporates herein the contentions and invalidity charts set forth in Defendant Google Inc.'s Fourth Supplemental Objections and Responses to Xerox's First Set of Interrogatories to Defendants (No. 7). By identifying in the charts where in a given prior art reference a limitation of an asserted claim is found, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case. The charts include representative examples of where each element may be found in a given prior art reference and are not intended to be an exhaustive description of each instance of an element in a given prior art reference. These charts are not, and should not be construed to be, an admission by Yahoo! that the claims asserted by Xerox should be construed in the manner asserted by Xerox, or in any particular manner so as to bring disclosed prior art within the scope of any particular claim limitation. Given that discovery is ongoing and that the Court has not yet issued an order construing the claims, Yahoo! reserves its rights to supplement or revise these charts.

May 11, 2011

By: 

Matthew B. Lehr (#2370)
Anthony I. Fenwick (admitted *pro hac vice*)
David J. Lisson (admitted *pro hac vice*)
Jeremy Brodsky (admitted *pro hac vice*)

DAVIS POLK & WARDWELL LLP
1600 El Camino Real
Menlo Park, CA 94025
650-752-2000
650-752-2111 (fax)

Attorneys For Defendants Yahoo! Inc. and Right Media LLC

Appendix A

By identifying in this chart where a limitation of an asserted claim is found in a reference, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case.

	Claims of the '979 Patent	**U.S. Patent No. 6,122,647**
1	A method for automatically generating a query from selected document content, comprising:	"The present invention overcomes the limitations of conventional information retrieval systems generally, and conventional Internet and intranet search engines particularly, by providing a system and method that dynamically generates contextual hypertext links in a source document to other topically relevant documents in response to the content of the source document or user-selected portion thereof. These new links are contextual links because they are generated in specific response to the content of a selected portion of the source document." 2:42-51. "[T]he present invention can treat any document (or portion of a document) a user is viewing as an inquiry and create new links in the source document to other related documents." 3:20-23. See Fig. 3. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 1, 2, 4, 15; Fig. 4 U.S. Patent No. 6,546,386 at Abstract, 1:47-54, 1:64-67, 2:10-14, 2:26-29, 2:42-45, 2:57-61, 3:29-35, 3:37-45, 3:55-57, 4:18-20; Fig. 1 PCT Application Pub. No. WO 01/44992 at 2:16-19, 5:8-12 U.S. Patent No. 6,236,768 at 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47

		Finkelstein at p. 406 U.S. Patent No. 6,473,752 at Abstract; Fig. 2 U.S. Patent No. 6,606,644 at 5:3-8, 5:30-39 U.S. Patent No. 6,829,780 at Abstract U.S. Patent No. 7,225,142 at 6:60-67 U.S. Patent No. 7,451,099 at 2:42-49, 44:8-19 Pretschner at p. 1.
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	"A knowledge base or other data repository stores information associating individual topics with sets of documents related to the topic, and with terms descriptive of the topic." 3:37-39. "The knowledge base 130 is a persistent data store that system 100 uses to store topic information. The knowledge base 130 comprises an arbitrary number of topics. Each topic is associated with one or more terms that are synonyms for each other. A term is a word or series of words (e.g., a noun phrase) that refer to a topic. A topic describes a possible subject annotation for documents in the document collection 140." 5:49-56. "Each topic in the knowledge base 130 may have a unique topic ID code for cross-referencing in other tables." 5:60-62. "[E]ach document (and document reference) has an association with at least one topic in the knowledge base 130, and preferably with many topics." 6:28-30. "Also, as a further embodiment, hierarchical topic menus may be created. Here, each menu item may be a topic in the knowledge base 130, with a submenu of related topics, and each related topic may have its own submenu of links to target documents. This hierarchical approach provides the user the ability to explore the entire document collection 140." 11:8-15. "The knowledge base 130 is a persistent data store that system 100 uses to store topic information. The knowledge base 130 comprises an arbitrary number of topics. Each topic is associated with one or

		more terms that are synonyms for each other. A term is a word or series of words (e.g., a noun phrase) that refer to a topic. A topic describes a possible subject annotation for documents in the document collection 140." 5:49-56. See Fig. 5. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 14, 15 Mase at Abstract; Fig. 1; p. 377-378 PCT Application Pub. No. WO 01/44992 at 15:8-11, 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 2:15-24, 4:20-27, 4:45-55 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19, 19:51-56; Fig. 8c Weiss at Fig. 3; pgs. 181, 184 Finkelstein at p. 410 U.S. Patent No. 6,473,752 at 1:44-53, 2:12-26, 6:24-44, 6:45-60 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 5:55-6:10, 11:16-40 U.S. Patent No. 7,225,142 at 9:33-45 U.S. Patent No. 7,451,099 at 9:20-25, 13:25-55, 14:4-9, 27:14-29, 29:25-29; Figs 3, 8, 23, 25C Oracle Text White Paper, p. 11, 18, 19

		Pretschner at p. 1 and 2.
	automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"The tagging module 120 is responsible for analyzing a selected portion of a document, identifying a set of terms and topics that are relevant to the selected portion or about the selected portion, and generating tags in the document which associate the terms and topics. The tagging module 120 may apply a variety of linguistic analysis techniques to identify the relevant topics." 6:49-55. "From the selected portion 304 of the source document 300, a number of new contextual links to target documents 310 relevant to the selected portion 304 are generated 306. These new contextual links are associated with selected terms of the source document, typically, but not necessarily only, in the selected portion 304." 7:19-25. "The terms selected for the links may be any terms of the selected portion, or may be other terms not necessarily appearing in the selected portion 304, but associated with topics that are most relevant to the selected portion 304." 7:46-49. "Referring now to FIG. 7, there is shown a flowgraph of one embodiment of a process for selecting topics and generating tags to such topics, as preferably implemented by the tagging module 120. The input to the process is a selected portion of a source document." 8:40-43. "Referring to the figure, the tagging module 120 tokenizes 702 the selected portion, dividing the selected portion into words and/or word phrases, each of which constitutes a token." 8:50-53. "The tagging module 120 then parses 706 the tokens to recognize groups of grammatical noun phrases." 8:62-63. "The tagging module 120 then selects 712 terms from the set of unified terms, based on threshold parameters for the number of terms to be selected." 9:13-15. "For each selected term, the tagging module 120 looks up 716 the term in the knowledge base 130." 9:27-28. "The contextual links provide the user with access to target documents that are relevant to the selected portion. The contextual links are generated from a linguistic analysis of the selected portion which identifies particular terms or keywords that are relevant to or about the selected portion." 12:27-32. See Figs. 4c-4d, 5, 7. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.:

		U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 15 PCT Application Pub. No. WO 01/44992 at 14:19 to 15:5 U.S. Patent No. 6,236,768 at 1:56-2:6, 2:15-54, 13:19-34 U.S. Patent No. 6,546,386 at 2:26-29, 2:42-45, 3:29-35 Finkelstein at pgs. 410, 408 U.S. Patent No. 6,473,752 at 3:5-8, 3:14-23, 5:62-6:7, 9:26-29; Figs. 2, 5, 8, 10B U.S. Patent No. 6,606,644 at 11:4-11, 12:52-63 U.S. Patent No. 6,829,780 at 5:55-6:14, 9:34-36 U.S. Patent No. 7,225,142 at 9:24-32 U.S. Patent No. 7,451,099 at 4:56-64, 7:22-25, 27:14-29, 27:65-28:5; Figs. 16A, 16B Oracle Text White Paper, p. 8, 9 Pretschner at p. 2.
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	"In one embodiment, the present invention provides a computer-implemented process in which a portion of text of a source document is analyzed and a number of topics are determined as being representative of what the selected portion is about. Topic analysis may be determined by various syntactic and semantic processes, such as identification and frequency analysis of terms of the selected portion. For each of the topics, a new tag is added to the source document. A tag includes a term, preferably from the text of the document, and a reference to the topic associated with the term. These tags are preferably stored with the source document." 3:25-35. "The user selects a portion 304 of the source document 300, which may be the entire source document 300, or any lesser portion of it, such as a selected set of words, a sentence, paragraph, or the like. The

		selected portion 304 is provided to the tagging module 120, which is coupled to the knowledge base 130." 7:59-65. "The tagging module 120 determines the topics in the knowledge base 130 that are about the selected portion 304. Preferably the tagging module 120 applies some type of linguistic analysis to the selected portion, including either syntactic or semantic analysis methods to determine the topics that are most representative or relevant to the selected portion 304." 7:66-8:5. See Figs. 6-8. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 Mase at p. 377-379; Fig. 1 PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 5:12-28 U.S. Patent No. 6,546,386 at 2:10-18, 2:57-61 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-18, 20:19-29 Weiss at pgs. 181, 182, 185, 186, 191 Finkelstein at pgs. 408, 410 U.S. Patent No. 6,473,752 at 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 11:16-40 U.S. Patent No. 7,225,142 at 9:33-37

		U.S. Patent No. 7,451,099 at 27:14-34, 27:46-51; Fig 16A Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 2.
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	"The tagging module receives as input a user selected portion of a source document and determines the topics relevant to the user selected portion, and creates a set of tags, each tag associating a term of the selected portion to one or more topics in the knowledge base. The presentation module receives the document and its set of tags, determines from the knowledge base the topics associated with each tag, and the target documents associated with each topic. The presentation module adds links to these target documents to the source document, either as links in the document body where one or more of the keywords appears, or in a separate navigational component." 3:64-4:8. "The presentation module 150 is responsible for determining a set of target documents associated with the generated tags in a document. . . ." 6:56-58. "Preferably the tagging module 120 applies some type of linguistic analysis to the selected portion, including either syntactic or semantic analysis methods to determine the topics that are most representative or relevant to the selected portion 304. The tagging module 120 adds a tag to the source document 300 for each of these topics, the tag specifying the topic in the knowledge base 130 . . . ." 8:1-7. "The presentation module 150 receives a source document including the set of tags in the document, and creates new links in the source document to other target documents. The presentation module 150 uses the knowledge base 130 to access the topics in knowledge base 130 associated with the tags. The presentation module 150 uses the document collection 140 to obtain references to the target documents associated with these topics." 8:17-24. "The presentation module 150 receives the source document 300 and the set of tags created by the tagging module 120. For each tag (802), the presentation module 150 looks up 804 the topic(s) in the knowledge base 130 specified in the tag. For each such topic the presentation module 150 retrieves 806 the documents associated with the topic, or more particularly, retrieves the URL (or other specification

		of network location) for the document." 10:11-18. See Figs. 5-6, 8. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15; Fig. 4 PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:14-22 U.S. Patent No. 6,236,768 at 5:12-28 U.S. Patent No. 6,546,386 at 3:37-45, 3:55-57 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-29 Weiss at pgs. 181, 182, 184, 185, 186 Finkelstein at p. 410 U.S. Patent No. 6,473,752 at 4:23-29, 7:51-8:4, 10:34-38, 15:19-23; Figs. 2, 7, 10B U.S. Patent No. 6,606,644 at 11:4-17, 12:2-7 U.S. Patent No. 6,829,780 at 13:40-50, 13:64-67, 14:43-52 U.S. Patent No. 7,225,142 at 4:20-22, 9:38-45 U.S. Patent No. 7,451,099 at 27:46-28:5, 28:30-44 Pretschner at p. 1 and 2.

2	The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	"From the selected portion 304 of the source document 300, a number of new contextual links to target documents 310 relevant to the selected portion 304 are generated 306. These new contextual links are associated with selected terms of the source document, typically, but not necessarily only, in the selected portion 304." 7:19-25. "The terms selected for the links may be any terms of the selected portion, or may be other terms not necessarily appearing in the selected portion 304, but associated with topics that are most relevant to the selected portion 304." 7:46-49. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.
3	The method according to claim 2, wherein the number of terms added is limited to a	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.:

	predefined number	U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.
5	The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	"The knowledge base 130 further contains a number of topics linked together in various hierarchical interconnected graphs by relations. A topic may have any number of relationships to other topics. Each relationship defines a semantic relationship between two topics, and has a predefined type. Relationship types include, for example, parent and child relationships (e.g., "is-a" or "type-of")." 5:65-6:4. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g. U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15

		U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19; Fig. 8c Weiss at pgs. 180, 184; Fig 3 U.S. Patent No. 6,236,768 at 12:41-49, 11:62-12:4, 13:1-8, 13:15-19 U.S. Patent No. 7,451,099 at 3:1-3, 14:4-9, 14:25-36, 27:30-34; Figs. 3, 23 Finkelstein at p. 410 Oracle Text White Paper, p. 19 Pretschner at p. 2.
10	The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary	"A knowledge base or other data repository stores information associating individual topics with sets of documents related to the topic, and with terms descriptive of the topic." 3:37-39. "Each topic is associated with one or more terms that are synonyms for each other. A term is a word or series of words (e.g., a noun phrase) that refer to a topic. A topic describes a possible subject annotation for documents in the document collection 140." 5:50-55. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 16:5-9 Weiss at p. 185 Mase at pgs. 377, 379

		U.S. Patent No. 6,236,768 at 4:45-55 U.S. Patent No. 6,473,752 at 6:45-60, 12:3-11; Fig. 9 Oracle Text White Paper, p. 19 Pretschner at p. 2.
18	An article of manufacture for use in a computer system, comprising:	"Referring now FIG. 2 there is shown an illustration of the software architecture of an information retrieval system 100 in accordance with the present invention. The information retrieval system 100 includes various functional software modules and structures that execute on a conventional computer system." 5:25-30. See Figs. 2, 9.
	a memory;	"The computer system includes a processor 105, addressable memory 103, operating system 107, display device 109, and user input device, such as a keyboard 111 or a mouse." 5:30-33. See Fig. 2 at 103.
	instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	See Chart for Claim 1 (above). "Referring now FIG. 2 there is shown an illustration of the software architecture of an information retrieval system 100 in accordance with the present invention. The information retrieval system 100 includes various functional software modules and structures that execute on a conventional computer system." 5:25-30. See Fig. 2 at 110, 120, 150.
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding	See Chart for Claim 1 (above).

	to a category of information in an information retrieval system;	
	automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	See Chart for Claim 1 (above).
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).
19	The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	See Chart for Claim 2 (above).

Appendix B

By identifying in this chart where a limitation of an asserted claim is found in a reference, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case.

	Claims of the '979 Patent	**U.S. Patent No. 6,473,752**
1	A method for automatically generating a query from selected document content, comprising:	"A system for locating computer documents or data of interest to a user without specification by the user of topics of interest. The system detects the selection of computer documents by the user of the system, and analyzes the contents of the selected computer documents to identify topics to which the contents are related. . . . The system then proceeds without user intervention, using the identified topics . . . to generate topics of interest to the user. . . . The system then uses the prioritized generated topics of user interest to locate documents whose contents are of interest to the user, and makes the located documents available to the user for selection. . . . For example, a computer document search engine can be used to locate additional documents by generating an appropriate search query." Abstract. See Fig. 2. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 1, 2, 4, 15; Fig. 4 U.S. Patent No. 6,546,386 at Abstract, 1:47-54, 1:64-67, 2:10-14, 2:26-29, 2:42-45, 2:57-61, 3:29-35, 3:37-45, 3:55-57, 4:18-20; Fig. 1 PCT Application Pub. No. WO 01/44992 at 2:16-19, 5:8-12

		U.S. Patent No. 6,236,768 at 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein at p. 406 U.S. Patent No. 6,122,647 at 2:42-51, 3:20-23 U.S. Patent No. 6,606,644 at 5:3-8, 5:30-39 U.S. Patent No. 6,829,780 at Abstract U.S. Patent No. 7,225,142 at 6:60-67 U.S. Patent No. 7,451,099 at 2:42-49, 44:8-19 Pretschner at p. 1.
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	"[V]arious companies have developed information search engines which can automatically index and organize information that is accessible from a computer. This accessible information may be located on any networked computer or storage device that the computer can access, or may be located on the computer system itself. After the information is indexed or organized, these search engines can then search the indexed or organized information to locate particular information of interest." 1:44-53. "The search engine will analyze the contents of the documents, and create an index of some or all of the terms in the documents. The search engine may also attempt to identify one or more general topics to which the entire document relates. The search engine will next search the documents for references to other computer documents. Upon finding such references, the search engine will access those referenced documents and continue the same process. In this manner, the search engines can eventually traverse and index all computer documents that are interconnected with the first documents given to the search engine. After creating this comprehensive index, the search engine can locate documents by receiving a search query containing terms or topics of interest to a user, and by searching the index to locate documents with corresponding terms or topics." 2:12-26. "After usage information is recorded and relevant terms are extracted, the Document Access Monitor 131 then forwards the recorded usage information and the extracted document terms to the Topic

		Analyzer 132. The Topic Generator 221 first receives the information from the Document Access Monitor 131, end generates topics related to areas of interest to the computer user. The Topic Generator 221 begins by generating related terms for the extracted terms. . . . Those skilled in the art will appreciate that related terms can be generated in a variety of ways, including the use of a thesaurus or the use of empirical testing to determine how terms are actually used." 6:24-44. "The related terms . . . assist in generating topics that are related to the contents of accessed documents. For example, if an accessed document contained only the term 'bat,' it would be difficult to determine which meaning of 'bat' was of interest to the user. However, if other extracted terms from this or other accessed documents included related terms such as 'Count Dracula' or 'a flying mammal' but not terms related to baseball, then it is likely that the user is interested in topics related to the mammal but not to the sport. Therefore, after the related terms are generated, the extracted terms are grouped together and used to determine topics of interest to the user. Those skilled in the art will appreciate that this grouping and determination can be performed in a variety of ways." 6:45-60. "Those skilled in the art will appreciate that other methods of generating topics related to document contents are possible. These may include methods that do not use extracted terms, other methods of grouping terms, topics that include a single term, or generating a new term that is broad enough to include multiple extracted and related terms." 12:18-24. "In one embodiment of the present invention, the search for new documents is conducted within a large set of accessible documents whose contents vary only slightly or not at all. In this situation, an index of the contents of these documents can be created, and the documents can later be retrieved based on the index." 15:35-40. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 14, 15 Mase at Abstract; Fig. 1; p. 377-378 PCT Application Pub. No. WO 01/44992 at 15:8-11, 17:15-32, 18:16-22

		U.S. Patent No. 6,236,768 at Table 2; 2:15-24, 4:20-27, 4:45-55 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19, 19:51-56; Fig. 8c Weiss at Fig. 3; pgs. 181, 184 Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56; Fig. 5 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 5:55-6:10, 11:16-40 U.S. Patent No. 7,225,142 at 9:33-45 U.S. Patent No. 7,451,099 at 9:20-25, 13:25-55, 14:4-9, 27:14-29, 29:25-29; Figs 3, 8, 23, 25C Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 1 and 2.
	automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"The system also analyzes the contents of the selected computer documents to identify relevant terms in the contents of the documents, and more generally to identify topics to which the contents are related." 3:5-8. "The system then attempts to locate additional computer documents, on any computer or device that is accessible to the system, whose contents are related to these prioritized generated topics of user interest. One method that the system may use to locate these documents involves identifying a computer document search engine, generating an appropriate search query, and requesting the search engine to perform the search on the generated search query." 3:14-23. "The Document Term Extractor 215 receives the detected document access notifications and extracts

		relevant terms from the contents of the documents that are accessed. In one embodiment, the detected document access notifications indicate only that a document is being accessed by the user, and the Document Term Extractor 215 accesses the document to ascertain its contents. In another embodiment, the detected document access notifications not only indicate that a document is being accessed, but also include the document contents that are being accessed. Typically, only relevant terms are extracted from the document contents, with relevance measured by the degree of relation between a term and the contents." 5:62- 6:7. "Those skilled in the art will appreciate that relevant terms can be generated in a variety of ways, and can be extracted from the contents of the entire document or only from the content of the portions of the document with which the user interacts." 9:25-29. See Figs. 2, 5, 8, 10B. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 15 PCT Application Pub. No. WO 01/44992 at 14:19 to 15:5 U.S. Patent No. 6,236,768 at 1:56-2:6, 2:15-54, 13:19-34 U.S. Patent No. 6,546,386 at 2:26-29, 2:42-45, 3:29-35 Finkelstein at pgs. 410, 408 U.S. Patent No. 6,122,647 at 6:49-55, 7:46-49, 9:27-28, 12:27-32; Figs. 4c-4d, 5, 7 U.S. Patent No. 6,606,644 at 11:4-11, 12:52-63 U.S. Patent No. 6,829,780 at 5:55-6:14, 9:34-36

		U.S. Patent No. 7,225,142 at 9:24-32 U.S. Patent No. 7,451,099 at 4:56-64, 7:22-25, 27:14-29, 27:65-28:5; Figs. 16A, 16B Oracle Text White Paper, p. 8, 9 Pretschner at p. 2.
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	"The present invention relates generally to locating computer documents and more particularly to determining topics of interest to a user and locating documents related to those topics." 1:6-9. "The system also analyzes the contents of the selected computer documents to identify relevant terms in the contents of the documents, and more generally to identify topics to which the contents are related." 3:5-8. "Those skilled in the art will appreciate that relevant terms can be generated in a variety of ways, and can be extracted from the contents of the entire document or only from the content of the portions of the document with which the user interacts." 9:25-29. "Those skilled in the art will appreciate that topics of user interest can be generated in a variety of ways, that the importance of a topic can be calculated in a variety of ways (i.e., using of a variety of importance measures), and that the topics can be prioritized in a variety of ways." 9:65-10:2. "Those skilled in the art will appreciate that other methods of generating topics related to document contents are possible. These may include methods that do not use extracted terms . . . or generating a new term that is broad enough to include multiple extracted and related terms." 12:18-24. See Figs. 2, 5, 6, 9. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 Mase at p. 377-379; Fig. 1

		PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 5:12-28 U.S. Patent No. 6,546,386 at 2:10-18, 2:57-61 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-18, 20:19-29 Weiss at pgs. 181, 182, 185, 186, 191 Finkelstein at pgs. 408, 410 U.S. Patent No. 6,122,647 at 3:25-35, 7:66-8:5; Figs. 6-8 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 11:16-40 U.S. Patent No. 7,225,142 at 9:33-37 U.S. Patent No. 7,451,099 at 27:14-34, 27:46-51; Fig 16A Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 2.
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	"The New Document Identifier receives the prioritized topics from the Topic Analyzer, generates a search query related to the prioritized topics, performs a search of accessible documents using the search query, and identifies documents that contain terms or topics related to the search query. These identified documents are then made available to the user for selection." 4:23-29. "Those skilled in the art will appreciate that a variety of search queries can be formulated to identify documents related to a given set of topics, and that these different search queries will often identify

		different groups of documents. For example, a search could be generated to identify only those documents whose contents contain every prioritized generated topic, or instead a search could be generated to identify those documents whose contents contain any prioritized generated topic. The generated search query is forwarded to the Search Engine 233, which uses the search query to perform a search on accessible computer documents. Those skilled in the art will appreciate that a variety of search engines are known in the art, including search engines from companies such as Infoseek, Excite, and Digital Equipment Corporation (DEC). Each search engine has its own rules and syntax for the search queries used by it. Thus, the particular search query that is generated will depend on the search engine to be used and on a determination of which available search query for that engine is most likely to return the desired documents." 7:51-8:4. "Those skilled in the art will appreciate that a search query can be generated in a variety of ways, including varying syntax to reflect a particular search engine and varying the topics in the search query in an attempt to identify different groups of documents." 10:34-38. "Generated search query 1050 illustrates that although one preferred embodiment prioritized topics based only on extracted terms, the generated search query includes both extracted terms and related terms." 15:19-23. See Figs. 2, 7, 10B. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15; Fig. 4. PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:14-22 U.S. Patent No. 6,236,768 at 5:12-28 U.S. Patent No. 6,546,386 at 3:37-45, 3:55-57

		U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-29 Weiss at pgs. 181, 182, 184, 185, 186 Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18; Figs. 5-6, 8 U.S. Patent No. 6,606,644 at 11:4-17, 12:2-7 U.S. Patent No. 6,829,780 at 13:40-50, 13:64-67, 14:43-52 U.S. Patent No. 7,225,142 at 4:20-22, 9:38-45 U.S. Patent No. 7,451,099 at 27:46-28:5, 28:30-44 Pretschner at p. 1 and 2.
2	The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	"The related terms not only provide additional terms that could be used to broaden a generated search query…" 6:45-46. Note that a term can itself be an extracted term and also be a related term for other extracted terms, such as the terms "White House" and "Hillary Clinton." 12:11-14. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45

		Weiss at p. 186 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.
3	The method according to claim 2, wherein the number of terms added is limited to a predefined number	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.

5	The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19; Fig. 8c Weiss at pgs. 180, 184; Fig 3 U.S. Patent No. 6,236,768 at 12:41-49, 11:62-12:4, 13:1-8, 13:15-19 U.S. Patent No. 6,122,647 at 5:65-6:4 U.S. Patent No. 7,451,099 at 3:1-3, 14:4-9, 14:25-36, 27:30-34; Figs. 3, 23 Finkelstein at p. 410 Oracle Text White Paper, p. 19 Pretschner at p. 2.
10	The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary	"The related terms . . . assist in generating topics that are related to the contents of accessed documents. For example, if an accessed document contained only the term 'bat,' it would be difficult to determine which meaning of 'bat' was of interest to the user. However, if other extracted terms from this or other accessed documents included related terms such as 'Count Dracula' or 'a flying mammal' but not terms related to baseball, then it is likely that the user is interested in topics related to the mammal but not to the sport. Therefore, after the related terms are generated, the extracted terms are grouped together and used to determine topics of interest to the user. Those skilled in the art will appreciate that this grouping and determination can be performed in a variety of ways." 6:45-60. "After the related terms are generated, the extracted and related terms are used to create the generated topics 950. Generated topics 950 comprises seven groups, with Group 1 including the extracted terms

		"President Clinton," "Hillary Clinton," "White House," and "U.S.," as well as related terms "Bill Clinton" and "Arkansas Governor." Group 2 includes the extracted terms of "educational system," "testing," "teachers," "standards," and "schools," as well as the related terms of "Secretary of Education," "universities," "tests," "exams," and "recess." Groups 3-7 contain similar sets of extracted and related terms." 12:3-11. See Fig. 9. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 16:5-9 Weiss at p. 185 Mase at pgs. 377, 379 U.S. Patent No. 6,236,768 at 4:45-55 U.S. Patent No. 6,122,647 at 3:37-39, 5:50-55 Oracle Text White Paper, p. 19 Pretschner at p. 2.
18	An article of manufacture for use in a computer system, comprising:	See Fig. 1 (depicting system implemented on computer system). "The computer system 110 includes the CPU 120, the memory 130, the input/output devices 140 and the bus 148. The input/output devices 140 include a storage device 141, a display 142, a keyboard 143 and a computer-readable media drive 144. The memory 130 includes the RDS system 135, which

		comprises the Document Access Monitor component 131, the Topic Analyzer component 132, and the New Document Identifier component 133." 4:34-41.
	a memory;	See Fig. 1 at 130. "The computer system 110 includes the CPU 120, the memory 130, the input/output devices 140 and the bus 148. The input/output devices 140 include a storage device 141, a display 142, a keyboard 143 and a computer-readable media drive 144. The memory 130 includes the RDS system 135, which comprises the Document Access Monitor component 131, the Topic Analyzer component 132, and the New Document Identifier component 133." 4:34-41.
	instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	See Chart for Claim 1 (above). Instructions for the system are inherently stored in the memory depicted in figure 1. "The computer system 110 includes the CPU 120, the memory 130, the input/output devices 140 and the bus 148. The input/output devices 140 include a storage device 141, a display 142, a keyboard 143 and a computer-readable media drive 144. The memory 130 includes the RDS system 135, which comprises the Document Access Monitor component 131, the Topic Analyzer component 132, and the New Document Identifier component 133." 4:34-41.
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	See Chart for Claim 1 (above).
	automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized	See Chart for Claim 1 (above).

	classification of content; and	
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).
19	The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content."	See Chart for Claim 2 (above).

Appendix C

By identifying in this chart where a limitation of an asserted claim is found in a reference, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case.

	Claims of the '979 Patent	**U.S. Patent No. 6,606,644**
1	A method for automatically generating a query from selected document content, comprising:	"The present invention provides a method and apparatus for automatically gathering, summarizing, and indexing real-time information derived from real-time communication on the Internet, such as Internet chat sessions, and to make that data readily available for immediate analysis and use such as targeted advertising." 5:3-8. "The present invention will give advertisers the ability to dynamically monitor the conversation being held in any given chat room, and be able to display advertising banners that match the theme of the conversation, thus, eliciting greater attention and interest from users. For example, a sudden occurrence of keywords such as 'car', 'automobile', 'drive', 'convertible', 'coupe', etc., may signal to an automobile manufacturer that now is an opportune time to display a banner advertising a special sale on convertible automobiles." 5:30-39. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 1, 2, 4, 15; Fig. 4 U.S. Patent No. 6,546,386 at Abstract, 1:47-54, 1:64-67, 2:10-14, 2:26-29, 2:42-45, 2:57-61, 3:29-35, 3:37-45, 3:55-57, 4:18-20; Fig. 1 PCT Application Pub. No. WO 01/44992 at 2:16-19, 5:8-12

		U.S. Patent No. 6,236,768 at 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein at p. 406 U.S. Patent No. 6,122,647 at 2:42-51, 3:20-23 U.S. Patent No. 6,473,752 at Abstract; Fig. 2 U.S. Patent No. 6,829,780 at Abstract U.S. Patent No. 7,225,142 at 6:60-67 U.S. Patent No. 7,451,099 at 2:42-49, 44:8-19. Pretschner at p. 1.
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	"As shown in FIG. 3, this aspect of the present invention [i.e., serving contextually-relevant advertisements] will extract and analyze the summary information generated by LISA 50 and then respond with an appropriate message or advertisement responsive to the subject matter of the summary information. Specifically, based upon the user's specifications, the present invention will look for the temporally contiguous occurrence of a particular keyword or set of keywords. The appearance of such keywords would be used to infer the general topic or subject matter of the communication or conversation. Once the general subject area of the conversation has been (probabilistically) deduced, the present invention may be used to present advertisements in the chat room that match the theme of the conversation." 11:4-17. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 14, 15 Mase at Abstract; Fig. 1; p. 377-378 PCT Application Pub. No. WO 01/44992 at 15:8-11, 17:15-32, 18:16-22

		U.S. Patent No. 6,236,768 at Table 2; 2:15-24, 4:20-27, 4:45-55 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19, 19:51-56; Fig. 8c Weiss at Fig. 3; pgs. 181, 184 Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56; Fig. 5 U.S. Patent No. 6,473,752 at 1:44-53, 2:12-26, 6:24-44, 6:45-60 U.S. Patent No. 6,829,780 at 5:55-6:10, 11:16-40 U.S. Patent No. 7,225,142 at 9:33-45 U.S. Patent No. 7,451,099 at 9:20-25, 13:25-55, 14:4-9, 27:14-29, 29:25-29; Figs 3, 8, 23, 25C Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 1 and 2.
	automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"As shown in FIG. 3, this aspect of the present invention [i.e., serving contextually-relevant advertisements] will extract and analyze the summary information generated by LISA 50 and then respond with an appropriate message or advertisement responsive to the subject matter of the summary information. Specifically, based upon the user's specifications, the present invention will look for the temporally contiguous occurrence of a particular keyword or set of keywords." 11:4-11. "In particular, using the information it has received from database manager 34, client interface 36 can present the advertisement in extremely customized, personalized, and interactive manner. For example, 'Bob' is a participant in a conversation in which the key words 'linux', 'windows', 'operating systems', 'software', and 'application', have been detected by LISA 50. Client interface 36 may now present Bob with a personal message that contains an advertisement for 'red hat linux' products, and

		which may state something like 'Hi Bob, I felt you might be interested in this !" 12:52-63. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 15 PCT Application Pub. No. WO 01/44992 at 14:19 to 15:5 U.S. Patent No. 6,236,768 at 1:56-2:6, 2:15-54, 13:19-34 U.S. Patent No. 6,546,386 at 2:26-29, 2:42-45, 3:29-35 Finkelstein at pgs. 410, 408 U.S. Patent No. 6,122,647 at 6:49-55, 7:46-49, 9:27-28, 12:27-32; Figs. 4c-4d, 5, 7 U.S. Patent No. 6,473,752 at 3:5-8, 3:14-23, 5:62-6:7, 9:26-29; Figs. 2, 5, 8, 10B U.S. Patent No. 6,829,780 at 5:55-6:14, 9:34-36 U.S. Patent No. 7,225,142 at 9:24-32 U.S. Patent No. 7,451,099 at 4:56-64, 7:22-25, 27:14-29, 27:65-28:5; Figs. 16A, 16B Oracle Text White Paper, p. 8, 9 Pretschner at p. 2.
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized	"As shown in FIG. 3, this aspect of the present invention [i.e., serving contextually-relevant advertisements] will extract and analyze the summary information generated by LISA 50 and then respond with an appropriate message or advertisement responsive to the subject matter of the summary information. Specifically, based upon the user's specifications, the present invention will look for the temporally contiguous occurrence of a particular keyword or set of keywords. The appearance of such

	classification of content; and	keywords would be used to infer the general topic or subject matter of the communication or conversation. Once the general subject area of the conversation has been (probabilistically) deduced, the present invention may be used to present advertisements in the chat room that match the theme of the conversation." 11:4-17. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 Mase at p. 377-379; Fig. 1 PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 5:12-28 U.S. Patent No. 6,546,386 at 2:10-18, 2:57-61 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-18, 20:19-29 Weiss pgs. 181, 182, 185, 186, 191 Finkelstein at pgs. 408, 410 U.S. Patent No. 6,122,647 at 3:25-35, 7:66-8:5; Figs. 6-8 U.S. Patent No. 6,473,752 at 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 U.S. Patent No. 6,829,780 at 11:16-40 U.S. Patent No. 7,225,142 at 9:33-37 U.S. Patent No. 7,451,099 at 27:14-34, 27:46-51; Fig 16A

		Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 2.
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	"As shown in FIG. 3, this aspect of the present invention [i.e., serving contextually-relevant advertisements] will extract and analyze the summary information generated by LISA 50 and then respond with an appropriate message or advertisement responsive to the subject matter of the summary information. Specifically, based upon the user's specifications, the present invention will look for the temporally contiguous occurrence of a particular keyword or set of keywords. The appearance of such keywords would be used to infer the general topic or subject matter of the communication or conversation. Once the general subject area of the conversation has been (probabilistically) deduced, the present invention may be used to present advertisements in the chat room that match the theme of the conversation." 11:4-17. "In a targeted advertising aspect of the present invention, database manager 34 is adapted to be able to determine which advertisement(s) from the database 60 of advertisements should be displayed in response to a particular message being exchanged in real-time between two 'chatters' or clients 70." 12:2-7. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15; Fig. 4. PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:14-22 U.S. Patent No. 6,236,768 at 5:12-28 U.S. Patent No. 6,546,386 at 3:37-45, 3:55-57 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-29 Weiss at pgs. 181, 182, 184, 185, 186

		Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18; Figs. 5-6, 8 U.S. Patent No. 6,473,752 at 4:23-29, 7:51-8:4, 10:34-38, 15:19-23; Figs. 2, 7, 10B U.S. Patent No. 6,829,780 at 13:40-50, 13:64-67, 14:43-52 U.S. Patent No. 7,225,142 at 4:20-22, 9:38-45 U.S. Patent No. 7,451,099 at 27:46-28:5, 28:30-44 Pretschner at p. 1 and 2.
2	The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c

		Finkelstein at pgs. 408, 410 Pretschner at p. 2.
3	The method according to claim 2, wherein the number of terms added is limited to a predefined number.	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.
5	The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2

		U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Oracle Text White Paper, p. 19 Pretschner at p. 2.
10	The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary.	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19; Fig. 8c Weiss at pgs. 180, 184; Fig 3

		U.S. Patent No. 6,236,768 at 12:41-49, 11:62-12:4, 13:1-8, 13:15-19 U.S. Patent No. 6,122,647 at 5:65-6:4 U.S. Patent No. 7,451,099 at 3:1-3, 14:4-9, 14:25-36, 27:30-34; Figs. 3, 23 Finkelstein at p. 410 Oracle Text White Paper, p. 19 Pretschner at p. 2.
18	An article of manufacture for use in a computer system, comprising:	"In another aspect, the present invention is directed to an article of manufacture comprising a computer usable medium having computer readable program code means embodied therein for causing a computer to automatically gather, summarize and index real-time communication on a computer network between at least a first and second user." 3:34-39. See also Fig. 3.
	a memory;	"The program storage devices of the present invention may be devised, made and used as a component of a machine utilizing optics, magnetic properties and/or electronics to perform the method steps of the present invention. Program storage devices include, but are not limited to, magnetic media such as a diskette or computer hard drive, which is readable and executable by a computer(s), optical disks, Read Only Memory (ROM), floppy disks, and semiconductor chips." 13:11-19.
	instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	See Chart for Claim 1 (above). "The program storage devices of the present invention may be devised, made and used as a component of a machine utilizing optics, magnetic properties and/or electronics to perform the method steps of the present invention. Program storage devices include, but are not limited to, magnetic media such as a diskette or computer hard drive, which is readable and executable by a computer(s), optical disks, Read Only Memory (ROM), floppy disks, and semiconductor chips." 13:11-19.
	defining an organized classification of	See Chart for Claim 1 (above).

	document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	
	automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	See Chart for Claim 1 (above).
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).
19	The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content."	See Chart for Claim 2 (above).

Appendix D

By identifying in this chart where a limitation of an asserted claim is found in a reference, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case.

	Claims of the '979 Patent	**U.S. Patent No. 6,829,780**
1	A method for automatically generating a query from selected document content, comprising:	"Transparently to the user, the system continuously operates in the background to adapt banner advertisements based on the detection of competing ads. The system includes a page analyzer that translates the hosted ad's web page into a document that can be analyzed for the presence of competing ads. An ad comparison unit compares the output of the page analyzer with information stored in the competitor ad database to detect competing ads. . . . The ad summary evaluator identifies competing ads and devises a counter strategy for banner ad display." Abstract. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 1, 2, 4, 15; Fig. 4 U.S. Patent No. 6,546,386 at Abstract, 1:47-54, 1:64-67, 2:10-14, 2:26-29, 2:42-45, 2:57-61, 3:29-35, 3:37-45, 3:55-57, 4:18-20; Fig. 1 PCT Application Pub. No. WO 01/44992 at 2:16-19, 5:8-12 U.S. Patent No. 6,236,768 at 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein at p. 406 U.S. Patent No. 6,122,647 at 2:42-51, 3:20-23

		U.S. Patent No. 6,473,752 at Abstract; Fig. 2 U.S. Patent No. 6,606,644 at 5:3-8, 5:30-39 U.S. Patent No. 7,225,142 at 6:60-67 U.S. Patent No. 7,451,099 at 2:42-49, 44:8-19 Pretschner at p. 1.
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	"In one embodiment, the system 10 operates with a banner display module 200 that displays the content of the adaptive banner ad, a web server 15, and an ad proxy router 205 that provides secure communication link between banner display module and the advertiser's server, and is generally comprised of: . . . a hosted ad database 220 that contains information about each banner ad hosted, including primary competitors; . . . a competitor ad database 235 that contains data about competing advertising;" 5:55-6:10. "The competitor ad database 235 contains data about competitors' advertisements that have [been] collected from the various advertisers and entered either manually or automatically. This data comprises a sample of the ad (e.g. an image) along with additional data useful for ad analysis. The system 10 performs database queries utilizing available image/multimedia comparison algorithms to locate a match. For the example of FIG. 4, the ad comparison unit 230 detects a match and the ad for the Mac Store is marked in the summary list as a <MATCHED_COMPETITOR>: <COMPETITOR id=4711> <MEDIA_URL>http://ad.doubleclicknet/ad/buy.prod.sm/homepage;cat=homepage_5;sz=100x60;tile=5;ord=16115127561</MEDIA_URL> <TARGET_URL>http://ad.doubleclick.net/jump/buy.prod.sm/homepage;cat=homepage_5;sz=100x60;ti1e=5;ord=16115127561</TARGET_URL>

		<title>BUYCOMP.COM - The Computer Superstore</title> <KEYWORDS>macintosh, mac, store</KEYWORDS> <MATCHED_COMPETITOR>Apple Inc.</MATCHED_COMPETITOR> <COMPETITOR>" 11:16-40. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 14, 15 Mase at Abstract; Fig. 1; p. 377-378 PCT Application Pub. No. WO 01/44992 at 15:8-11, 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 2:15-24, 4:20-27, 4:45-55 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19, 19:51-56; Fig. 8c Weiss at Fig. 3; pgs. 181, 184 Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56; Fig. 5 U.S. Patent No. 6,473,752 at 1:44-53, 2:12-26, 6:24-44, 6:45-60 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 7,225,142 at 9:33-45 U.S. Patent No. 7,451,099 at 9:20-25, 13:25-55, 14:4-9, 27:14-29, 29:25-29; Figs 3, 8, 23, 25C Oracle Text White Paper, p. 11, 18, 19

		Pretschner at p. 1 and 2.
	automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"In one embodiment, the system 10 operates with a banner display module 200 that displays the content of the adaptive banner ad, a web server 15, and an ad proxy router 205 that provides secure communication link between banner display module and the advertiser's server, and is generally comprised of: an ad identification manager 215 that performs competitive analysis of all the ads on the hosted ad's web page; . . . a page analyzer 225 that translates the hosted ad's web page into a document that can be analyzed for the presence of competing advertising; . . . an ad comparison unit 230 that compares the output of the page analyzer 225 with information stored in the competitor ad database 235 to detect competing advertising;" 5:55-6:14. "As illustrated in the foregoing example, the page analyzer 225 also extracts keywords and alternate representations from the web page and adds them to the page summary." 9:34-36. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 15 PCT Application Pub. No. WO 01/44992 at 14:19 to 15:5 U.S. Patent No. 6,236,768 at 1:56-2:6, 2:15-54, 13:19-34 U.S. Patent No. 6,546,386 at 2:26-29, 2:42-45, 3:29-35 Finkelstein at pgs. 410, 408 U.S. Patent No. 6,122,647 at 6:49-55, 7:46-49, 9:27-28, 12:27-32; Figs. 4c-4d, 5, 7

		U.S. Patent No. 6,473,752 at 3:5-8, 3:14-23, 5:62-6:7, 9:26-29; Figs. 2, 5, 8, 10B U.S. Patent No. 6,606,644 at 11:4-11, 12:52-63 U.S. Patent No. 7,225,142 at 9:24-32 U.S. Patent No. 7,451,099 at 4:56-64, 7:22-25, 27:14-29, 27:65-28:5; Figs. 16A, 16B Oracle Text White Paper, p. 8, 9 Pretschner at p. 2.
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	"The competitor ad database 235 contains data about competitors' advertisements that have [been] collected from the various advertisers and entered either manually or automatically. This data comprises a sample of the ad (e.g. an image) along with additional data useful for ad analysis. The system 10 performs database queries utilizing available image/multimedia comparison algorithms to locate a match. For the example of FIG. 4, the ad comparison unit 230 detects a match and the ad for the Mac Store is marked in the summary list as a <MATCHED_COMPETITOR>: <COMPETITOR id=4711> <MEDIA_URL>http://ad.doubleclicknet/ad/buy.prod.sm/homepage;cat=homepage _5;sz=100x60;tile=5;ord=16115127561</MEDIA_URL> <TARGET_URL>http://ad.doubleclick.net/jump/buy.prod.sm/homepage;cat=homep age_5;sz=100x60;ti1e=5;ord=16115127561</TARGET_URL> <title>BUYCOMP.COM - The Computer Superstore</title> <KEYWORDS>macintosh, mac, store</KEYWORDS> <MATCHED_COMPETITOR>Apple Inc.</MATCHED_COMPETITOR> <COMPETITOR>" 11:16-40. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.:

		U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 Mase at p. 377-379; Fig. 1 PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 5:12-28 U.S. Patent No. 6,546,386 at 2:10-18, 2:57-61 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-18, 20:19-29 Weiss at pgs. 181, 182, 185, 186, 191 Finkelstein at pgs. 408, 410 U.S. Patent No. 6,122,647 at 3:25-35, 7:66-8:5; Figs. 6-8 U.S. Patent No. 6,473,752 at 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 7,225,142 at 9:33-37 U.S. Patent No. 7,451,099 at 27:14-34, 27:46-51; Fig 16A Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 2.
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the	"The web page summary now contains all the information necessary . . . to define an appropriate ad strategy for the web page: Ad ID used to identify the hosted ad in the hosted ad database 220 and associated information; a list of all potential competitors on the web page including the media URL, target URL and

	information retrieval system identified by the assigned classification label.	additional metadata gained from document analysis or OCR shown highlighted in bold letters; and a list of competitors clearly identified by the ad comparison unit 230 as <MATCHED COMPETITOR>." 13:40-50. "Method 300 then determines from the hosted ad database 220, at decision step 350, if any of the ads on the web page are key competitors by having the page analyzer 225 pass the web page summary to the ad summary evaluator 255." 13:64-67. "The strategy or strategies used by the ad summary evaluator 255 are implemented as rules that are stored in the ad rules database 260. A rules engine is used to interpret these rules. For the example of FIG. 4, the ad summary evaluator 255 identifies the Mac Store as a key competitor. The ad summary evaluator 255 ignores the other ads because they offer services in different markets and thus are not competitors. From the ad rules database 260, the ad summary evaluator 255 obtains a rule which instructs the system 10 to use a specific Mac-targeted ad instead of the original ad." 14:43-52. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15; Fig. 4. PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:14-22 U.S. Patent No. 6,236,768 at 5:12-28 U.S. Patent No. 6,546,386 at 3:37-45, 3:55-57 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-29 Weiss at pgs. 181, 182, 184, 185, 186 Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18; Figs. 5-6, 8

		U.S. Patent No. 6,473,752 at 4:23-29, 7:51-8:4, 10:34-38, 15:19-23; Figs. 2, 7, 10B U.S. Patent No. 6,606,644 at 11:4-17, 12:2-7 U.S. Patent No. 7,225,142 at 4:20-22, 9:38-45 U.S. Patent No. 7,451,099 at 27:46-28:5, 28:30-44 Pretschner at p. 1 and 2.
2	The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	"At this point, the page analyzer 225 can also download the target URLs of each potential candidate and analyze these documents. Typically, these documents contain useful information (e.g. page title, abstracts etc.) that can be used to enrich the metadata of the summary document…" 10:32-36. The final XML representation created by the page analyzer 225 for the example of FIG. 4 is as follows: … <COMPETITOR id~4711> <MEDIA_URL>http://ad.doubleclicknet/adibuy.prod.sm/homepage;cat~homepage_ 5;sz~100x60;tile~5;ord~16115127561</MEDIA_URL> <TARGET_URL>http://ad.doubleclicknet/jump/buy.prod.sm/homepage;cat~homepage_ 5;sz~100x60;tile~5;ord~16115127561</TARGET_URL> <title>BUYCOMP.COM -The Computer Superstore</title> <KEYWORDS>macintosh, mac, store</KEYWORDS> <MATCHED_COMPETITOR>Apple [nc.</MATCHED_ COMPETITOR> <COMPETITOR> ... 13:2-40. "If, however, none of the identified candidates are competitors, other unidentified competitors may still reside on the web page. Consequently, a supplemental analysis might be required, and the page analyzer 225 invokes the OCR engine 240 at step 330 to convert the image data into text data. Performing OCR analysis on the web page retrieves additional information that can enhance the XML representation of potential candidates." 12:13-21.

		"The web page summary now contains all the information necessary . . . to define an appropriate ad strategy for the web page: Ad ID used to identify the hosted ad in the hosted ad database 220 and associated information; a list of all potential competitors on the web page including the media URL, target URL and additional metadata gained from document analysis or OCR shown highlighted in bold letters; and a list of competitors clearly identified by the ad comparison unit 230 as <MATCHED COMPETITOR>." 13:40-50. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.
3	The method according to claim 2, wherein the number of terms added is limited to a	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.:

	predefined number.	U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.
5	The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19; Fig. 8c Weiss at pgs. 180, 184; Fig 3 U.S. Patent No. 6,236,768 at 12:41-49, 11:62-12:4, 13:1-8, 13:15-19

		U.S. Patent No. 6,122,647 at 5:65-6:4 U.S. Patent No. 7,451,099 at 3:1-3, 14:4-9, 14:25-36, 27:30-34; Figs. 3, 23 Finkelstein at p. 410 Oracle Text White Paper, p. 19 Pretschner at p. 2.
10	The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary.	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 16:5-9 Weiss at p. 185 Mase at pgs. 377, 379 U.S. Patent No. 6,236,768 at 4:45-55 U.S. Patent No. 6,122,647 at 3:37-39, 5:50-55 U.S. Patent No. 6,473,752 at 6:45-60, 12:3-11; Fig. 9 Oracle Text White Paper, p. 19 Pretschner at p. 2.
18	An article of manufacture for use in a computer system, comprising:	"[T]his invention pertains to a computer software product for dynamically adapting, enhancing, and optimizing the appearance and content of a banner advertisement based on the automatic detection of competing advertising within a document." 1:23-27.

		See also Fig. 1.
	a memory;	"[T]his invention pertains to a computer software product for dynamically adapting, enhancing, and optimizing the appearance and content of a banner advertisement based on the automatic detection of competing advertising within a document." 1:23-27. A computer software product inherently requires the use of a memory.
	instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	See Chart for Claim 1 (above). "[T]his invention pertains to a computer software product for dynamically adapting, enhancing, and optimizing the appearance and content of a banner advertisement based on the automatic detection of competing advertising within a document." 1:23-27. A computer software product inherently requires the use of instructions stored in memory for it to be executable.
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	See Chart for Claim 1 (above).
	automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	See Chart for Claim 1 (above).
	automatically formulating the query to restrict a search at the information retrieval	See Chart for Claim 1 (above).

	system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	
19	The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content."	See Chart for Claim 2 (above).

Appendix E

By identifying in this chart where a limitation of an asserted claim is found in a reference, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case.

	Claims of the '979 Patent	**U.S. Patent No. 7,225,142**
1	A method for automatically generating a query from selected document content, comprising:	"Because the advertisements are streamed from a server rather than downloaded as a set and played to the user in a loop, the present invention can make choices about which advertisements to display to the user that are responsive to the user's current viewing habits. Thus, if a user is selecting and viewing pages in the browser area 31 concerning outdoor activities, the present invention can select advertisements for camping gear . . . ." 6:60-67. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 1, 2, 4, 15; Fig. 4 U.S. Patent No. 6,546,386 at Abstract, 1:47-54, 1:64-67, 2:10-14, 2:26-29, 2:42-45, 2:57-61, 3:29-35, 3:37-45, 3:55-57, 4:18-20; Fig. 1 PCT Application Pub. No. WO 01/44992 at 2:16-19, 5:8-12 U.S. Patent No. 6,236,768 at 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein at p. 406 U.S. Patent No. 6,122,647 at 2:42-51, 3:20-23

		U.S. Patent No. 6,473,752 at Abstract; Fig. 2 U.S. Patent No. 6,606,644 at 5:3-8, 5:30-39 U.S. Patent No. 6,829,780 at Abstract U.S. Patent No. 7,451,099 at 2:42-49, 44:8-19 Pretschner at p. 1.
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	"After extracting the keywords, the keywords are compared to a database index, which cross-references keywords with topic names. Thus, in the present example, the keyword 'surfing' matches topics 'outdoor adventure' and 'water sports.' 'Molokai' matches the topic 'Hawaii.' Each topic in the database is correlated with a series of URLs for advertisements that relate to the topic. Thus, the topic 'Hawaii' corresponds advertisements for the 'Airline Deals to Hawaii by TravelNow' and 'Luau Hawaiian Hotels,' which are now streamed to the user and displayed in the advertising area 37. In this way, the user's viewing habits are used to effectively target advertisements to the user that are pertinent to the user's interests." 9:33-45. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 14, 15 Mase at Abstract; Fig. 1; p. 377-378 PCT Application Pub. No. WO 01/44992 at 15:8-11, 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 2:15-24, 4:20-27, 4:45-55 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19, 19:51-56; Fig. 8c Weiss at Fig. 3; pgs. 181, 184

		Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56; Fig. 5 U.S. Patent No. 6,473,752 at 1:44-53, 2:12-26, 6:24-44, 6:45-60 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 5:55-6:10, 11:16-40 U.S. Patent No. 7,451,099 at 9:20-25, 13:25-55, 14:4-9, 27:14-29, 29:25-29; Figs 3, 8, 23, 25C Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 1 and 2.
	automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"In one embodiment, the present invention carries out this content-sensitive advertising by conducting a keyword search of a page requested to be displayed on the client computer by the user. Keywords are obtained by noting words that appear between TITLE headers in HTML documents. For example, a page that contains the code: <TITLE>Bill's Favorite Surfing Spots on Molokai</TITLE> the keywords 'surfing' and 'Molokai' would be extracted as keywords." 9:24-32. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 15 PCT Application Pub. No. WO 01/44992 at 14:19 to 15:5 U.S. Patent No. 6,236,768 at 1:56-2:6, 2:15-54, 13:19-34 U.S. Patent No. 6,546,386 at 2:26-29, 2:42-45, 3:29-35

		Finkelstein at pgs. 410, 408 U.S. Patent No. 6,122,647 at 6:49-55, 7:46-49, 9:27-28, 12:27-32; Figs. 4c-4d, 5, 7 U.S. Patent No. 6,473,752 at 3:5-8, 3:14-23, 5:62-6:7, 9:26-29; Figs. 2, 5, 8, 10B U.S. Patent No. 6,606,644 at 11:4-11, 12:52-63 U.S. Patent No. 6,829,780 at 5:55-6:14, 9:34-36 U.S. Patent No. 7,451,099 at 4:56-64, 7:22-25, 27:14-29, 27:65-28:5; Figs. 16A, 16B Oracle Text White Paper, p. 8, 9 Pretschner at p. 2.
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	"After extracting the keywords, the keywords are compared to a database index, which cross-references keywords with topic names. Thus, in the present example, the keyword 'surfing' matches topics 'outdoor adventure' and 'water sports.' 'Molokai' matches the topic 'Hawaii.'" 9:33-37. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 Mase at p. 377-379; Fig. 1 PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 5:12-28 U.S. Patent No. 6,546,386 at 2:10-18, 2:57-61

		U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-18, 20:19-29 Weiss at pgs. 181, 182, 185, 186, 191 Finkelstein at pgs. 408, 410 U.S. Patent No. 6,122,647 at 3:25-35, 7:66-8:5; Figs. 6-8 U.S. Patent No. 6,473,752 at 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 11:16-40 U.S. Patent No. 7,451,099 at 27:14-34, 27:46-51; Fig 16A Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 2.
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	"Advertisements may also be selected by deducing user areas of interest based upon the content of pages on the web selected by the user for viewing." 4:20-22. "Each topic in the database is correlated with a series of URLs for advertisements that relate to the topic. Thus, the topic 'Hawaii' corresponds advertisements for the 'Airline Deals to Hawaii by TravelNow' and 'Luau Hawaiian Hotels,' which are now streamed to the user and displayed in the advertising area 37. In this way, the user's viewing habits are used to effectively target advertisements to the user that are pertinent to the user's interests." 9:38-45. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15; Fig. 4.

		PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:14-22 U.S. Patent No. 6,236,768 at 5:12-28 U.S. Patent No. 6,546,386 at 3:37-45, 3:55-57 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-29 Weiss at pgs. 181, 182, 184, 185, 186 Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18; Figs. 5-6, 8 U.S. Patent No. 6,473,752 at 4:23-29, 7:51-8:4, 10:34-38, 15:19-23; Figs. 2, 7, 10B U.S. Patent No. 6,606,644 at 11:4-17, 12:2-7 U.S. Patent No. 6,829,780 at 13:40-50, 13:64-67, 14:43-52 U.S. Patent No. 7,451,099 at 27:46-28:5, 28:30-44 Pretschner at p. 1 and 2.
2	The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49

		U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.
3	The method according to claim 2, wherein the number of terms added is limited to a predefined number.	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5

		U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.
5	The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19; Fig. 8c Weiss at pgs. 180, 184; Fig 3 U.S. Patent No. 6,236,768 at 12:41-49, 11:62-12:4, 13:1-8, 13:15-19 U.S. Patent No. 6,122,647 at 5:65-6:4 U.S. Patent No. 7,451,099 at 3:1-3, 14:4-9, 14:25-36, 27:30-34; Figs. 3, 23 Finkelstein at p. 410 Oracle Text White Paper, p. 19 Pretschner at p. 2.
10	The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary.	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15

		PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 16:5-9 Weiss at p. 185 Mase at pgs. 377, 379 U.S. Patent No. 6,236,768 at 4:45-55 U.S. Patent No. 6,122,647 at 3:37-39, 5:50-55 U.S. Patent No. 6,473,752 at 6:45-60, 12:3-11; Fig. 9 Oracle Text White Paper, p. 19 Pretschner at p. 2.
18	An article of manufacture for use in a computer system, comprising:	See Figs. 3, 4; Claim 33. "A client computer for presenting advertising to a user, comprising:. . . [A] memory that stores browser software adapted to be executed . . . ." Claim 33.
	a memory;	See Figs. 3, 4. The computers, which are used to carry out the advertising selection processes disclosed in the patent, inherently include memory. "A client computer for presenting advertising to a user, comprising: . . . [A] memory that stores browser software adapted to be executed . . . ." Claim 33.
	instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	See Chart for Claim 1 (above). See Figs. 3, 4. The advertising selection process inherently requires the storage of instructions in memory that are executed on the computer.

		"A client computer for presenting advertising to a user, comprising: . . . [A] memory that stores browser software adapted to be executed . . . ." Claim 33.
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	See Chart for Claim 1 (above).
	automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	See Chart for Claim 1 (above).
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).
19	The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content."	See Chart for Claim 2 (above).

Appendix F

By identifying in this chart where a limitation of an asserted claim is found in a reference, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case.

	Claims of the '979 Patent	**U.S. Patent No. 7,451,099**
1	A method for automatically generating a query from selected document content, comprising:	"When a new document (e.g. a web page) is displayed on the client system to an end user, selected context associated from the document is analyzed for selected keywords. Specific context in the document may then be identified using the selected keyword information. Based upon the identified context in the document, a selected pop-up advertisement may be automatically displayed on the client system." 2:42-49. "It will be appreciated that the technique of the present invention enables businesses and advertisers to proactively interact with existing and potential on-line customers by marking up (e.g. underlining, highlighting, displaying additional text, graphics, and/or sound) selected keywords or phrases on any document, web page or web page which is currently being displayed on the user's computer system. In this way, static HTML pages may be converted at the user's computer system into customized, dynamic information which provides the ability for businesses and advertisers to proactively deliver dynamic, targeted and customized service to the end users via additional information." 44:8-19. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 1, 2, 4, 15; Fig. 4 U.S. Patent No. 6,546,386 at Abstract, 1:47-54, 1:64-67, 2:10-14, 2:26-29, 2:42-45, 2:57-61, 3:29-35, 3:37-45, 3:55-57, 4:18-20; Fig. 1

		PCT Application Pub. No. WO 01/44992 at 2:16-19, 5:8-12 U.S. Patent No. 6,236,768 at 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein at p. 406 U.S. Patent No. 6,122,647 at 2:42-51, 3:20-23 U.S. Patent No. 6,473,752 at Abstract; Fig. 2 U.S. Patent No. 6,606,644 at 5:3-8, 5:30-39 U.S. Patent No. 6,829,780 at Abstract U.S. Patent No. 7,225,142 at 6:60-67 Pretschner at p. 1.
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	"Briefly, the contextual inventory is organized and categorized into Super Categories 302, Sub-Categories 304, and Keywords 306, as shown in FIG. 3. In accordance with one specific embodiment, this organizational tree is applied to organize the Keywords and/or phrases under their appropriate product Categories." 14:4-9. "The EZ Gateway 204 also performs category management tasks such as permitting the Ad Campaign Provider to enhance their ontology (the database of categories and keywords) on an ongoing basis, in real time." 9:20-25. "According to a specific embodiment, the contextual pop-ups media feature is based on the ability to identify keywords on the page, classify them into categories, and using the category assign a matching category to a given page. In order to illustrate this aspect of the present invention, an example will now be described in which it is assumed that a document (e.g. web page) is displayed on the user's computer system which includes the following text: truck, car, vehicle, SUV, sport car. In this particular example, the document may be classified as a page corresponding to the category name 'Auto'. Accordingly, in one implementation, it will be appropriate to display information from the 'Auto'

		category to the end user. In this way, the technique in the present invention provides a benefit of automatically displaying advertisements which match specific context of the page or documents displayed to the end user." 27:14-29. "According to a specific embodiment, the Category ID field 802 may be used to identify a specific category (e.g. 304 of FIG. 3) associated with specific keywords, key phrases, or titles. In one implementation, the Category ID value may be represented as a 4-byte integer." 29:25-29. See Figs. 3, 8, 23, 25C; 13:25-55. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 14, 15 Mase at Abstract; Fig. 1; p. 377-378 PCT Application Pub. No. WO 01/44992 at 15:8-11, 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 2:15-24, 4:20-27, 4:45-55 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19, 19:51-56; Fig. 8c Weiss at Fig. 3; pgs. 181, 184 Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56; Fig. 5 U.S. Patent No. 6,473,752 at 1:44-53, 2:12-26, 6:24-44, 6:45-60 U.S. Patent No. 6,606,644 at 11:4-17

		U.S. Patent No. 6,829,780 at 5:55-6:10, 11:16-40 U.S. Patent No. 7,225,142 at 9:33-45 Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 1 and 2.
	automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"According to a specific embodiment, the search engine is designed to support different business requirements. It may operate in a variety of search modes, including an exact search mode and a fuzzy search mode. The search engine may search the document text, WEB PAGE, title, Meta tags, or any other property of the selected document for selected key words or phrases. In one embodiment, a search is conducted by analyzing words in the text of a selected document to see if it includes specified keywords or phrases." 4:56-64. "Accordingly, when one of the Clients 110 is surfing the Internet, regardless of what web page they are viewing, the Client Application scans the text of the web page, analyzes the context, and marks up keywords and/or phrases." 7:22-25. "According to a specific embodiment, the contextual pop-ups media feature is based on the ability to identify keywords on the page, classify them into categories, and using the category assign a matching category to a given page. In order to illustrate this aspect of the present invention, an example will now be described in which it is assumed that a document (e.g. web page) is displayed on the user's computer system which includes the following text: truck, car, vehicle, SUV, sport car. In this particular example, the document may be classified as a page corresponding to the category name 'Auto'. Accordingly, in one implementation, it will be appropriate to display information from the 'Auto' category to the end user. In this way, the technique in the present invention provides a benefit of automatically displaying advertisements which match specific context of the page or documents displayed to the end user." 27:14-29. "Further, according to one implementation, different types of context within the document (e.g. document title, Meta keywords, Meta information, document text, etc.) may be weighted differently to emphasize each type's particular relevance. If more than one advertisement is associated with a particular campaign, selection of the appropriate advertisement may be based upon different

		mechanisms such as, for example, assigned priority, round robin, relative age, etc." 27:65-28:5. See Figs. 16A, 16B. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 15 PCT Application Pub. No. WO 01/44992 at 14:19 to 15:5 U.S. Patent No. 6,236,768 at 1:56-2:6, 2:15-54, 13:19-34 U.S. Patent No. 6,546,386 at 2:26-29, 2:42-45, 3:29-35 Finkelstein at pgs. 410, 408 U.S. Patent No. 6,122,647 at 6:49-55, 7:46-49, 9:27-28, 12:27-32; Figs. 4c-4d, 5, 7 U.S. Patent No. 6,473,752 at 3:5-8, 3:14-23, 5:62-6:7, 9:26-29; Figs. 2, 5, 8, 10B U.S. Patent No. 6,606,644 at 11:4-11, 12:52-63 U.S. Patent No. 6,829,780 at 5:55-6:14, 9:34-36 U.S. Patent No. 7,225,142 at 9:24-32 Oracle Text White Paper, p. 8, 9 Pretschner at p. 2.
	automatically categorizing the selected document content using the organized classification of document content for	"According to a specific embodiment, the contextual pop-ups media feature is based on the ability to identify keywords on the page, classify them into categories, and using the category assign a matching category to a given page. In order to illustrate this aspect of the present invention, an example will now

	assigning the selected document content a classification label from the organized classification of content; and	be described in which it is assumed that a document (e.g. web page) is displayed on the user's computer system which includes the following text: truck, car, vehicle, SUV, sport car. In this particular example, the document may be classified as a page corresponding to the category name 'Auto'. Accordingly, in one implementation, it will be appropriate to display information from the 'Auto' category to the end user. In this way, the technique in the present invention provides a benefit of automatically displaying advertisements which match specific context of the page or documents displayed to the end user." 27:14-29. "In a specific embodiment, the MAIN application 520 may be configured to analyze a selected document for keywords, categories and/or super categories in order to find a match for an appropriate pop-up advertisement or window to be displayed." 27:30-34. "According to a specific embodiment, one or more algorithms may be used for determining the most appropriate matching category for the selected document being analyzed. For example, in one algorithm, a variety of different parameters relating to the current document may be analyzed in order to determine the most appropriate matching category." 27:46-51. See Fig. 16A. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 Mase at p. 377-379; Fig. 1 PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 5:12-28 U.S. Patent No. 6,546,386 at 2:10-18, 2:57-61 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-18, 20:19-29

		Weiss at pgs. 181, 182, 185, 186, 191 Finkelstein at pgs. 408, 410 U.S. Patent No. 6,122,647 at 3:25-35, 7:66-8:5; Figs. 6-8 U.S. Patent No. 6,473,752 at 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 11:16-40 U.S. Patent No. 7,225,142 at 9:33-37 Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 2.
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	"According to a specific embodiment, one or more algorithms may be used for determining the most appropriate matching category for the selected document being analyzed. For example, in one algorithm, a variety of different parameters relating to the current document may be analyzed in order to determine the most appropriate matching category. For example, the current document may be analyzed and assigned a specific context score (CS) that is then compared with specific campaign requirement included in the campaign update files. If the context score is greater than or equal to a predetermined threshold value TH, then a pop-up ad (or other media type ads) may be displayed. According to one implementation, keywords which are identified in different elements of the document may be scored appropriately. The cumulative score of all the keywords that are found may be used to determine the CS value. If the identified keywords match a specific category of an ad campaign, and the cumulative CS value is above the threshold for that campaign, then a pop-up advertisement for that campaign may be displayed. Further, according to one implementation, different types of context within the document (e.g. document title, Meta keywords, Meta information, document text, etc.) may be weighted differently to emphasize each type's particular relevance. If more than one advertisement is associated with a particular campaign, selection of the appropriate advertisement may be based upon different mechanisms such as, for example, assigned priority, round robin, relative age, etc." 27:46-

		28:5. "FIG. 7 shows a specific embodiment of a flow diagram illustrating how various information flows are passed between the client system and the server system of the present invention. Initially, at (30) it is assumed that the user has clicked or selected a particular portion of text which has been marked up in accordance with the technique of the present invention. According to at least one embodiment, when the user clicks on a particular portion of marked up text, a pop-up layer (e.g. dynamic browser control layer) may be displayed (31) to the user providing the user with additional information relating to the topic of the marked up text portion. An example of one type of pop-up layer is illustrated in FIG. 21 of the drawings. According to specific embodiments, the pop-up layer may include, for example, one or more links, audio information, video information, and/or textual information." 28:30-44. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15; Fig. 4. PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:14-22 U.S. Patent No. 6,236,768 at 5:12-28 U.S. Patent No. 6,546,386 at 3:37-45, 3:55-57 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-29 Weiss at pgs. 181, 182, 184, 185, 186 Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18; Figs. 5-6, 8 U.S. Patent No. 6,473,752 at 4:23-29, 7:51-8:4, 10:34-38, 15:19-23; Figs. 2, 7, 10B

		U.S. Patent No. 6,606,644 at 11:4-17, 12:2-7 U.S. Patent No. 6,829,780 at 13:40-50, 13:64-67, 14:43-52 U.S. Patent No. 7,225,142 at 4:20-22, 9:38-45 Pretschner at p. 1 and 2.
2	The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	"Further, according to one implementation, different types of context within the document (e.g. document title, Meta keywords, Meta information, document text, etc.) may be weighted differently to emphasize each type's particular relevance. If more than one advertisement is associated with a particular campaign, selection of the appropriate advertisement may be based upon different mechanisms such as, for example, assigned priority, round robin, relative age, etc." 27:65-28:5. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410

		Pretschner at p. 2.
3	The method according to claim 2, wherein the number of terms added is limited to a predefined number	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.
5	The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	"FIG. 3 is a schematic diagram of the context hierarchy in accordance with a specific embodinment of the present invention." 3:1-3. "Briefly, the contextual inventory is organized and categorized into Super Categories 302, Sub-Categories 304, and Keywords 306, as shown in FIG. 3. In accordance with one specific embodiment, this organizational tree is applied to organize the Keywords and/or phrases under their appropriate product Categories." 14:4-9.

		"In one implementation, the hierarchy of the Super-Category 302 is designed to provide Keywords that can apply to multiple Categories at a highest level and for very specific Keywords at a lower level. For example, a Credit Card company may be offered Keywords in the "Personal Finance" Super-Category, such as Keyword "credit", and then be offered Keywords in the Sub-Categories "Personal Finance-Credit Cards" and "Personal Finance-Credit Cards-Low Rate", such as the Keywords "credit card" and "low rate credit card", respectively. This is but one organizational example, and it will be appreciated that the such categorization be adjusted according to need. 14:25-36 "In a specific embodiment, the MAIN application 520 may be configured to analyze a selected document for keywords, categories and/or super categories in order to find a match for an appropriate pop-up advertisement or window to be displayed." 27:30-34. See Figs. 3, 23. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19; Fig. 8c Weiss at pgs. 180, 184; Fig 3 U.S. Patent No. 6,236,768 at 12:41-49, 11:62-12:4, 13:1-8, 13:15-19 U.S. Patent No. 6,122,647 at 5:65-6:4 Finkelstein at p. 410 Oracle Text White Paper, p. 19 Pretschner at p. 2.

10	The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 16:5-9 Weiss at p. 185 Mase at pgs. 377, 379 U.S. Patent No. 6,236,768 at 4:45-55 U.S. Patent No. 6,122,647 at 3:37-39, 5:50-55 U.S. Patent No. 6,473,752 at 6:45-60, 12:3-11; Fig. 9 Oracle Text White Paper, p. 19 Pretschner at p. 2.
18	An article of manufacture for use in a computer system, comprising:	"Generally, the various techniques of the present invention may be implemented on software and/or hardware." 45:37-38. See Fig. 22.
	a memory;	"A software or software/hardware hybrid implementation of the various technique of this invention may be implemented on a general-purpose programmable machine selectively activated or reconfigured by a computer program stored in memory." 45:46-50. See Fig. 22.

	instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	See Chart for Claim 1 (above). "A software or software/hardware hybrid implementation of the various technique of this invention may be implemented on a general-purpose programmable machine selectively activated or reconfigured by a computer program stored in memory." 45:46-50. See Fig. 22.
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	See Chart for Claim 1 (above).
	automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	See Chart for Claim 1 (above).
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).
19	The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the	See Chart for Claim 2 (above).

	query by adding terms relating to context information surrounding the set of entities in the selected document content."	

Appendix G

By identifying in this chart where a limitation of an asserted claim is found in a reference, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case.

	Claims of the '979 Patent	**Finkelstein et al., *Placing Search in Context: The Concept Revisited*, Proc. of the 10th International World Wide Web Conference (May 1-5, 2001)**
1	A method for automatically generating a query from selected document content, comprising:	"In the IntelliZap system we developed, search is initiated from a text query marked by the user in a document she views, and is guided by the text surrounding the marked query in that document ('the context'). The context-guided information retrieval process involves semantic keyword extraction and clustering to automatically generate new, augmented queries. The latter are submitted to a host of general and domain-specific search engines." Page 406. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 1, 2, 4, 15; Fig. 4 U.S. Patent No. 6,546,386 at Abstract, 1:47-54, 1:64-67, 2:10-14, 2:26-29, 2:42-45, 2:57-61, 3:29-35, 3:37-45, 3:55-57, 4:18-20; Fig. 1 PCT Application Pub. No. WO 01/44992 at 2:16-19, 5:8-12 U.S. Patent No. 6,236,768 at 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 2:42-51, 3:20-23 U.S. Patent No. 6,473,752 at Abstract; Fig. 2

		U.S. Patent No. 6,606,644 at 5:3-8, 5:30-39 U.S. Patent No. 6,829,780 at Abstract U.S. Patent No. 7,225,142 at 6:60-67 U.S. Patent No. 7,451,099 at 2:42-49, 44:8-19 Pretschner at p. 1.
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	"The classification algorithm classifies the context to a limited number of high-level domains[11] (e.g., medicine or law). A probabilistic analysis determines the amount of similarity between the domain signatures and the query context. The *a priori* assignment of search engines to domains is performed offline. Some of the search engines (such as AltaVista) allow limiting the search to a specific category. In such cases, categorizing the query in order to further constrain the search usually yields superior results." Page 410. FN 11: "Currently, nine domains are defined, each of which is mapped to two or three search engines." To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 14, 15 Mase at Abstract; Fig. 1; p. 377-378 PCT Application Pub. No. WO 01/44992 at 15:8-11, 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 2:15-24, 4:20-27, 4:45-55 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19, 19:51-56; Fig. 8c

		Weiss at Fig. 3; pgs. 181, 184 U.S. Patent No. 6,122,647 at 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56; Fig. 5 U.S. Patent No. 6,473,752 at 1:44-53, 2:12-26, 6:24-44, 6:45-60 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 5:55-6:10, 11:16-40 U.S. Patent No. 7,225,142 at 9:33-45 U.S. Patent No. 7,451,099 at 9:20-25, 13:25-55, 14:4-9, 27:14-29, 29:25-29; Figs 3, 8, 23, 25C Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 1 and 2.
	automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"The algorithm utilizes the semantic network to extract keywords from the context surrounding the user-selected text. These keywords are added to the text to form an augmented query. . . ." Page 410 "The IntelliZap system has three main components based on the semantic network: 1. Extracting keywords from the captured text and context. …" Page 410. "The context may include the sentence containing the query word or phrase, a few sentences surrounding the query term, the paragraph in which it resides, or even the whole document." Page 408. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 15

		PCT Application Pub. No. WO 01/44992 at 14:19 to 15:5 U.S. Patent No. 6,236,768 at 1:56-2:6, 2:15-54, 13:19-34 U.S. Patent No. 6,546,386 at 2:26-29, 2:42-45, 3:29-35 U.S. Patent No. 6,122,647 at 6:49-55, 7:46-49, 9:27-28, 12:27-32; Figs. 4c-4d, 5, 7 U.S. Patent No. 6,473,752 at 3:5-8, 3:14-23, 5:62-6:7, 9:26-29; Figs. 2, 5, 8, 10B U.S. Patent No. 6,606,644 at 11:4-11, 12:52-63 U.S. Patent No. 6,829,780 at 5:55-6:14, 9:34-36 U.S. Patent No. 7,225,142 at 9:24-32 U.S. Patent No. 7,451,099 at 4:56-64, 7:22-25, 27:14-29, 27:65-28:5; Figs. 16A, 16B Oracle Text White Paper, p. 8, 9 Pretschner at p. 2.
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	"[W]e attempt to classify the captured context in order to select domain-specific search engines that stand a good chance of providing more specialized results. The classification algorithm classifies the context to a limited number of high-level domains (e.g., medicine or law). A probabilistic analysis determines the amount of similarity between the domain signatures and the query context." Page 410 (footnote omitted). "The IntelliZap system has three main components based on the semantic network: … 2. High-level classification of the query to a small set of predefined domains. …" Page 410.

		"The context may include the sentence containing the query word or phrase, a few sentences surrounding the query term, the paragraph in which it resides, or even the whole document." Page 408. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 Mase at p. 377-379; Fig. 1 PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 5:12-28 U.S. Patent No. 6,546,386 at 2:10-18, 2:57-61 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-18, 20:19-29 Weiss at pgs. 181, 182, 185, 186, 191 U.S. Patent No. 6,122,647 at 3:25-35, 7:66-8:5; Figs. 6-8 U.S. Patent No. 6,473,752 at 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 11:16-40 U.S. Patent No. 7,225,142 at 9:33-37 U.S. Patent No. 7,451,099 at 27:14-34, 27:46-51; Fig 16A Oracle Text White Paper, p. 11, 18, 19

		Pretschner at p. 2.
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	"The algorithm utilizes the semantic network to extract keywords from the context surrounding the user-selected text. These keywords are added to the text to form an augmented query. . . ." Page 410 "[W]e attempt to classify the captured context in order to select domain-specific search engines that stand a good chance of providing more specialized results. . . . Some of the search engines (such as AltaVista) allow limiting the search to a specific category. In such cases, categorizing the query in order to further constrain the search usually yields superior results." Page 410. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15; Fig. 4. PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:14-22 U.S. Patent No. 6,236,768 at 5:12-28 U.S. Patent No. 6,546,386 at 3:37-45, 3:55-57 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-29 Weiss at pgs. 181, 182, 184, 185, 186 U.S. Patent No. 6,122,647 at 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18; Figs. 5-6, 8 U.S. Patent No. 6,473,752 at 4:23-29, 7:51-8:4, 10:34-38, 15:19-23; Figs. 2, 7, 10B U.S. Patent No. 6,606,644 at 11:4-17, 12:2-7 U.S. Patent No. 6,829,780 at 13:40-50, 13:64-67, 14:43-52

		U.S. Patent No. 7,225,142 at 4:20-22, 9:38-45 U.S. Patent No. 7,451,099 at 27:46-28:5, 28:30-44 Pretschner at p. 1 and 2.
2	The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	"The algorithm utilizes the semantic network to extract keywords from the context surrounding the user-selected text. These keywords are added to the text to form an augmented query. . . ." Page 410 "The IntelliZap system has three main components based on the semantic network: 3. Extracting keywords from the captured text and context. …" Page 410. "The context may include the sentence containing the query word or phrase, a few sentences surrounding the query term, the paragraph in which it resides, or even the whole document." Page 408. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50

		U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Pretschner at p. 2.
3	The method according to claim 2, wherein the number of terms added is limited to a predefined number.	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Pretschner at p. 2.
5	The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	"Some of the search engines (such as AltaVista) allow limiting the search to a specific category. In such cases, categorizing the query in order to further constrain the search usually yields superior results." Page 410. To the extent this reference does not teach this claim element, this reference in combination with the

		knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19; Fig. 8c Weiss at pgs. 180, 184; Fig 3 U.S. Patent No. 6,236,768 at 12:41-49, 11:62-12:4, 13:1-8, 13:15-19 U.S. Patent No. 6,122,647 at 5:65-6:4 U.S. Patent No. 7,451,099 at 3:1-3, 14:4-9, 14:25-36, 27:30-34; Figs. 3, 23 Oracle Text White Paper, p. 19 Pretschner at p. 2.
10	The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary.	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 16:5-9 Weiss at p. 185 Mase at pgs. 377, 379 U.S. Patent No. 6,236,768 at 4:45-55

		U.S. Patent No. 6,122,647 at 3:37-39, 5:50-55 U.S. Patent No. 6,473,752 at 6:45-60, 12:3-11; Fig. 9 Oracle Text White Paper, p. 19 Pretschner at p. 2.
18	An article of manufacture for use in a computer system, comprising:	"Our system (named IntelliZap) is based on the client-server paradigm, where a client application running on user's computer captures the context around the text highlighted by the user. The server-based algorithms analyze the context, selecting most important context words and performing word sense disambiguation, and then prepare a set of augmented queries for subsequent search." Page 406.
	a memory;	"Our system (named IntelliZap) is based on the client-server paradigm, where a client application running on user's computer captures the context around the text highlighted by the user. The server-based algorithms analyze the context, selecting most important context words and performing word sense disambiguation, and then prepare a set of augmented queries for subsequent search." Page 406. The client and server computers inherently include a memory.
	instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	See Chart for Claim 1 (above). "Our system (named IntelliZap) is based on the client-server paradigm, where a client application running on user's computer captures the context around the text highlighted by the user. The server-based algorithms analyze the context, selecting most important context words and performing word sense disambiguation, and then prepare a set of augmented queries for subsequent search." Page 406. The client and server computers inherently store the IntelliZap software in memory in order to make it executable.
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an	See Chart for Claim 1 (above).

	information retrieval system;	
	automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	See Chart for Claim 1 (above).
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).
19	The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content."	See Chart for Claim 2 (above).

Appendix H

By identifying in this chart where a limitation of an asserted claim is found in a reference, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case.

	Claims of the '979 Patent	**U.S. Patent No. 6,236,768**
1	A method for automatically generating a query from selected document content, comprising:	"The RA works in two stages. First, the user's collection of text documents is indexed into a database saved in a vector format. These form the reservoir of documents from which later suggestions of relevance are drawn; that is, stored documents will later be 'suggested' as being relevant to a document currently being edited or read. The store documents can be any sort of text document (notes, Usenet entries, webpages, e-mail, etc.). This indexing is usually performed automatically every night, and the index files are stored in a database. After the database is created, the other stage of the RA is run from Emacs, periodically taking a sample of text from the working buffer. The RA finds documents "similar" to the current sample according to word similarities; that is, the more times a word in the current sample is duplicated in a candidate database document, the greater will be assumed the relevance of that database document. The RA displays one-line summaries of the best few documents at the bottom of the Emacs window." 1:56-2:6. "Analysis module 133 first indexes all the documents in a corpus of data (which, again, are stored as files mass storage device 106, which is assumed for explanatory purposes to be a hard disk), and writes indices to disk. Unlike the RA, the invention preferably keeps several vectors for each document. These include not only the wordvec vector for text (if any) in the document but also vectors for meta-information, e.g., subject, people, time, date, day of week, location, etc." 10:42-51. "4. Determination of relevance For each element of each discrete vector in a query - the generation and vectorization of which is described below - the algorithm used by the RA may be used to determine relevance to documents in the corpus." 12:53-57.

		"5. Weighted addition of vectors The result of the foregoing operations is a single similarity value for each type of meta-information. These values are associated with each document in the indexed corpus, and are used to compute the overall similarity using bias values for query and document types, by the following formula: Query biases = bq pq sq lq dq etc. (i.e., body_query_bias, person_query_bias, etc.)" 13:1-8. "Each vector similarity is multiplied by its respective bias and the resulting biased similarity is summed, to produce an overall similarity between zero and one." 13:15-19. "Analysis module 133 supplies a ranked list of the most relevant documents, which may be continually, intermittently, or upon request presented to the user over display 126. If desired, or upon user command, the list may be pruned to include only documents whose relevance level exceeds a predetermined threshold." 13:42-47. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 1, 2, 4, 15; Fig. 4 U.S. Patent No. 6,546,386 at Abstract, 1:47-54, 1:64-67, 2:10-14, 2:26-29, 2:42-45, 2:57-61, 3:29-35, 3:37-45, 3:55-57, 4:18-20; Fig. 1 PCT Application Pub. No. WO 01/44992 at 2:16-19, 5:8-12 Finkelstein at p. 406 U.S. Patent No. 6,122,647 at 2:42-51, 3:20-23 U.S. Patent No. 6,473,752 at Abstract; Fig. 2

		U.S. Patent No. 6,606,644 at 5:3-8, 5:30-39 U.S. Patent No. 6,829,780 at Abstract U.S. Patent No. 7,225,142 at 6:60-67 U.S. Patent No. 7,451,099 at 2:42-49, 44:8-19 Pretschner at p. 1.
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	Table 2: (int) (width*uns int) (int) (uns int) (uns int) (uns int) NUM_WORDS, WORDCODE-1, NUM_DOCS=N1, DOC-1, DOC-2, . . . , DOC-N1, WORDCODE-2, NUM_DOCS=N2, DOC-1, DOC-2, . . . , DOC-N2, etc. "Briefly, the concept behind the indexing scheme used in RA is that any given document may be represented by a multidimensional vector, each dimension or entry of which corresponds to a single word and is equal in magnitude to the number of times that word appears in the document. ... The advantages gained by this representation are relatively speedy disk retrieval, and an easily computed quantity indicating similarity between two documents: the dot product of their (normalized) vectors." 2:15-24. "Experience with the RA has shown that actually performing a dot product with each indexed document is prohibitively slow for large databases. In preferred implementations, therefore, document vectors are not stored; instead, word vectors are stored. The 'wordvec' file contains each word appearing in the entire indexed corpus of documents followed by a list of each document that contains that particular word." 4:20-27. "Each word in the wordvec is represented by a unique numerical code, the 'width' indicating the

		number of integers in the code (the RA uses two integers per code). The NUM_DOCS field indicates the number of documents containing the word specified by the associated wordcode. The word-count variables DOC-1, DOC-2,..., DOC-N1 each correspond to a document containing the word, and reflect the number of occurrences of the word divided by the total number of words in the document." 4:45-55. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 14, 15 Mase at Abstract; Fig. 1; p. 377-378 PCT Application Pub. No. WO 01/44992 at 15:8-11, 17:15-32, 18:16-22 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19, 19:51-56; Fig. 8c Weiss at Fig. 3; pgs. 181, 184 Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56; Fig. 5 U.S. Patent No. 6,473,752 at 1:44-53, 2:12-26, 6:24-44, 6:45-60 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 5:55-6:10, 11:16-40 U.S. Patent No. 7,225,142 at 9:33-45 U.S. Patent No. 7,451,099 at 9:20-25, 13:25-55, 14:4-9, 27:14-29, 29:25-29; Figs 3, 8, 23, 25C

		Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 1 and 2.
	automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"Analysis module 133 preferably generates queries autonomously from the current document in document buffer 140 or by reference to a current context. In the former case, analysis module 133 classifies the document either by its header or by reference to a template, and extracts the appropriate meta-information. In the latter case, the user's physical or interpersonal surroundings furnish the meta-information upon which the query is based. It is not necessary for the documents searched or identified to correspond in type to a current document. Furthermore, the query may not be limited to meta-information. Instead, the invention may utilize both a meta-information component (with relevance to candidate documents determined as discussed above) and a text component (with relevance determined in accordance with RA)." 13:19-34. "The RA works in two stages. First, the user's collection of text documents is indexed into a database saved in a vector format. These form the reservoir of documents from which later suggestions of relevance are drawn; that is, stored documents will later be 'suggested' as being relevant to a document currently being edited or read. The store documents can be any sort of text document (notes, Usenet entries, webpages, e-mail, etc.). This indexing is usually performed automatically every night, and the index files are stored in a database. After the database is created, the other stage of the RA is run from Emacs, periodically taking a sample of text from the working buffer. The RA finds documents 'similar' to the current sample according to word similarities; that is, the more times a word in the current sample is duplicated in a candidate database document, the greater will be assumed the relevance of that database document. The RA displays one-line summaries of the best few documents at the bottom of the Emacs window." 1:56-2:6. "Briefly, the concept behind the indexing scheme used in RA is that any given document may be represented by a multidimensional vector, each dimension or entry of which corresponds to a single word and is equal in magnitude to the number of times that word appears in the document. ... RA creates vectors in three steps: ... Step 1: Remove stop words ... Step 2: Stem words ... Step 3: Make the document vector." 2:15-54. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.:

		U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 15 PCT Application Pub. No. WO 01/44992 at 14:19 to 15:5 U.S. Patent No. 6,546,386 at 2:26-29, 2:42-45, 3:29-35 Finkelstein at pgs. 410, 408 U.S. Patent No. 6,122,647 at 6:49-55, 7:46-49, 9:27-28, 12:27-32; Figs. 4c-4d, 5, 7 U.S. Patent No. 6,473,752 at 3:5-8, 3:14-23, 5:62-6:7, 9:26-29; Figs. 2, 5, 8, 10B U.S. Patent No. 6,606,644 at 11:4-11, 12:52-63 U.S. Patent No. 6,829,780 at 5:55-6:14, 9:34-36 U.S. Patent No. 7,225,142 at 9:24-32 U.S. Patent No. 7,451,099 at 4:56-64, 7:22-25, 27:14-29, 27:65-28:5; Figs. 16A, 16B Oracle Text White Paper, p. 8, 9 Pretschner at p. 2.
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	"Experience with the RA has shown that actually performing a dot product with each indexed document is prohibitively slow for large databases. In preferred implementations, therefore, document vectors are not stored; instead, word vectors are stored. The 'wordvec' file contains each word appearing in the entire indexed corpus of documents followed by a list of each document that contains that particular word." 4:20-27.

		(int) (width*uns int) (int) (uns int) (uns int) (uns int) NUM_WORDS, WORDCODE-1, NUM_DOCS=N1, DOC-1, DOC-2, . . . , DOC-N1, WORDCODE-2, NUM_DOCS=N2, DOC-1, DOC-2, . . . , DOC-N2, etc. 4:32-40. "Accordingly, for each word in the query vector, the RA first looks up the word in the word offset file, and from that the word's entry is looked up in the wordvec file. An array of document similarities is used to maintain a running tally of documents and their similarities, in terms of numbers of word matches, to the query vector. The array is sorted by similarity, with the most similar documents at the top of the list. Similarity is computed for each word in the query vector by taking the product of the query-vector entry and the weight of each document in the corresponding wordvec file. To normalize this product, it is then divided by the query-vector magnitude (computed in the same manner as the document magnitude) and also by the document magnitude. The final value is added to the current running-total similarity for that document, and the process is repeated for the next word in the query. In summary, the query vector is analyzed wordcode by wordcode, with the similarities array indicating the relevance to the query of each document." 5:12-28. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 Mase at p. 377-379; Fig. 1 PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:16-22 U.S. Patent No. 6,546,386 at 2:10-18, 2:57-61 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-18, 20:19-29

		Weiss at pgs. 181, 182, 185, 186, 191 Finkelstein at pgs. 408, 410 U.S. Patent No. 6,122,647 at 3:25-35, 7:66-8:5; Figs. 6-8 U.S. Patent No. 6,473,752 at 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 11:16-40 U.S. Patent No. 7,225,142 at 9:33-37 U.S. Patent No. 7,451,099 at 27:14-34, 27:46-51; Fig 16A Oracle Text White Paper, p. 11, 18, 1 Pretschner at p. 2.
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	"Accordingly, for each word in the query vector, the RA first looks up the word in the word offset file, and from that the word's entry is looked up in the wordvec file. An array of document similarities is used to maintain a running tally of documents and their similarities, in terms of numbers of word matches, to the query vector. The array is sorted by similarity, with the most similar documents at the top of the list. Similarity is computed for each word in the query vector by taking the product of the query-vector entry and the weight of each document in the corresponding wordvec file. To normalize this product, it is then divided by the query-vector magnitude (computed in the same manner as the document magnitude) and also by the document magnitude. The final value is added to the current running-total similarity for that document, and the process is repeated for the next word in the query. In summary, the query vector is analyzed wordcode by wordcode, with the similarities array indicating the relevance to the query of each document." 5:12-28.

		To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15; Fig. 4. PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:14-22 U.S. Patent No. 6,546,386 at 3:37-45, 3:55-57 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-29 Weiss at pgs. 181, 182, 184, 185, 186 Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18; Figs. 5-6, 8 U.S. Patent No. 6,473,752 at 4:23-29, 7:51-8:4, 10:34-38, 15:19-23; Figs. 2, 7, 10B U.S. Patent No. 6,606,644 at 11:4-17, 12:2-7 U.S. Patent No. 6,829,780 at 13:40-50, 13:64-67, 14:43-52 U.S. Patent No. 7,225,142 at 4:20-22, 9:38-45 U.S. Patent No. 7,451,099 at 27:46-28:5, 28:30-44 Pretschner at p. 1 and 2.
2	The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected	"Analysis module 133 preferably generates queries autonomously from the current document in document buffer 140 or by reference to a current context. In the former case, analysis module 133 classifies the document either by its header or by reference to a template, and extracts the appropriate meta-information. In the latter case, the user's physical or interpersonal surroundings furnish the meta-

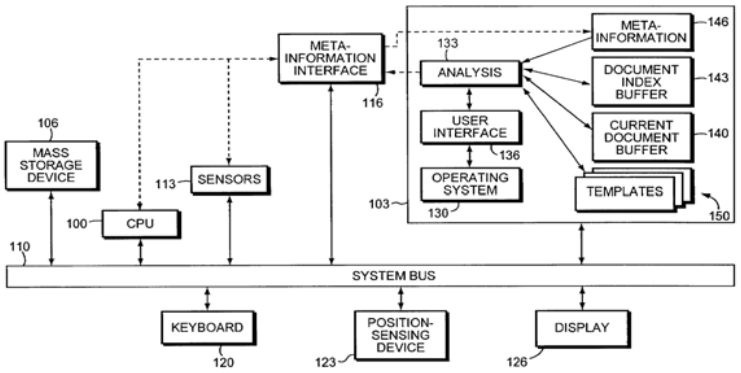
	document content.	information upon which the query is based. It is not necessary for the documents searched or identified to correspond in type to a current document. Furthermore, the query may not be limited to meta-information. Instead, the invention may utilize both a meta-information component (with relevance to candidate documents determined as discussed above) and a text component (with relevance determined in accordance with RA).” 13:19-34. “2. After the document is identified, different fields are extracted, again based on the template. For example, the email template continues: Delimiter {startline, "From"} Format {{anyorder {startline, "From: ", PERSON, ".backslash.n"} {startline, "Date: ", DATE, ".backslash.n"} optional {startline, "Subject: ", SUBJECT, ".backslash.n"}} ".backslash.n.backslash.n", BODY} } Bias 21100000 The delimiter command explicitly identifies the separator between one document of this template type and another, should they both reside in the same file. (For example, a plain e-mail archive may contain several pieces of mail in the same file, all separated by the word "From" plus a space at the start of a line.) The remainder of the template specifies that the "From:" line contains the person or people associated with this document. and the line starting with "Date:" contains the date/timestamp of the document.” 11:10-31. “Analysis module 133 first indexes all the documents in a corpus of data (which, again, are stored as files mass storage device 106, which is assumed for explanatory purposes to be a hard disk), and writes indices to disk. Unlike the RA, the invention preferably keeps several vectors for each document. These include not only the wordvec vector for text (if any) in the document but also vectors for metainformation, e.g., subject, people, time, date, day of week, location, etc.)” 10:42-51. “Any meta-information that can be represented by text (e.g., subject lines, room names, people names, bodies of text, etc.) is encoded in accordance with the above scheme. Like the body of text, each word in these text strings is encoded separately and added to a vector. Vectors of discrete (text) data are all

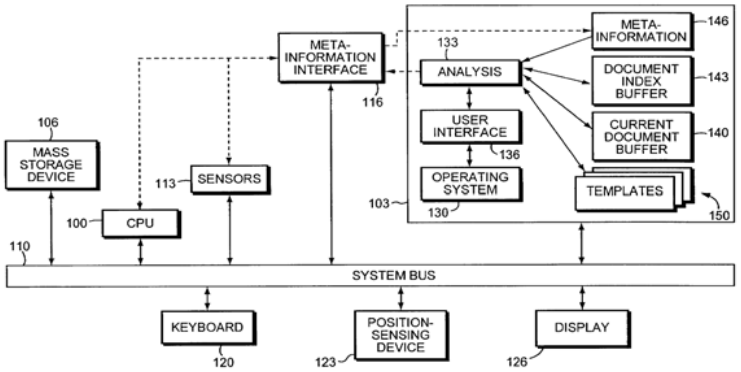
		stored in one file, but the vectors are still conceptually distinct and are distinguished by their type bits. The file format for discrete type information is the same as the wordvec file format." 12:41-49. "5. Weighted addition of vectors The result of the foregoing operations is a single similarity value for each type of meta-information. These values are associated with each document in the indexed corpus, and are used to compute the overall similarity using bias values for query and document types, by the following formula: Query biases = bq pq sq lq dq etc. (i.e., body_query_bias, person_query_bias, etc.)." 13:1-8. "Each vector similarity is multiplied by its respective bias and the resulting biased similarity is summed, to produce an overall similarity between zero and one." 13:15-19. "Analysis module 133 supplies a ranked list of the most relevant documents, which may be continually, intermittently, or upon request presented to the user over display 126. If desired, or upon user command, the list may be pruned to include only documents whose relevance level exceeds a predetermined threshold." 13:42-47. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50

		U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.
3	The method according to claim 2, wherein the number of terms added is limited to a predefined number.	To the extent this reference does not teach this claim element, the reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.
5	The method according to claim 1, wherein the organized classification of document	"Any meta-information that can be represented by text (e.g., subject lines, room names, people names, bodies of text, etc.) is encoded in accordance with the above scheme. Like the body of text, each word

	content is defined using a hierarchical organization.	in these text strings is encoded separately and added to a vector. Vectors of discrete (text) data are all stored in one file, but the vectors are still conceptually distinct and are distinguished by their type bits. The file format for discrete type information is the same as the wordvec file format." 12:41-49. "In accordance with the present invention, each type of meta-information is placed in its own vector, and a single vector represents each type of meta-information supported by the invention. The final entry in the template file is the bias number for the particular type of file, which ranks the fields of the file in terms of importance. In the e-mail example above, the bias means that the body of the e-mail is most important, person and date fields are secondary (in a ratio of 2 to 1 to1), and no other fields are used to compute similarity." 11:62-12:4. "5. Weighted addition of vectors The result of the foregoing operations is a single similarity value for each type of meta-information. These values are associated with each document in the indexed corpus, and are used to compute the overall similarity using bias values for query and document types, by the following formula: Query biases = bq pq sq lq dq etc. (i.e., body_query_bias, person_query_bias, etc.)." 13:1-8. "Each vector similarity is multiplied by its respective bias and the resulting biased similarity is summed, to produce an overall similarity between zero and one." 13:15-19. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19; Fig. 8c Weiss at pgs. 180, 184; Fig 3

		U.S. Patent No. 6,122,647 at 5:65-6:4 U.S. Patent No. 7,451,099 at 3:1-3, 14:4-9, 14:25-36, 27:30-34; Figs. 3, 23 Finkelstein at p. 410 Oracle Text White Paper, p. 19 Pretschner at p. 2.
10	The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary.	"Each word in the wordvec is represented by a unique numerical code, the "width" indicating the number of integers in the code (the RA uses two integers per code). The NUM_DOCS field indicates the number of documents containing the word specified by the associated wordcode. The word-count variables DOC‐1, DOC‐2, ⋯ , DOC‐N1 each correspond to a document containing the word, and reflect the number of occurrences of the word divided by the total number of words in the document." 4:45-55. To the extent this reference does not teach this claim element, the reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 16:5-9 Weiss at p. 185 Mase at pgs. 377, 379 U.S. Patent No. 6,122,647 at 3:37-39, 5:50-55

		U.S. Patent No. 6,473,752 at 6:45-60, 12:3-11; Fig. 9 Oracle Text White Paper, p. 19 Pretschner at p. 2.
18	An article of manufacture for use in a computer system, comprising:	Fig. 1:  "Refer now to FIG. 1, which illustrates, in block-diagram form, a hardware platform incorporating a representative, generalized embodiment of the invention. As indicated therein, the system includes a central-processing unit (‘CPU’) 100, which perform operations on and interacts with a main system memory 103 and components thereof. System memory 103 typically includes volatile or random-access memory (‘RAM’) for temporary storage of information, including buffers, executing programs, and portions of the computer's basic operating system. The platform typically also includes read-only memory (‘ROM’) for permanent storage of the computer's configuration and additional portions of the basic operating system, and at least one mass storage device 106, such a hard disk and/or CD-ROM drive. All components of the platform are interconnected by and communicate over, a bidirectional

		system bus 110." 9:18-34.
	a memory;	Fig. 1:  Rhodes, 9:18-34: "Refer now to FIG. 1, which illustrates, in block-diagram form, a hardware platform incorporating a representative, generalized embodiment of the invention. As indicated therein, the system includes a central-processing unit ('CPU') 100, which perform operations on and interacts with a main system memory 103 and components thereof. System memory 103 typically includes volatile or random- access memory ('RAM') for temporary storage of information, including buffers, executing programs, and portions of the computer's basic operating system. The platform typically also includes read-only memory ('ROM') for permanent storage of the computer's configuration and additional portions of the basic operating system, and at least one mass storage device 106, such a hard disk and/or CD-ROM drive. All components of the platform are interconnected by and communicate over, a bidirectional system bus 110."
	instructions stored in the memory for operating a method for automatically generating a query from selected document	"The main memory 103 contains a group of modules that control the operation of CPU 100 and its interaction with the other hardware components. These modules are implemented as executable machine instructions, running (by means of CPU 100) as active processes effectively capable of

	content, comprising:	interacting (i.e., exchanging data and control commands) as illustrated. An operating system 130 directs the execution of low-level, basic system functions such as memory allocation, file management, and operation of mass storage devices 106. At a higher level, an analyzer module 133 directs execution of the primary functions performed by the invention, as discussed below; and instructions defining a user interface 136 allow straightforward interaction over display 126. User interface 136 generates words or graphical images on display 126 to facilitate user action and examination of documents, and accepts user commands from keyboard 120 and/or position-sensing device 123." 10:1-17. See Chart for Claim 1 (above).
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	See Chart for Claim 1 (above).
	automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	See Chart for Claim 1 (above).
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).
19	The article of manufacture according to	See Chart for Claim 2 (above).

	claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	

Appendix I

By identifying in this chart where a limitation of an asserted claim is found in a reference, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case.

	Claims of the '979 Patent	**U.S. Patent Application Publication No. 2002/0147738**
1	A method for automatically generating a query from selected document content, comprising:	"Patent professionals often search for publications relevant to patents. Searches typically arise in two contexts: when looking for "prior art" publications that might invalidate a patent and when looking for publications that might disclose an infringement of a patent." ¶ 1. "An ever-increasing number of publications are being published on the Internet, for example, "white papers" published on companies' public websites. Thus, the Internet has become a more and more important resource for patent professionals looking for publications relevant to patents." ¶ 2. "However, patent professionals have for the most part relied on general Internet search techniques, such as applying keywords to general-purpose Internet search engines, to discover patent-relevant publications on the Internet." ¶ 2. "The present invention provides a highly automated search technique for discovering patent-relevant publications on the Internet. The high level of automation may be achieved with the expedient of a search client resident on an end-user station that initiates linked searches for patent data and Internet publication data in a manner transparent to a user. From the user's perspective, a patent-identifying attribute, such as an inventor name, assignee name or patent number, input on an end-user station automatically returns Internet publication data, such as Uniform Resource Locators (URLs) of Web documents. The invention thereby allows a user to find patent-relevant publications on the Internet by merely inputting a patent-identifying attribute. A patent-identifying attribute may be a patent family-identifying attribute, such as an inventor name or assignee name. Or a patent identifying-attribute may be a single patent-identifying attribute, such as a patent number. Or a patent identifying-attribute may be a patent claim-identifying attribute, such as a patent claim number. A basic method for finding patent-relevant documents published on the Internet in accordance with the present invention

		comprises the steps of: inputting a patent-identifying attribute on an end-user station; identifying patent data from the patent-identifying attribute; identifying Internet publication data from the patent data; and outputting the Internet publication data on the end-user station." ¶ 4. *See also* ¶ 15; Fig. 4. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 6,546,386 at Abstract, 1:47-54, 1:64-67, 2:10-14, 2:26-29, 2:42-45, 2:57-61, 3:29-35, 3:37-45, 3:55-57, 4:18-20; Fig. 1 PCT Application Pub. No. WO 01/44992 at 2:16-19, 5:8-12 U.S. Patent No. 6,236,768 at 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein at p. 406 U.S. Patent No. 6,122,647 at 2:42-51, 3:20-23 U.S. Patent No. 6,473,752 at Abstract; Fig. 2 U.S. Patent No. 6,606,644 at 5:3-8, 5:30-39 U.S. Patent No. 6,829,780 at Abstract U.S. Patent No. 7,225,142 at 6:60-67 U.S. Patent No. 7,451,099 at 2:42-49, 44:8-19 Pretschner at p. 1.
	defining an organized classification of	"Patent server 330 has patent database 332 and website database 334 resident thereon. Patent database

	document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	332 has entries stored thereon associating patent-identifying attributes, such as inventor names, assignee names and patent numbers, with patent classifications and patent language, such as patent claim text. Entries may include full-text patents. Website database 334 has entries stored thereon associating patent classifications with company website identifiers, such as URLs of company home pages." ¶ 14. "The patent classification may be a U.S. or international patent classification." ¶ 15. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Mase at Abstract; Fig. 1; p. 377-378 PCT Application Pub. No. WO 01/44992 at 15:8-11, 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 2:15-24, 4:20-27, 4:45-55 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19, 19:51-56; Fig. 8c Weiss at Fig. 3; pgs. 181, 184 Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56; Fig. 5 U.S. Patent No. 6,473,752 at 1:44-53, 2:12-26, 6:24-44, 6:45-60 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 5:55-6:10, 11:16-40 U.S. Patent No. 7,225,142 at 9:33-45 U.S. Patent No. 7,451,099 at 9:20-25, 13:25-55, 14:4-9, 27:14-29, 29:25-29; Figs 3, 8, 23, 25C

		Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 1 and 2.
	automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"Abstraction of Web document-identifying attributes from the patent language search result may be accomplished by any of numerous algorithms well known in the art. Abstraction may involve, for example, reduction of a full-text patent claim to keywords separated by Boolean operators, which keywords and operators may be selected taking into account the syntactic and lexico-semantic interdependency of the words (i.e,. context) of the full-text claim." ¶ 13. "Search client 314 extracts a company website identifier from the CW [Company Website] search result and abstracts Web document-identifying (WDI) attributes from the patent language portion of the PC-PL search result (435)" ¶ 15. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: PCT Application Pub. No. WO 01/44992 at 14:19 to 15:5 U.S. Patent No. 6,236,768 at 1:56-2:6, 2:15-54, 13:19-34 U.S. Patent No. 6,546,386 at 2:26-29, 2:42-45, 3:29-35 Finkelstein at pgs. 410, 408 U.S. Patent No. 6,122,647 at 6:49-55, 7:46-49, 9:27-28, 12:27-32; Figs. 4c-4d, 5, 7 U.S. Patent No. 6,473,752 at 3:5-8, 3:14-23, 5:62-6:7, 9:26-29; Figs. 2, 5, 8, 10B U.S. Patent No. 6,606,644 at 11:4-11, 12:52-63 U.S. Patent No. 6,829,780 at 5:55-6:14, 9:34-36

		U.S. Patent No. 7,225,142 at 9:24-32 U.S. Patent No. 7,451,099 at 4:56-64, 7:22-25, 27:14-29, 27:65-28:5; Figs. 16A, 16B Oracle Text White Paper, p. 8, 9 Pretschner at p. 2.
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	"A user of end-user station 310 inputs at least one patent-identifying (PI) attribute on user interface 312 (405). Search client 314 forms a patent-identifying search query using the one or more patent-identifying attributes (410). In this regard, search client 314 forms a search query targeted, when applied to patent database 332, to retrieve a patent classification/patent language search result that includes pairs of patent classifications and patent language from one or more patents relevant to the one or more patent-identifying attributes. The patent classification may be a U.S. or international patent classification. The patent-identifying search query is transmitted via network interface 316 and network 320 from end-user station 310 to patent server 330. Patent server 330 applies the patent-identifying search query to patent database 332 to generate patent classification/patent language (PC-PL) search result (415). Patent server 330 transmits the patent classification/patent language search result to end-user station 310. End-user station 310, particularly search client 314, extracts a patent classification (PC) attribute from the patent classification portion of the PC-PL search result (420) and forms a company website-identifying (CWI) search query using the patent classification attribute (425)." ¶ 15. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: Mase at p. 377-379; Fig. 1 PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 5:12-28 U.S. Patent No. 6,546,386 at 2:10-18, 2:57-61

		U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-18, 20:19-29 Weiss at pgs. 181, 182, 185, 186, 191 Finkelstein at pgs. 408, 410 U.S. Patent No. 6,122,647 at 3:25-35, 7:66-8:5; Figs. 6-8 U.S. Patent No. 6,473,752 at 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 11:16-40 U.S. Patent No. 7,225,142 at 9:33-37 U.S. Patent No. 7,451,099 at 27:14-34, 27:46-51; Fig 16A Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 2.
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	"Patent server 330 transmits the patent classification/patent language search result to end-user station 310. End-user station 310, particularly search client 314, extracts a patent classification attribute (PC) attribute from the patent classification portion of the PC-PL search result (420) and forms a company website-identifying (CWI) search query using the patent classification attribute (425). In this regard, end-user station 310 forms a search query targeted, when applied on patent server 330, to retrieve a company website search result that includes one or more company website identifiers, such as URLs of company home pages, relevant to the patent classification attribute. End-user station 310 transmits the CWI search query to patent server 330. Patent server 330 applies the CWI search query to website database 334 to generate company website (CW) search result (430). The CW search result is transmitted to end-user station 310. Search client 314 extracts a company website identifier from the CW search result and abstracts Web document-identifying (WDI) attributes from the patent language portion of the PC-PL search result (435). Search client 314 passes the company website identifier and

		WDI attributes to search agent 318 (440). Using the company website identifier and well known DNS addressing, search agent 318 contacts the appropriate one of Web hosts 340 and, using well known "Web crawler" techniques, searches the totality of full-text documents published on the associated company website for Web document language relevant to the WDI attributes (445). Upon completion of the search, search agent 318 generates a Web document (WD) search result including Web document identifiers, such as URLs, of the relevant Web documents (450). Search agent 318 passes the Web document search result to search client 314 (455). Search client 314 extracts Web document identifiers from the Web document search result (460) and outputs the Web document identifiers on user interface 312." ¶ 15. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:14-22 U.S. Patent No. 6,236,768 at 5:12-28 U.S. Patent No. 6,546,386 at 3:37-45, 3:55-57 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-29 Weiss at pgs. 181, 182, 184, 185, 186 Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18; Figs. 5-6, 8 U.S. Patent No. 6,473,752 at 4:23-29, 7:51-8:4, 10:34-38, 15:19-23; Figs. 2, 7, 10B U.S. Patent No. 6,606,644 at 11:4-17, 12:2-7 U.S. Patent No. 6,829,780 at 13:40-50, 13:64-67, 14:43-52

		U.S. Patent No. 7,225,142 at 4:20-22, 9:38-45 U.S. Patent No. 7,451,099 at 27:46-28:5, 28:30-44 Pretschner at p. 1 and 2.
2	The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	To the extent this reference does not teach this claim element, the reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.
3	The method according to claim 2, wherein the number of terms added is limited to a predefined number.	To the extent this reference does not teach this claim element, the reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.:

		U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.
5	The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	"Search client 314 forms a patent-identifying search query using the one or more patent-identifying attributes (410). In this regard, search client 314 forms a search query targeted, when applied to patent database 332, to retrieve a patent classification/patent language search result that includes pairs of patent classifications and patent language from one or more patents relevant to the one or more patent-identifying attributes. The patent classification may be a U.S. or international patent classification." ¶ 15. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15

		U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19; Fig. 8c Weiss at pgs. 180, 184; Fig 3 U.S. Patent No. 6,236,768 at 12:41-49, 11:62-12:4, 13:1-8, 13:15-19 U.S. Patent No. 6,122,647 at 5:65-6:4 U.S. Patent No. 7,451,099 at 3:1-3, 14:4-9, 14:25-36, 27:30-34; Figs. 3, 23 Finkelstein at p. 410 Oracle Text White Paper, p. 19. Pretschner at p. 2.
10	The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary	"The patent classification may be a U.S. or international patent classification" ¶ 15. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 16:5-9 Weiss at p. 185 Mase at pgs. 377, 379 U.S. Patent No. 6,236,768 at 4:45-55 U.S. Patent No. 6,122,647 at 3:37-39, 5:50-55 U.S. Patent No. 6,473,752 at 6:45-60, 12:3-11; Fig. 9

		Oracle Text White Paper, p. 19 Pretschner at p. 2.
18	An article of manufacture for use in a computer system, comprising:	"Turning now to FIG. 3, a communication system in which the present invention is operative in accordance with a second embodiment is shown. The communication system includes an end-user station (EUS) 310, such as a personal computer or workstation, having a user interface (UI) 312, a processor-implemented search client 314 and search agent 318 and a network interface (NI) 316. Search client 314 and search agent 318 are software applications. End-user station 310 has access to patent server 330 and Web hosts 340 via network 320 that may include local area networks (LANs) and wide area networks (WANs). Patent server 330 has patent database 332 and website database 334 resident thereon. Patent database 332 has entries stored thereon associating patent-identifying attributes, such as inventor names, assignee names and patent numbers, with patent classifications and patent language, such as patent claim text." ¶ 14.
	a memory;	"Turning now to FIG. 3, a communication system in which the present invention is operative in accordance with a second embodiment is shown. The communication system includes an end-user station (EUS) 310, such as a personal computer or workstation, having a user interface (UI) 312, a processor-implemented search client 314 and search agent 318 and a network interface (NI) 316. Search client 314 and search agent 318 are software applications. End-user station 310 has access to patent server 330 and Web hosts 340 via network 320 that may include local area networks (LANs) and wide area networks (WANs). Patent server 330 has patent database 332 and website database 334 resident thereon. Patent database 332 has entries stored thereon associating patent-identifying attributes, such as inventor names, assignee names and patent numbers, with patent classifications and patent language, such as patent claim text." ¶ 14.
	instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	"Turning now to FIG. 3, a communication system in which the present invention is operative in accordance with a second embodiment is shown. The communication system includes an end-user station (EUS) 310, such as a personal computer or workstation, having a user interface (UI) 312, a processor-implemented search client 314 and search agent 318 and a network interface (NI) 316. Search client 314 and search agent 318 are software applications. End-user station 310 has access to

		patent server 330 and Web hosts 340 via network 320 that may include local area networks (LANs) and wide area networks (WANs). Patent server 330 has patent database 332 and website database 334 resident thereon. Patent database 332 has entries stored thereon associating patent-identifying attributes, such as inventor names, assignee names and patent numbers, with patent classifications and patent language, such as patent claim text." ¶ 14. See Chart for Claim 1 (above).
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	See Chart for Claim 1 (above).
	automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	See Chart for Claim 1 (above).
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).
19	The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context	See Chart for Claim 2 (above).

	information surrounding the set of entities in the selected document content.	

Appendix J

By identifying in this chart where a limitation of an asserted claim is found in a reference, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case.

	Claims of the '979 Patent	**PCT Application Pub. No. WO 01/44992**
1	A method for automatically generating a query from selected document content, comprising:	"The client 12 makes a call to the match server 14. The call from client 12 to match server 14 includes at least a contextual query, i.e., a document or a portion of the document." 5:8-10. "Another object of the present invention is to provide a system and method which automatically and contextually matches products, advertisements or other content (hereinafter referred to as 'offers') to the content on a web page that a user has selected in real-time" 2:16-19. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 1, 2, 4, 15; Fig. 4 U.S. Patent No. 6,546,386 at Abstract, 1:47-54, 1:64-67, 2:10-14, 2:26-29, 2:42-45, 2:57-61, 3:29-35, 3:37-45, 3:55-57, 4:18-20; Fig. 1 U.S. Patent No. 6,236,768 at 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein at p. 406 U.S. Patent No. 6,122,647 at 2:42-51, 3:20-23 U.S. Patent No. 6,473,752 at Abstract; Fig. 2

		U.S. Patent No. 6,606,644 at 5:3-8, 5:30-39 U.S. Patent No. 6,829,780 at Abstract U.S. Patent No. 7,225,142 at 6:60-67 U.S. Patent No. 7,451,099 at 2:42-49, 44:8-19 Pretschner at p. 1.
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	"According to one embodiment, the present invention provides a novel approach to representing textual documents as high dimensional vectors. Such an approach provides an efficient means of indexing document collections, allowing retrieval of document (querying) based on keywords, grouping related documents (categorization). Additionally, this method supports such contextual queries and document groupings" 15:8-11. "After the vector generations, the document vector, or feature vector, must be compared with database vectors, or feature vectors. A naïve approach to product matching would be to compare the document vector to the vectors for every product in the database. This process becomes burdensome as the number of products in the database grows. Fortunately, the database vectors are not smoothly distributed throughout the vector space, but rather, tend to 'clump' together, leaving vast empty spaces between the clumps, or clusters as they are commonly known. For any given cluster, there is a sphere that bounds every point in the cluster. For all the various clusters, one could compute the center point and the radius of the sphere which bounds the cluster. Then, when one wishes to find the products that match a given document vector, one need only compare the document to the products in the cluster whose bounding sphere contains the document vector (or the nearest spheres, if no sphere contains the document vector). Thus, the computation is reduced from comparing the document vector to all product vectors to simply comparing the document vector to the center vectors for the spheres, followed by comparison to the product vectors for the products in the matching spheres." 17:15-32. "In accordance with an embodiment of the present invention, the query context vector is compared to the center vector of each cluster and the clusters with center vectors closest to the query context vector are selected. It is appreciated that these cluster IDs are then used to narrow the scope of products returned by the original metadata query as issued by the client 12. In other words, this narrowing

		qualification is added to the original metadata query to form an SQL query." 18:16-22. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 14, 15 Mase at Abstract; Fig. 1; p. 377-378 U.S. Patent No. 6,236,768 at Table 2; 2:15-24, 4:20-27, 4:45-55 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19, 19:51-56; Fig. 8c Weiss at Fig. 3; pgs. 181, 184 Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56; Fig. 5 U.S. Patent No. 6,473,752 at 1:44-53, 2:12-26, 6:24-44, 6:45-60 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 5:55-6:10, 11:16-40 U.S. Patent No. 7,225,142 at 9:33-45 U.S. Patent No. 7,451,099 at 9:20-25, 13:25-55, 14:4-9, 27:14-29, 29:25-29; Figs 3, 8, 23, 25C Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 1 and 2.

	automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"T]contextual matching engine 210 is composed of three subsystems: (1) the contextual matching server 300 …The contextual matching server 300 generates a query context vector, or feature vector, using a vector generation algorithm. Generally, vector based generation algorithms have certain features in common: (1) they all characterize documents based on the presence of keywords; (2) they all associate vectors with these keywords; and (3) they all form document vectors by combining the vectors of the keywords present in the document" 14:19 – 15:5. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 15 U.S. Patent No. 6,236,768 at 1:56-2:6, 2:15-54, 13:19-34 U.S. Patent No. 6,546,386 at 2:26-29, 2:42-45, 3:29-35 Finkelstein at pgs. 410, 408 U.S. Patent No. 6,122,647 at 6:49-55, 7:46-49, 9:27-28, 12:27-32; Figs. 4c-4d, 5, 7 U.S. Patent No. 6,473,752 at 3:5-8, 3:14-23, 5:62-6:7, 9:26-29; Figs. 2, 5, 8, 10B U.S. Patent No. 6,606,644 at 11:4-11, 12:52-63 U.S. Patent No. 6,829,780 at 5:55-6:14, 9:34-36 U.S. Patent No. 7,225,142 at 9:24-32 U.S. Patent No. 7,451,099 at 4:56-64, 7:22-25, 27:14-29, 27:65-28:5; Figs. 16A, 16B Oracle Text White Paper, p. 8, 9 Pretschner at p. 2.

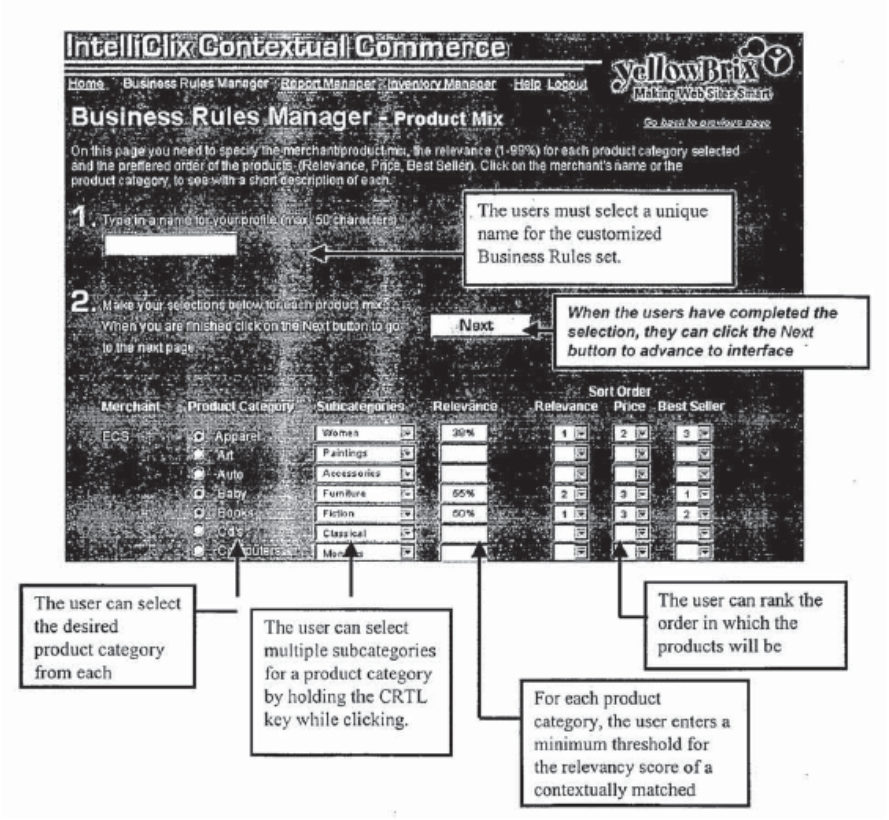
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	"After the vector generations, the document vector, or feature vector, must be compared with database vectors, or feature vectors. A naïve approach to product matching would be to compare the document vector to the vectors for every product in the database. This process becomes burdensome as the number of products in the database grows. Fortunately, the database vectors are not smoothly distributed throughout the vector space, but rather, tend to 'clump' together, leaving vast empty spaces between the clumps, or clusters as they are commonly known. For any given cluster, there is a sphere that bounds every point in the cluster. For all the various clusters, one could compute the center point and the radius of the sphere which bounds the cluster. Then, when one wishes to find the products that match a given document vector, one need only compare the document to the products in the cluster whose bounding sphere contains the document vector (or the nearest spheres, if no sphere contains the document vector). Thus, the computation is reduced from comparing the document vector to all product vectors to simply comparing the document vector to the center vectors for the spheres, followed by comparison to the product vectors for the products in the matching spheres." 17:15-32. "In accordance with an embodiment of the present invention, the query context vector is compared to the center vector of each cluster and the clusters with center vectors closest to the query context vector are selected. It is appreciated that these cluster IDs are then used to narrow the scope of products returned by the original metadata query as issued by the client 12. In other words, this narrowing qualification is added to the original metadata query to form an SQL query. Preferably, the contextual matching server 300 returns N most relevant to the client or the E-commerce applet 12, along with their associated relevance" 18:16-22. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 Mase at p. 377-379; Fig. 1 U.S. Patent No. 6,236,768 at Table 2; 5:12-28 U.S. Patent No. 6,546,386 at 2:10-18, 2:57-61

		U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-18, 20:19-29 Weiss at pgs. 181, 182, 185, 186, 191 Finkelstein at pgs. 408, 410 U.S. Patent No. 6,122,647 at 3:25-35, 7:66-8:5; Figs. 6-8 U.S. Patent No. 6,473,752 at 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 11:16-40 U.S. Patent No. 7,225,142 at 9:33-37 U.S. Patent No. 7,451,099 at 27:14-34, 27:46-51; Fig 16A Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 2.
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	"After the vector generations, the document vector, or feature vector, must be compared with database vectors, or feature vectors. A naïve approach to product matching would be to compare the document vector to the vectors for every product in the database. This process becomes burdensome as the number of products in the database grows. Fortunately, the database vectors are not smoothly distributed throughout the vector space, but rather, tend to 'clump' together, leaving vast empty spaces between the clumps, or clusters as they are commonly known. For any given cluster, there is a sphere that bounds every point in the cluster. For all the various clusters, one could compute the center point and the radius of the sphere which bounds the cluster. Then, when one wishes to find the products that match a given document vector, one need only compare the document to the products in the cluster whose bounding sphere contains the document vector (or the nearest spheres, if no sphere contains the document vector). Thus, the computation is reduced from comparing the document vector to all product vectors to simply comparing the document vector to the center vectors for the spheres,

		followed by comparison to the product vectors for the products in the matching spheres." 17:15-32. "The contextual matching server 300 then compares the query context vector to pre-determined item context vectors to narrow the search to focus on products that are most likely relevant to the selected text. In accordance with an embodiment of the present invention, the query context vector is compared to the center vector of each cluster and the clusters with center vectors closest to the query context vector are selected. It is appreciated that these cluster IDs are then used to narrow the scope of products returned by the original metadata query as issued by the client 12. In other words, this narrowing qualification is added to the original metadata query to form an SQL query." 18:14-22. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15; Fig. 4. U.S. Patent No. 6,236,768 at 5:12-28 U.S. Patent No. 6,546,386 at 3:37-45, 3:55-57 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-29 Weiss at pgs. 181, 182, 184, 185, 186 Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18; Figs. 5-6, 8 U.S. Patent No. 6,473,752 at 4:23-29, 7:51-8:4, 10:34-38, 15:19-23; Figs. 2, 7, 10B U.S. Patent No. 6,606,644 at 11:4-17, 12:2-7 U.S. Patent No. 6,829,780 at 13:40-50, 13:64-67, 14:43-52

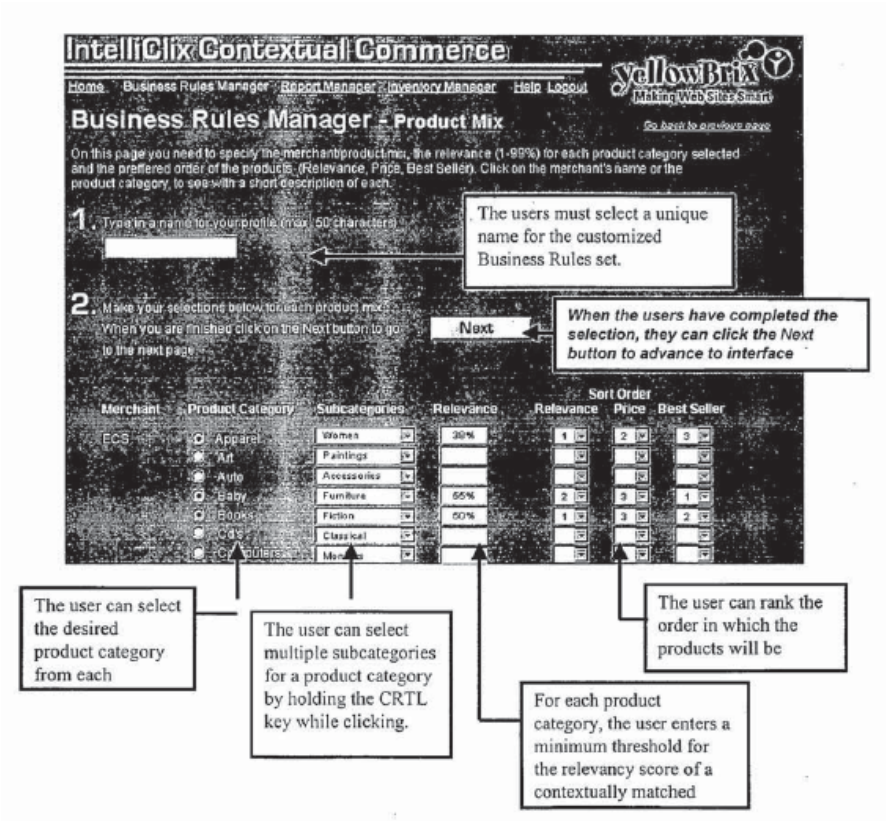
		U.S. Patent No. 7,225,142 at 4:20-22, 9:38-45 U.S. Patent No. 7,451,099 at 27:46-28:5, 28:30-44 Pretschner at p. 1 and 2.
2	The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	To the extent this reference does not teach this claim element, the reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.
3	The method according to claim 2, wherein the number of terms added is limited to a predefined number.	To the extent this reference does not teach this claim element, the reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.:

		U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.
5	The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	Fig 15:



"Preferably, the user can, as an option, set the matching process to return results based on the default matching, or based upon user - defined business rules 240. If the user selects results based on general topic matching, then the system will employ the predefined (default) business rules 240. Whereas, if the user selects results based on business rules 240, then the system will pass the text to the vector generation subsystem 310. The vector generation subsystem 310 employs user - defined business rules 240 to return results that are filtered, i.e., biased towards the business rules or instructions 240. In

		accordance with an embodiment of the present invention, the business rules 240 are applied on the front end of the contextual analysis process to define the inventory sources to match to and the minimum relevancy score acceptable for contextual matches." 18:24-19:1. "For example, a screen shot of the Business Rules Manager for customizing business rules 240 shown in Fig. 15. … Each merchant will have the applicable Product Categories in a separate column. These categories are designed to be options so there are duplicate categories among merchants. Each category has a corresponding subcategory select box. The selection box allows the users to choose more than one subcategory for each corresponding category. After all of the desired subcategories are selected, the users can enter the Relevance score. This is a percentage that determines the minimum relevancy score of the product to the context of the content." 22:3-15. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19; Fig. 8c Weiss at pgs. 180, 184; Fig 3 U.S. Patent No. 6,236,768 at 12:41-49, 11:62-12:4, 13:1-8, 13:15-19 U.S. Patent No. 6,122,647 at 5:65-6:4 U.S. Patent No. 7,451,099 at 3:1-3, 14:4-9, 14:25-36, 27:30-34; Figs. 3, 23 Finkelstein at p. 410 Oracle Text White Paper, p. 19 Pretschner at p. 2.
10	The method according to claim 1, wherein	Fig. 15:

	each class in the organized classification of document content has associated therewith a characteristic vocabulary.	 "Preferably, the user can, as an option, set the matching process to return results based on the default matching, or based upon user - defined business rules 240. If the user selects results based on general topic matching, then the system will employ the predefined (default) business rules 240. Whereas, if the user selects results based on business rules 240, then the system will pass the text to the vector generation subsystem 310. The vector generation subsystem 310 employs user - defined business rules

		240 to return results that are filtered, i.e., biased towards the business rules or instructions 240. In accordance with an embodiment of the present invention, the business rules 240 are applied on the front end of the contextual analysis process to define the inventory sources to match to and the minimum relevancy score acceptable for contextual matches." 18:24-19:1. "For example, a screen shot of the Business Rules Manager for customizing business rules 240 shown in Fig. 15. … Each merchant will have the applicable Product Categories in a separate column. These categories are designed to be options so there are duplicate categories among merchants. Each category has a corresponding subcategory select box. The selection box allows the users to choose more than one subcategory for each corresponding category. After all of the desired subcategories are selected, the users can enter the Relevance score. This is a percentage that determines the minimum relevancy score of the product to the context of the content." 22:3-15. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 U.S. Patent No. 7,225,180 at 16:5-9 Weiss at p. 185 Mase at pgs. 377, 379 U.S. Patent No. 6,236,768 at 4:45-55 U.S. Patent No. 6,122,647 at 3:37-39, 5:50-55 U.S. Patent No. 6,473,752 at 6:45-60, 12:3-11; Fig. 9 Oracle Text White Paper, p. 19 Pretschner at p. 2.

18	An article of manufacture for use in a computer system, comprising:	"The invention finds ready application in virtually all commercial communications and/or computer networks, including but not limited to world wide web (Internet), intranet, local area network (LAN), wide area network (WAN), wireless network and wired cable transmission systems." 4:27-5:1
	a memory;	"The offers to be contextually matched to that content or queries 200 are sent from the database to the intelligence engine 14, where they are then categorized and stored." 8:4-5.
	instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	"The invention finds ready application in virtually all commercial communications and/or computer networks, including but not limited to world wide web (Internet), intranet, local area network (LAN), wide area network (WAN), wireless network and wired cable transmission systems." 4:27-5:1. Implementation of the invention inherently includes instructions stored in memory. See Chart for Claim 1 (above).
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	See Chart for Claim 1 (above).
	automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	See Chart for Claim 1 (above).
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).

19	The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	See Chart for Claim 2 (above).

Appendix K

By identifying in this chart where a limitation of an asserted claim is found in a reference, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case.

	Claims of the '979 Patent	**U.S. Patent No. 6,546,386 ("Black") and Weiss et al., *HyPursuit: A Heirarchical Network Search Engine that Exploits Content-Link Hypertext Clustering*, Proc. of the 7th ACM Conference on Hypertext (March 1996), pages 180-193 ("Weiss")**
1	A method for automatically generating a query from selected document content, comprising:	Black at 1:64-67. "A system for conducting queries from any document displayed on any computer device." Black at Abstract. "[A]utomatic ... statistical and empirical analysis" of a body of selected content that may be "text, such as magazine articles, news stories or any other text" where the text can be an online article. Black at 2:43, 1:47-48. See also id. at Fig. 1, 4:18-20. "Brilliant queries require a preparation process that analyzes any text to enhance and generate a set of suggested searches based on that analysis and certain pre-set user parameters." Black at 1:50-54. "The hook is the concept, primary subject matter or main topic for a body of text. The hook is used to define a query as narrowly as possible on a particular topic for a selected information source." Black at 2:10-14. "Automatic Generation of the Hook - One embodiment of the brilliant query to enable an automatic process for generating brilliant queries for a body of text, is to determine the hook by extracting the highest frequency proper names from the text body." Black at 2:57-61. "A brilliant query requires a list of keywords that are generated by automatic ... statistical and

		empirical analysis of the body of content to be enhanced or a comparable body of content." Black at 2:42-45. "Keywords are simply a collection of words, generated automatically..., that are deemed to be indicative of the topic matter or one of the topics for a given content selection." Black at 2:26-29. "Automatic Generation of Keywords - A word frequency analysis is done on all of the text, with stopwords excluded, and the resulting words, by order of frequency are compared to a pre-selected keyword list. Those that match, based upon a desired frequency become keywords to be combined with the hook to form focused, optimal queries." Black at 3:29-35. "Once the keywords have been selected and the hook for a body of text has been determined or automatically generated, the searches are created by generating a link for every keyword extracted from the body of text and combining it with the hook in a search that results in a result set that is the logical intersection of the results generated by the hook and the keyword. Basically, each entry in the list of search results must contain both the hook and the keyword and not just one or the other." Black at 3:37-45. "After the hook and the keywords have been established the query is conducted selecting one of the hook-keyword sets." Black at 3:55-57. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 1, 2, 4, 15; Fig. 4 PCT Application Pub. No. WO 01/44992 at 2:16-19, 5:8-12 U.S. Patent No. 6,236,768 at 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein at p. 406

		U.S. Patent No. 6,122,647 at 2:42-51, 3:20-23 U.S. Patent No. 6,473,752 at Abstract; Fig. 2 U.S. Patent No. 6,606,644 at 5:3-8, 5:30-39 U.S. Patent No. 6,829,780 at Abstract U.S. Patent No. 7,225,142 at 6:60-67 U.S. Patent No. 7,451,099 at 2:42-49, 44:8-19 Pretschner at p. 1.
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	Weiss at Fig. 3. "The HyPursuit prototype is a scalable system that uses content-link hypertext clustering, based on document contents and link information, to structure the information space and to support the entire range of search activities. Content-link clustering automatically computes sets of related documents called clusters. HyPursuit admits multiple coexisting cluster hierarchies based on different principles of grouping documents, such as the Library of Congress catalog scheme and institutional structures. These hierarchies may be constructed automatically or manually" Weiss at p. 184 col. 2. "For example, documents can be clustered based on institutional boundaries or based on Library of Congress catalog subjects." Weiss at p. 184 col. 1. "Clusters also provide convenient units for the partitioning of work and resource allocation among the distributed components of the system. For example, a separate information server on a separate host may represent each individual cluster, performing operations on its local data." Weiss at p. 181, col. 1.

		See also Fig. 3. "Each content router users its abstraction functions to compute a content label that summarizes its associated cluster." Weiss at p. 184. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 14, 15 Mase at Abstract; Fig. 1; p. 377-378 PCT Application Pub. No. WO 01/44992 at 15:8-11, 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 2:15-24, 4:20-27, 4:45-55 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19, 19:51-56; Fig. 8c Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56; Fig. 5 U.S. Patent No. 6,473,752 at 1:44-53, 2:12-26, 6:24-44, 6:45-60 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 5:55-6:10, 11:16-40 U.S. Patent No. 7,225,142 at 9:33-45 U.S. Patent No. 7,451,099 at 9:20-25, 13:25-55, 14:4-9, 27:14-29, 29:25-29; Figs 3, 8, 23, 25C

		Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 1 and 2.
	automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"Keywords are simply a collection of words, generated automatically … , that are deemed to be indicative of the topic matter or one of the topics for a given content selection." Black at 2:26-29. "A brilliant query requires a list of keywords that are generated by automatic ... statistical and empirical analysis of the body of content to be enhanced or a comparable body of content." Black at 2:42-45. "Automatic Generation of Keywords - A word frequency analysis is done on all of the text, with stopwords excluded, and the resulting words, by order of frequency are compared to a pre-selected keyword list. Those that match, based upon a desired frequency become keywords to be combined with the hook to form focused, optimal queries." Black at 3:29-35. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 15 PCT Application Pub. No. WO 01/44992 at 14:19 to 15:5 U.S. Patent No. 6,236,768 at 1:56-2:6, 2:15-54, 13:19-34 Finkelstein at pgs. 410, 408 U.S. Patent No. 6,122,647 at 6:49-55, 7:46-49, 9:27-28, 12:27-32; Figs. 4c-4d, 5, 7 U.S. Patent No. 6,473,752 at 3:5-8, 3:14-23, 5:62-6:7, 9:26-29; Figs. 2, 5, 8, 10B U.S. Patent No. 6,606,644 at 11:4-11, 12:52-63

		U.S. Patent No. 6,829,780 at 5:55-6:14, 9:34-36 U.S. Patent No. 7,225,142 at 9:24-32 U.S. Patent No. 7,451,099 at 4:56-64, 7:22-25, 27:14-29, 27:65-28:5; Figs. 16A, 16B Oracle Text White Paper, p. 8, 9 Pretschner at p. 2.
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	"The hook is the concept, primary subject matter or main topic for a body of text. The hook is used to define a query as narrowly as possible on a particular topic for a selected information source. To determine a "hook", a content layer must exist for which a context can be determined. There must be a perceivable structure to the information source and each content entry must have an associated context or place or places within the structure of the information source." Black at 2:10-18. "Automatic Generation of the Hook - One embodiment of the brilliant query to enable an automatic process for generating brilliant queries for a body of text, is to determine the hook by extracting the highest frequency proper names from the text body." Black at 2:57-61. "To support a variety of query processing operations, HyPursuit uses query routing to identify relevant clusters, forward queries to the information servers for those clusters, and merge the results." Weiss at p. 182 col. 1. "HyPursuit uses query routing to support the search operations. Query routing uses the content labels stored in the content router to determine which of the child servers are likely to contain documents related to the user query. The query is then forwarded to these servers, and the results from each server are merged into a single result set. Documents returned by more than one child server are displayed only once." Weiss at p. 186 col. 2. "The abstraction function for query routing, on the other hand, computes a manageable set of terms that are used for identifying portions of the information space relevant to particular queries." Weiss at p.

		185 col. 1. "To support operations like query processing in a scalable way, HyPursuit uses manageable summaries of cluster contents, called content labels, to approximate complete knowledge of the information space." Weiss at p. 181 col. 2. "To support scalable query processing, HyPursuit uses manageable summaries of cluster contents, called content labels, to approximate complete knowledge of the information space." Weiss at p. 191 col. 2. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 Mase at p. 377-379; Fig. 1 PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 5:12-28 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-18, 20:19-29 Finkelstein at pgs. 408, 410 U.S. Patent No. 6,122,647 at 3:25-35, 7:66-8:5; Figs. 6-8 U.S. Patent No. 6,473,752 at 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 11:16-40 U.S. Patent No. 7,225,142 at 9:33-37

		U.S. Patent No. 7,451,099 at 27:14-34, 27:46-51; Fig 16A Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 2.
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	"Once the keywords have been selected and the hook for a body of text has been determined or automatically generated, the searches are created by generating a link for every keyword extracted from the body of text and combining it with the hook in a search that results in a result set that is the logical intersection of the results generated by the hook and the keyword. Basically, each entry in the list of search results must contain both the hook and the keyword and not just one or the other." Black at 3:37-45. "After the hook and the keywords have been established the query is conducted selecting one of the hook-keyword sets." Black at 3:55-57. "To support a variety of query processing operations, HyPursuit uses query routing to identify relevant clusters, forward queries to the information servers for those clusters, and merge the results." Weiss at p. 182 col. 1. "HyPursuit uses query routing to support the search operations. Query routing uses the content labels stored in the content router to determine which of the child servers are likely to contain documents related to the user query. The query is then forwarded to these servers, and the results from each server are merged into a single result set. Documents returned by more than one child server are displayed only once." Weiss at p. 186 col. 2. "The abstraction function for query routing, on the other hand, computes a manageable set of terms that are used for identifying portions of the information space relevant to particular queries." Weiss at p. 185 col. 1. "To support operations like query processing in a scalable way, HyPursuit uses manageable summaries

		of cluster contents, called content labels, to approximate complete knowledge of the information space" Weiss at p. 184 col. 1. "To support scalable query processing, HyPursuit uses manageable summaries of cluster contents, called content labels, to approximate complete knowledge of the information space." Weiss at p. 181 col. 2. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15; Fig. 4. PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:14-22 U.S. Patent No. 6,236,768 at 5:12-28 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-29 Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18; Figs. 5-6, 8 U.S. Patent No. 6,473,752 at 4:23-29, 7:51-8:4, 10:34-38, 15:19-23; Figs. 2, 7, 10B U.S. Patent No. 6,606,644 at 11:4-17, 12:2-7 U.S. Patent No. 6,829,780 at 13:40-50, 13:64-67, 14:43-52 U.S. Patent No. 7,225,142 at 4:20-22, 9:38-45 U.S. Patent No. 7,451,099 at 27:46-28:5, 28:30-44

		Pretschner at p. 1 and 2.
2	The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	"For example, an article on the SR-71 Blackbird Airplane might have the following brilliant queries: 1. Search for more information on BLACKBIRD and AVIATION 2. Search for more information on BLACKBIRD and ELINT 3. Search for more information on BLACKBIRD and RECONNAISSANCE 4. Search for more information on BLACKBIRD and TRANSPORT The book is BLACKBIRD and the keywords are AVIATION, ELINT (electronic intelligence) RECONNAISSANCE and TRANSPORT." Black at 1:64-2:9. "Query Refinement HyPursuit uses term information about sub-clusters to dynamically compute recall- and precision-enhancing terms related to a user query. Figure 6 shows the interface of our system after an interaction with the search facilities to produce a result set and a subsequent query refinement operation. The region titled suggested terms in Figure 6 contains three scrollable lists of terms. A content router suggests query refinement terms using the subclusters in the content labels of its child servers. Collocated terms are the highest weighted terms from the subclusters that match the query. HyPursuit's term weights approximate conditional probabilities of term collocation. Term collocation in sub-clusters approximates term collocation in documents." Weiss at p. 186 col. 2. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49

		U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.
5	The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	Weiss Fig. 3:

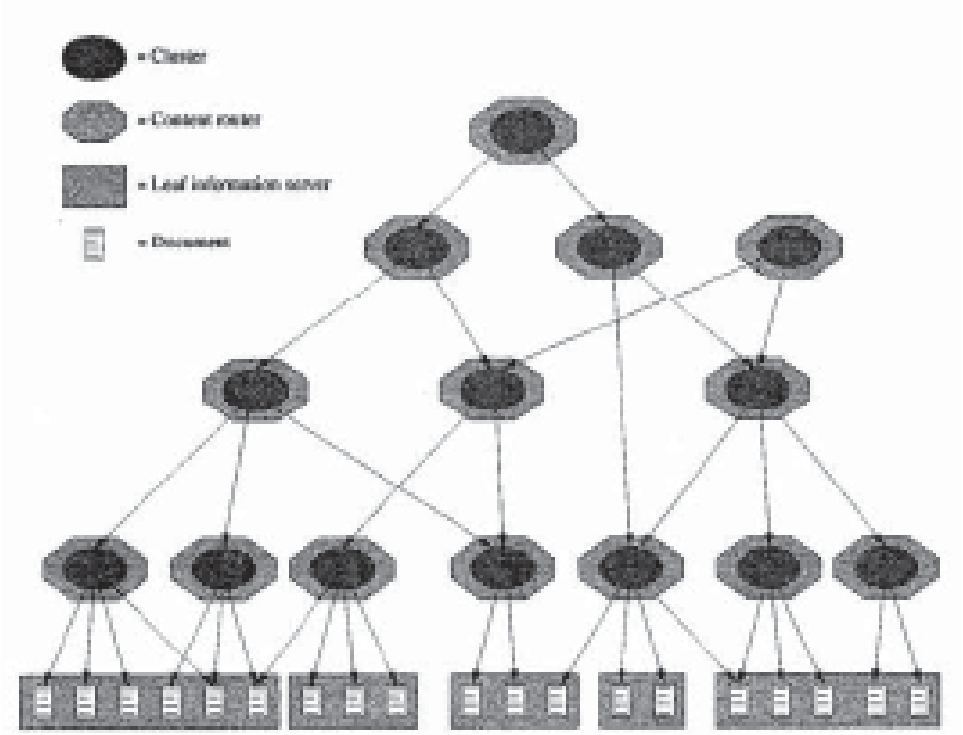


Figure 3: Content Routing as Cluster Hierarchies

"The HyPursuit prototype is a scalable system that uses content-link hypertext clustering, based on document contents and link information, to structure the information space and to support the entire range of search activities. Content-link clustering automatically computes sets of related documents called clusters. HyPursuit admits multiple coexisting cluster hierarchies based on different principles of grouping documents, such as the Library of Congress catalog scheme and institutional structures.

		These hierarchies may be constructed automatically or manually." Weiss at p. 180 col. 2. "For example, documents can be clustered based on institutional boundaries or based on Library of Congress catalog subjects." Weiss at p. 184 col. 1. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19; Fig. 8c U.S. Patent No. 6,236,768 at 12:41-49, 11:62-12:4, 13:1-8, 13:15-19 U.S. Patent No. 6,122,647 at 5:65-6:4 U.S. Patent No. 7,451,099 at 3:1-3, 14:4-9, 14:25-36, 27:30-34; Figs. 3, 23 Finkelstein at p. 410 Oracle Text White Paper, p. 19 Pretschner at p. 2.
10	The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary.	"The abstraction function for query routing, on the other hand, computes a manageable set of terms that are used for identifying portions of the information space relevant to particular queries. The abstraction function uses term and term frequency information in the children's content labels to compute term weights. The abstraction function then selects the most heavily weighted terms for generating the content router's content label. The abstraction function may also choose to add additional terms that characterize the information space but were not among the terms transmitted upon the hierarchy. For example, the abstraction function could add a term describing a poetry cluster as

		literature even though none of the poems mention literature explicitly." Weiss at p. 185 col. 1. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 16:5-9 Mase at pgs. 377, 379 U.S. Patent No. 6,236,768 at 4:45-55 U.S. Patent No. 6,122,647 at 3:37-39, 5:50-55 U.S. Patent No. 6,473,752 at 6:45-60, 12:3-11; Fig. 9 Oracle Text White Paper, p. 19 Pretschner at p. 2.
18	An article of manufacture for use in a computer system, comprising:	Black at 4:58-67, 1:50-57, 4:30-43. Weiss at p. 189 col. 2.
	a memory;	Black at 1:50-57, 4:58-67. Weiss at p. 189 col. 2.

	instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	Weiss at p. 189 col. 2. See Chart for Claim 1 (above).
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	See Chart for Claim 1 (above).
	automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	See Chart for Claim 1 (above).
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).
19	The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	Black at 1:50-57, 4:58-67. Weiss, p. 189 col. 2. See Chart for Claim 2 (above)

Appendix L

By identifying in this chart where a limitation of an asserted claim is found in a reference, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case.

	Claims of the '979 Patent	**U.S. Patent No. 6,546,386 ("Black") and U.S. Patent No. 7,225,180 ("Donaldson")**
1	A method for automatically generating a query from selected document content, comprising:	Black at 1:64-67. "A system for conducting queries from any document displayed on any computer device." Black at Abstract. "[A]utomatic ... statistical and empirical analysis" of a body of selected content that may be "text, such as magazine articles, news stories or any other text" where the text can be an online article. Black at 2:43, 1:47-48. See also id. at Fig. 1, 4:18-20. "Brilliant queries require a preparation process that analyzes any text to enhance and generate a set of suggested searches based on that analysis and certain pre-set user parameters." Black at 1:50-54. "The hook is the concept, primary subject matter or main topic for a body of text. The hook is used to define a query as narrowly as possible on a particular topic for a selected information source." Black at 2:10-14. "Automatic Generation of the Hook - One embodiment of the brilliant query to enable an automatic process for generating brilliant queries for a body of text, is to determine the hook by extracting the highest frequency proper names from the text body." Black at 2:57-61. "A brilliant query requires a list of keywords that are generated by automatic ... statistical and empirical analysis of the body of content to be enhanced or a comparable body of content." Black at

		2:42-45. "Keywords are simply a collection of words, generated automatically..., that are deemed to be indicative of the topic matter or one of the topics for a given content selection." Black at 2:26-29. "Automatic Generation of Keywords - A word frequency analysis is done on all of the text, with stopwords excluded, and the resulting words, by order of frequency are compared to a pre-selected keyword list. Those that match, based upon a desired frequency become keywords to be combined with the hook to form focused, optimal queries." Black at 3:29-35. "Once the keywords have been selected and the hook for a body of text has been determined or automatically generated, the searches are created by generating a link for every keyword extracted from the body of text and combining it with the hook in a search that results in a result set that is the logical intersection of the results generated by the hook and the keyword. Basically, each entry in the list of search results must contain both the hook and the keyword and not just one or the other." Black at 3:37-45. "After the hook and the keywords have been established the query is conducted selecting one of the hook-keyword sets." Black at 3:55-57. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 1, 2, 4, 15; Fig. 4 PCT Application Pub. No. WO 01/44992 at 2:16-19, 5:8-12 U.S. Patent No. 6,236,768 at 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein at p. 406 U.S. Patent No. 6,122,647 at 2:42-51, 3:20-23

		U.S. Patent No. 6,473,752 at Abstract; Fig. 2 U.S. Patent No. 6,606,644 at 5:3-8, 5:30-39 U.S. Patent No. 6,829,780 at Abstract U.S. Patent No. 7,225,142 at 6:60-67 U.S. Patent No. 7,451,099 at 2:42-49, 44:8-19 Pretschner at p. 1.
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	Donaldson at 15:38-46, 17:17-19, 19:51-56. "Each category may include a listing of sub-categories 865 and web sites 875 within those categories." Donaldson at 17:17-19; Fig. 8c. "For example, in one implementation, the hierarchy of category identifiers may include a hierarchy of category names, where groups of the category names are linked together in a hierarchical relationship. In this instance, names in the hierarchy represent categories, the names of which are linked together using sub-categories. The hierarchy of category identifiers also may include other related information, such as a list of web sites that are related to the category by name, description, or otherwise." Donaldson at 15:38-46. "Each electronic information store may contain content that has been classified and stored based on a specified type or types of classification criteria. For instance, the first electronic information store 992 may include content classified as non-offensive and the second electronic information store 994 may include content classified as offensive. Other types of content classification criteria may be implemented in addition to or separate from criteria based on offensive and non-offensive classifications. Other criteria that may be used, for example, include medical and non-medical, legal and non-legal, and sports and non-sports. In one implementation, the first electronic information includes contents relating to non-offensive web sites, and the second electronic information includes

		contents relating to offensive web sites." Donaldson at 19:51-65. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 14, 15 Mase at Abstract; Fig. 1; p. 377-378 PCT Application Pub. No. WO 01/44992 at 15:8-11, 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 2:15-24, 4:20-27, 4:45-55 Weiss at Fig. 3; pgs. 181, 184 Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56; Fig. 5 U.S. Patent No. 6,473,752 at 1:44-53, 2:12-26, 6:24-44, 6:45-60 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 5:55-6:10, 11:16-40 U.S. Patent No. 7,225,142 at 9:33-45 U.S. Patent No. 7,451,099 at 9:20-25, 13:25-55, 14:4-9, 27:14-29, 29:25-29; Figs 3, 8, 23, 25C Oracle Text White Paper, p. 11, 18, 19

		Pretschner at p. 1 and 2.
	automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"Keywords are simply a collection of words, generated automatically … , that are deemed to be indicative of the topic matter or one of the topics for a given content selection." Black at 2:26-29. “A brilliant query requires a list of keywords that are generated by automatic ... statistical and empirical analysis of the body of content to be enhanced or a comparable body of content." Black at 2:42-45. "Automatic Generation of Keywords - A word frequency analysis is done on all of the text, with stopwords excluded, and the resulting words, by order of frequency are compared to a pre-selected keyword list. Those that match, based upon a desired frequency become keywords to be combined with the hook to form focused, optimal queries." Black at 3:29-35. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 15 PCT Application Pub. No. WO 01/44992 at 14:19 to 15:5 U.S. Patent No. 6,236,768 at 1:56-2:6, 2:15-54, 13:19-34 Finkelstein at pgs. 410, 408 U.S. Patent No. 6,122,647 at 6:49-55, 7:46-49, 9:27-28, 12:27-32; Figs. 4c-4d, 5, 7 U.S. Patent No. 6,473,752 at 3:5-8, 3:14-23, 5:62-6:7, 9:26-29; Figs. 2, 5, 8, 10B U.S. Patent No. 6,606,644 at 11:4-11, 12:52-63 U.S. Patent No. 6,829,780 at 5:55-6:14, 9:34-36

		U.S. Patent No. 7,225,142 at 9:24-32 U.S. Patent No. 7,451,099 at 4:56-64, 7:22-25, 27:14-29, 27:65-28:5; Figs. 16A, 16B Oracle Text White Paper, p. 8, 9 Pretschner at p. 2.
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	"The hook is the concept, primary subject matter or main topic for a body of text. The hook is used to define a query as narrowly as possible on a particular topic for a selected information source. To determine a "hook", a content layer must exist for which a context can be determined. There must be a perceivable structure to the information source and each content entry must have an associated context or place or places within the structure of the information source." Black at 2:10-18. "Automatic Generation of the Hook - One embodiment of the brilliant query to enable an automatic process for generating brilliant queries for a body of text, is to determine the hook by extracting the highest frequency proper names from the text body." Black at 2:57-61. "Classifying the search term (922) generally includes classifying the received search term among one or more categories, with a first category and a second category being described and shown for illustrative purposes. If several search terms are grouped as a single string, the search terms may be collectively classified as a single string based on the grouping of search terms, or they may be classified individually based on each individual search term." Donaldson at 18:47-54; Figs. 9b and 9c. “Comparing the search terms (step 924) generally includes comparing the search term to first electronic information within a first electronic information store when the search term is classified within the first category. By contrast, comparing the search term (step 926) generally includes comparing the search term to the second electronic information within the second electronic information store to determine whether matches exist when the search term is classified within the second category." Donaldson at 18:55-63. "The following describes an example applying the described search methods of FIG. 9b to this implementation. A user of a client system enters a search term (step 910). The search term is

		classified as either being offensive or non-offensive (step 922). If the term is classified as being non-offensive, then only the contents of the first electronic information store are searched (924) and results from the search are communicated for display to the user (step 930). In this example, the first electronic information store only contains contents that previously have been classified as non-offensive. If the search term entered by the user is classified as being offensive, the contents of either the second electronic information store or both the first and second electronic information stores are searched (step 926) and the results are communicated for display to the user (step 930)." Donaldson at 20:3-18. "The described filtering of results between offensive content and non-offensive content based on the classification of the search term may allow a web host to implement a parental type of control in determining what search results are displayed to the user. Because the offensive and non-offensive contents are stored in different electronic information stores, the ability to restrict access is enhanced. For instance, parental control can be exercised by blocking the access of a user to one or more electronic information stores. Other forms of data filtering also are enabled through this process and related techniques." Donaldson at 20:19-29. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 Mase at p. 377-379; Fig. 1 PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 5:12-28 Weiss at pgs. 181, 182, 185, 186, 191 Finkelstein at pgs. 408, 410 U.S. Patent No. 6,122,647 at 3:25-35, 7:66-8:5; Figs. 6-8

		U.S. Patent No. 6,473,752 at 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 11:16-40 U.S. Patent No. 7,225,142 at 9:33-37 U.S. Patent No. 7,451,099 at 27:14-34, 27:46-51; Fig 16A Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 2.
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	"Once the keywords have been selected and the hook for a body of text has been determined or automatically generated, the searches are created by generating a link for every keyword extracted from the body of text and combining it with the hook in a search that results in a result set that is the logical intersection of the results generated by the hook and the keyword. Basically, each entry in the list of search results must contain both the hook and the keyword and not just one or the other." Black at 3:37-45. "After the hook and the keywords have been established the query is conducted selecting one of the hook-keyword sets." Black at 3:55-57. "Classifying the search term (922) generally includes classifying the received search term among one or more categories, with a first category and a second category being described and shown for illustrative purposes. If several search terms are grouped as a single string, the search terms may be collectively classified as a single string based on the grouping of search terms, or they may be classified individually based on each individual search term." Donaldson at 18:47-54. "Comparing the search terms (step 924) generally includes comparing the search term to first electronic information within a first electronic information store when the search term is classified within the first category. By contrast, comparing the search term (step 926) generally includes comparing the search term to the second electronic information within the second electronic information store to determine

		whether matches exist when the search term is classified within the second category." Donaldson at 18:55-63; Figs. 9b and 9c. "The following describes an example applying the described search methods of FIG. 9b to this implementation. A user of a client system enters a search term (step 910). The search term is classified as either being offensive or non-offensive (step 922). If the term is classified as being non-offensive, then only the contents of the first electronic information store are searched (924) and results from the search are communicated for display to the user (step 930). In this example, the first electronic information store only contains contents that previously have been classified as non-offensive. If the search term entered by the user is classified as being offensive, the contents of either the second electronic information store or both the first and second electronic information stores are searched (step 926) and the results are communicated for display to the user (step 930)." Donaldson at 20:3-18. "The described filtering of results between offensive content and non-offensive content based on the classification of the search term may allow a web host to implement a parental type of control in determining what search results are displayed to the user. Because the offensive and non-offensive contents are stored in different electronic information stores, the ability to restrict access is enhanced. For instance, parental control can be exercised by blocking the access of a user to one or more electronic information stores. Other forms of data filtering also are enabled through this process and related techniques." Donaldson at 20:19-29. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15; Fig. 4. PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:14-22 U.S. Patent No. 6,236,768 at 5:12-28 Weiss at pgs. 181, 182, 184, 185, 186

		Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18; Figs. 5-6, 8 U.S. Patent No. 6,473,752 at 4:23-29, 7:51-8:4, 10:34-38, 15:19-23; Figs. 2, 7, 10B U.S. Patent No. 6,606,644 at 11:4-17, 12:2-7 U.S. Patent No. 6,829,780 at 13:40-50, 13:64-67, 14:43-52 U.S. Patent No. 7,225,142 at 4:20-22, 9:38-45 U.S. Patent No. 7,451,099 at 27:46-28:5, 28:30-44 Pretschner at p. 1 and 2.
5	The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	"Each category may include a listing of sub-categories 865 and web sites 875 within those categories." Donaldson at 17:17-19; Fig. 8c. "For example, in one implementation, the hierarchy of category identifiers may include a hierarchy of category names, where groups of the category names are linked together in a hierarchical relationship. In this instance, names in the hierarchy represent categories, the names of which are linked together using sub - categories. The hierarchy of category identifiers also may include other related information, such as a list of web sites that are related to the category by name, description, or otherwise." Donaldson at 15:38-46. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15

		Weiss at pgs. 180, 184; Fig 3 U.S. Patent No. 6,236,768 at 12:41-49, 11:62-12:4, 13:1-8, 13:15-19 U.S. Patent No. 6,122,647 at 5:65-6:4 U.S. Patent No. 7,451,099 at 3:1-3, 14:4-9, 14:25-36, 27:30-34; Figs. 3, 23 Finkelstein at p. 410 Oracle Text White Paper, p. 19 Pretschner at p. 2.
10	The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary.	"Comparing the search term with terms related to one or more categories to determine whether matches exist (step 830) may include using information related to categories, such as a name of a web site corresponding to a category, a description of the web site, or other related terms." Donaldson at 16:5-9. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 Weiss at p. 185 Mase at pgs. 377, 379 U.S. Patent No. 6,236,768 at 4:45-55

		U.S. Patent No. 6,122,647 at 3:37-39, 5:50-55 U.S. Patent No. 6,473,752 at 6:45-60, 12:3-11; Fig. 9 Oracle Text White Paper, p. 19 Pretschner at p. 2.
18	An article of manufacture for use in a computer system, comprising:	Black at 4:58-67, 1:50-57, 4:30-43. Donaldson at Fig. 4; 13:62-14:6, 6:53-7:10, 3:45-58.
	a memory;	Black at 1:50-57, 4:58-67. Donaldson at Fig. 4; 13:62-14:6, 6:53-7:10, 3:45-58.
	instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	Donaldson at Fig. 4; 3:45-58, 6:53-7:10, 13:62-14:6. See Chart for Claim 1 (above).
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	See Chart for Claim 1 (above).
	automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
	automatically categorizing the selected document content using the organized	See Chart for Claim 1 (above).

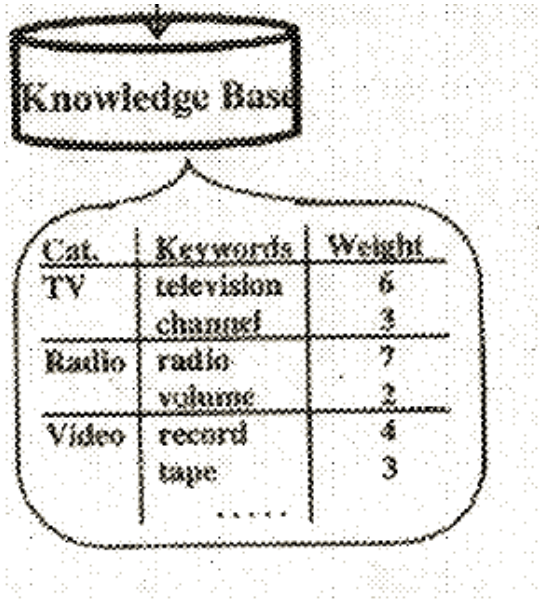
	classification of document content for assigning the selected document content a classification label from the organized classification of content; and	
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).

Appendix M

By identifying in this chart where a limitation of an asserted claim is found in a reference, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case.

	Claims of the '979 Patent	**U.S. Patent Application Publication No. 2002/0147738 ("Reader") and Mase et al., *Experimental Simulation for Automatic Patent Categorization*, Advances in Production management Systems, Kyoto, Japan, Nov. 1996, pages 377-382 ("Mase")**
1	A method for automatically generating a query from selected document content, comprising:	"Patent professionals often search for publications relevant to patents. Searches typically arise in two contexts: when looking for "prior art" publications that might invalidate a patent and when looking for publications that might disclose an infringement of a patent." Reader ¶ 1. "An ever-increasing number of publications are being published on the Internet, for example, "white papers" published on companies' public websites. Thus, the Internet has become a more and more important resource for patent professionals looking for publications relevant to patents." Reader ¶ 2. "However, patent professionals have for the most part relied on general Internet search techniques, such as applying keywords to general-purpose Internet search engines, to discover patent-relevant publications on the Internet." Reader ¶ 2. "The present invention provides a highly automated search technique for discovering patent-relevant publications on the Internet. The high level of automation may be achieved with the expedient of a search client resident on an end-user station that initiates linked searches for patent data and Internet publication data in a manner transparent to a user. From the user's perspective, a patent-identifying attribute, such as an inventor name, assignee name or patent number, input on an end-user station automatically returns Internet publication data, such as Uniform Resource Locators (URLs) of Web documents. The invention thereby allows a user to find patent-relevant publications on the Internet by merely inputting a patent-identifying attribute. A patent-identifying attribute may be a patent family-identifying attribute, such as an inventor name or assignee name. Or a patent identifying-attribute may be a single patent-identifying attribute, such as a patent number. Or a patent identifying-attribute may be a patent claim-identifying attribute, such as a patent claim number. A basic method for finding

		patent-relevant documents published on the Internet in accordance with the present invention comprises the steps of: inputting a patent-identifying attribute on an end-user station; identifying patent data from the patent-identifying attribute; identifying Internet publication data from the patent data; and outputting the Internet publication data on the end-user station." Reader ¶ 4. *See also* Reader ¶ 15; Fig. 4. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 6,546,386 at Abstract, 1:47-54, 1:64-67, 2:10-14, 2:26-29, 2:42-45, 2:57-61, 3:29-35, 3:37-45, 3:55-57, 4:18-20; Fig. 1 PCT Application Pub. No. WO 01/44992 at 2:16-19, 5:8-12 U.S. Patent No. 6,236,768 at 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 Finkelstein at p. 406 U.S. Patent No. 6,122,647 at 2:42-51, 3:20-23 U.S. Patent No. 6,473,752 at Abstract; Fig. 2 U.S. Patent No. 6,606,644 at 5:3-8, 5:30-39 U.S. Patent No. 6,829,780 at Abstract U.S. Patent No. 7,225,142 at 6:60-67 U.S. Patent No. 7,451,099 at 2:42-49, 44:8-19 Pretschner at p. 1.

	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	Mase at Fig. 1 and p. 378:  "This paper describes keywords-based patent categorization using our text classification support tool called FLUTE and discusses a simulation study applied to 154,000 patents. FLUTE automatically generates a classification knowledge base from sample patent texts. Experimental simulation results show that FLUTE is powerful enough to support the patent classification work of indexing experts." Mase at Abstract. "(1) Automatic knowledge base initialization FLUTE is able to obtain word statistics from electronic documents. It first derives keyword candidates for each document. Then, it removes stop-words which are obviously not keywords ('thing', 'is', etc.) and common words appearing over every category ('invention' and 'patent' appear in all patent documents). Next, it identifies keywords for each category

		by applying weights. Finally, it generates classification rules, which include a certain factor. Since the KB structure is simple, a person can verify and modify the knowledge base." Mase at p. 378, col. 1. "At present most patents are applied electronically. If an intelligent system could read a patent, recognize its purpose, and finally categorize it, the patent management process could be restructured. To classify patents into appropriate categories, the system would have to have a powerful knowledge base, which is difficult to build. We have been researching automatic patent categorization as a submitted research from Industrial Property Cooperation Center. This paper presents our text classification support tool called FLUTE (Mase, et al. 1996a), the customization of FLUTE for patent categorization and its evaluation using 154,000 patents" Mase at p. 377 col. 2. "Obtaining patents is important for any organization that needs to maintain intellectual property rights. In order to manage the review, search, and citation of patents at the patent office efficiently, patent applications should be appropriately indexed as soon as possible." Mase at p. 377 col. 1. "(2) Classification certainty FLUTE presents three kinds of classification results according to values corresponding to the degree of confidence (certainty). If one category's certainty is high, FLUTE presents a unique solution. If the certainties of more than two categories are high, FLUTE offers alternatives. Otherwise FLUTE offers no solution and asks an expert to intervene." Mase at p. 378. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: PCT Application Pub. No. WO 01/44992 at 15:8-11, 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 2:15-24, 4:20-27, 4:45-55 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19, 19:51-56; Fig. 8c Weiss at Fig. 3; pgs. 181, 184

		Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56; Fig. 5 U.S. Patent No. 6,473,752 at 1:44-53, 2:12-26, 6:24-44, 6:45-60 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 5:55-6:10, 11:16-40 U.S. Patent No. 7,225,142 at 9:33-45 U.S. Patent No. 7,451,099 at 9:20-25, 13:25-55, 14:4-9, 27:14-29, 29:25-29; Figs 3, 8, 23, 25C Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 1 and 2.
	automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"Abstraction of Web document-identifying attributes from the patent language search result may be accomplished by any of numerous algorithms well known in the art. Abstraction may involve, for example, reduction of a full-text patent claim to keywords separated by Boolean operators, which keywords and operators may be selected taking into account the syntactic and lexico-semantic interdependency of the words (i.e,. context) of the full-text claim." Reader ¶ 13. "Search client 314 extracts a company website identifier from the CW [Company Website] search result and abstracts Web document-identifying (WDI) attributes from the patent language portion of the PC-PL search result (435)" Reader ¶ 15. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: PCT Application Pub. No. WO 01/44992 at 14:19 to 15:5 U.S. Patent No. 6,236,768 at 1:56-2:6, 2:15-54, 13:19-34

		U.S. Patent No. 6,546,386 at 2:26-29, 2:42-45, 3:29-35 Finkelstein at pgs. 410, 408 U.S. Patent No. 6,122,647 at 6:49-55, 7:46-49, 9:27-28, 12:27-32; Figs. 4c-4d, 5, 7 U.S. Patent No. 6,473,752 at 3:5-8, 3:14-23, 5:62-6:7, 9:26-29; Figs. 2, 5, 8, 10B U.S. Patent No. 6,606,644 at 11:4-11, 12:52-63 U.S. Patent No. 6,829,780 at 5:55-6:14, 9:34-36 U.S. Patent No. 7,225,142 at 9:24-32 U.S. Patent No. 7,451,099 at 4:56-64, 7:22-25, 27:14-29, 27:65-28:5; Figs. 16A, 16B Oracle Text White Paper, p. 8, 9 Pretschner at p. 2.
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	"Obtaining patents is important for any organization that needs to maintain intellectual property rights. In order to manage the review, search, and citation of patents at the patent office efficiently, patent applications should be appropriately indexed as soon as possible. However, there are problems, as follows: - There are approximately 3,000 patent categories. This makes it impossible for any indexing expert to perform categorization work over every technical field. – In Japan, there were 370,000 patent applications in 1994. Despite such a large number, it is difficult to increase the number of indexing experts. – Most patents include over 5,000 words. In Japan, patent categorization requires experts to read all documents, which is time consuming work.

		– To maintain an applicant's proprietary information, only persons entrusted with confidentiality can be assigned to indexing work. At present, most patents are applied electronically. If an intelligent system could read a patent, recognize its purpose, and finally categorize it, the patent management process could be restructured. To classify patents into appropriate categories, the system would have to have a powerful knowledge base, which is difficult to build." Mase at p. 377. "We have been researching automatic patent categorization as a submitted research from Industrial Property Cooperation Center. This paper presents our text classification support tool called FLUTE (Mase, et al. 1996a), the customization of FLUTE for patent categorization, and its evaluation using 154,000 patents." Mase at p. 377 col. 2. Mase at Fig 1:

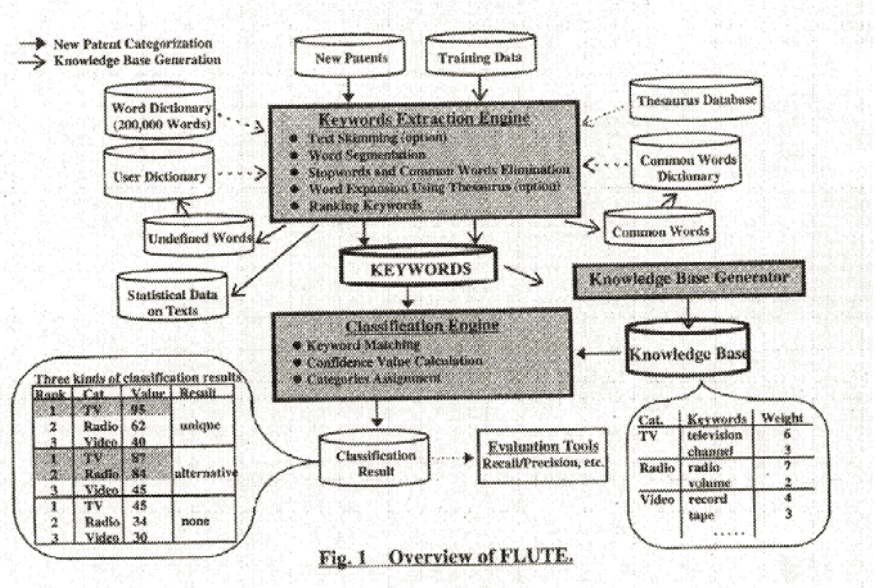
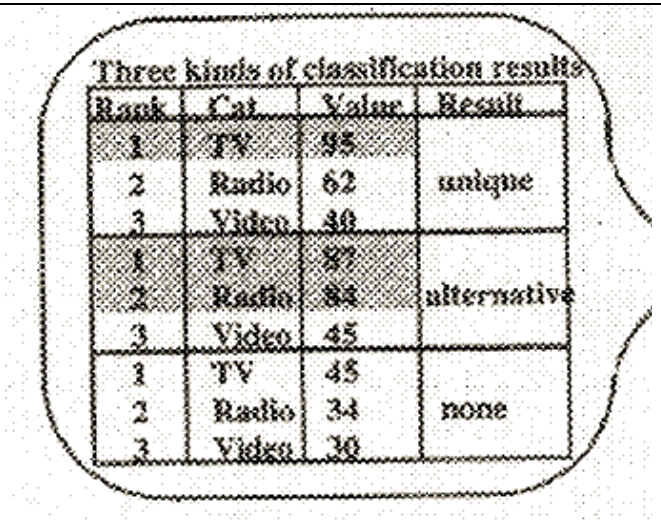


Fig. 1 Overview of FLUTE.

"The classification engine calculates the similarity between inputted text and each category by comparing the keywords extracted from the inputted text with those stored in the KB." Mase at p. 379, col. 1.

"(2) *Classification certainty*
FLUTE presents three kinds of classification results according to values corresponding to the degree of confidence (certainty). If one category's certainty is high, FLUTE presents a unique solution. If the certainties of more than two categories are high, FLUTE offers alternatives. Otherwise FLUTE offers no solution and asks an expert to intervene." Mase at p. 378.

Mase at Fig. 1:



Three kinds of classification results

Rank	Cat.	Value	Result
1	TV	95	
2	Radio	62	unique
3	Video	40	
1	TV	87	
2	Radio	83	alternative
3	Video	45	
1	TV	45	
2	Radio	34	none
3	Video	30	

To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.:

PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:16-22

U.S. Patent No. 6,236,768 at Table 2; 5:12-28

U.S. Patent No. 6,546,386 at 2:10-18, 2:57-61

U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-18, 20:19-29

Weiss at pgs. 181, 182, 185, 186, 191

Finkelstein at pgs. 408, 410

		U.S. Patent No. 6,122,647 at 3:25-35, 7:66-8:5; Figs. 6-8 U.S. Patent No. 6,473,752 at 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 11:16-40 U.S. Patent No. 7,225,142 at 9:33-37 U.S. Patent No. 7,451,099 at 27:14-34, 27:46-51; Fig 16A Oracle Text White Paper, p. 11, 18, 19 Pretschner at p. 2.
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	"Patent server 330 transmits the patent classification/patent language search result to end-user station 310. End-user station 310, particularly search client 314, extracts a patent classification attribute (PC) attribute from the patent classification portion of the PC-PL search result (420) and forms a company website-identifying (CWI) search query using the patent classification attribute (425). In this regard, end-user station 310 forms a search query targeted, when applied on patent server 330, to retrieve a company website search result that includes one or more company website identifiers, such as URLs of company home pages, relevant to the patent classification attribute. End-user station 310 transmits the CWI search query to patent server 330. Patent server 330 applies the CWI search query to website database 334 to generate company website (CW) search result (430). The CW search result is transmitted to end-user station 310. Search client 314 extracts a company website identifier from the CW search result and abstracts Web document-identifying (WDI) attributes from the patent language portion of the PC-PL search result (435). Search client 314 passes the company website identifier and WDI attributes to search agent 318 (440). Using the company website identifier and well known DNS addressing, search agent 318 contacts the appropriate one of Web hosts 340 and, using well known "Web crawler" techniques, searches the totality of full-text documents published on the associated company website for Web document language relevant to the WDI attributes (445). Upon completion of the search, search agent 318 generates a Web document (WD) search result including Web document identifiers, such as URLs, of the relevant Web documents (450). Search agent 318 passes the

		Web document search result to search client 314 (455). Search client 314 extracts Web document identifiers from the Web document search result (460) and outputs the Web document identifiers on user interface 312." Reader ¶ 15. See also Fig. 4. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:14-22 U.S. Patent No. 6,236,768 at 5:12-28 U.S. Patent No. 6,546,386 at 3:37-45, 3:55-57 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-29 Weiss at pgs. 181, 182, 184, 185, 186 Finkelstein at p. 410 U.S. Patent No. 6,122,647 at 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18; Figs. 5-6, 8 U.S. Patent No. 6,473,752 at 4:23-29, 7:51-8:4, 10:34-38, 15:19-23; Figs. 2, 7, 10B U.S. Patent No. 6,606,644 at 11:4-17, 12:2-7 U.S. Patent No. 6,829,780 at 13:40-50, 13:64-67, 14:43-52 U.S. Patent No. 7,225,142 at 4:20-22, 9:38-45 U.S. Patent No. 7,451,099 at 27:46-28:5, 28:30-44 Pretschner at p. 1 and 2.

5	The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	"Search client 314 forms a patent-identifying search query using the one or more patent-identifying attributes (410). In this regard, search client 314 forms a search query targeted, when applied to patent database 332, to retrieve a patent classification/patent language search result that includes pairs of patent classifications and patent language from one or more patents relevant to the one or more patent-identifying attributes. The patent classification may be a U.S. or international patent classification." Reader ¶ 15. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19; Fig. 8c Weiss at pgs. 180, 184; Fig 3 U.S. Patent No. 6,236,768 at 12:41-49, 11:62-12:4, 13:1-8, 13:15-19 U.S. Patent No. 6,122,647 at 5:65-6:4 U.S. Patent No. 7,451,099 at 3:1-3, 14:4-9, 14:25-36, 27:30-34; Figs. 3, 23 Finkelstein at p. 410 Oracle Text White Paper, p. 19 Pretschner at p. 2.
10	The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary.	"The patent classification may be a U.S. or international patent classification." Reader ¶ 15. "FLUTE adopts a keywords-based classification approach, and is applicable to system building satisfying the following premises: • Categories are pre-defined and exclusive of each other. • Each category has characteristic keywords." Mase at p. 377 col. 2.

		"The classification engine calculates the similarity between inputted text and each category by comparing the keywords extracted from the inputted text with those stored in the KB." Mase at p. 379, col. 1. To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 16:5-9 Weiss at p. 185 U.S. Patent No. 6,236,768 at 4:45-55 U.S. Patent No. 6,122,647 at 3:37-39, 5:50-55 U.S. Patent No. 6,473,752 at 6:45-60, 12:3-11; Fig. 9 Oracle Text White Paper, p. 19 Pretschner at p. 2.
18	An article of manufacture for use in a computer system, comprising:	"Turning now to FIG. 3, a communication system in which the present invention is operative in accordance with a second embodiment is shown. The communication system includes an end-user station (EUS) 310, such as a personal computer or workstation, having a user interface (UI) 312, a processor-implemented search client 314 and search agent 318 and a network interface (NI) 316. Search client 314 and search agent 318 are software applications. End-user station 310 has access to patent server 330 and Web hosts 340 via network 320 that may include local area networks (LANs) and wide area networks (WANs). Patent server 330 has patent database 332 and website database 334 resident thereon. Patent database 332 has entries stored thereon associating patent-identifying attributes, such as inventor names, assignee names and patent numbers, with patent classifications and patent language, such as patent claim text." ¶ 14.

		Mase at p. 382, col. 1.
	a memory;	"Turning now to FIG. 3, a communication system in which the present invention is operative in accordance with a second embodiment is shown. The communication system includes an end-user station (EUS) 310, such as a personal computer or workstation, having a user interface (UI) 312, a processor-implemented search client 314 and search agent 318 and a network interface (NI) 316. Search client 314 and search agent 318 are software applications. End-user station 310 has access to patent server 330 and Web hosts 340 via network 320 that may include local area networks (LANs) and wide area networks (WANs). Patent server 330 has patent database 332 and website database 334 resident thereon. Patent database 332 has entries stored thereon associating patent-identifying attributes, such as inventor names, assignee names and patent numbers, with patent classifications and patent language, such as patent claim text." ¶ 14. Mase at p. 382, col. 1.
	instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	"Turning now to FIG. 3, a communication system in which the present invention is operative in accordance with a second embodiment is shown. The communication system includes an end-user station (EUS) 310, such as a personal computer or workstation, having a user interface (UI) 312, a processor-implemented search client 314 and search agent 318 and a network interface (NI) 316. Search client 314 and search agent 318 are software applications. End-user station 310 has access to patent server 330 and Web hosts 340 via network 320 that may include local area networks (LANs) and wide area networks (WANs). Patent server 330 has patent database 332 and website database 334 resident thereon. Patent database 332 has entries stored thereon associating patent-identifying attributes, such as inventor names, assignee names and patent numbers, with patent classifications and patent language, such as patent claim text." ¶ 14. Mase at p. 382, col. 1. See Chart for Claim 1 (above).
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification	See Chart for Claim 1 (above).

	label; each classification label corresponding to a category of information in an information retrieval system;	
	automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;	See Chart for Claim 1 (above).
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	See Chart for Claim 1 (above).
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).

Appendix N

By identifying in this chart where a limitation of an asserted claim is found in a reference, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case.

	Claims of the '979 Patent	**PCT Application Pub. No. WO 01/44992 ("Wieser") and U.S. Patent No. 7,089,236 ("Stibel")**
2	The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	"(the client 12 generates a query composed of all or a portion of a document (e.g., a web page) and sends the query to a match server 14)." Wieser at 5:8-12. "Another object of the present invention is to provide a system and method which automatically and contextually matches products, advertisements or other content (hereinafter referred to as 'offers') to the content on a web page that a user has selected in real-time." Wieser at 2:16-19. Stibel, Table 2: TABLE 2 Search Terms Generated by the Relational Knowledgebase for User Query "java" java (required) coffee (meaning) cafe au lait cafe noir demitasse decaffeinated coffee decaf espresso capuccino coffee capuccino iced coffee ice coffee mocha Turkish coffee cafe royale beverage coffee royal Irish coffee

		"The systems and methods described herein include systems that, inter alia, operate as a front end to a database search engine or engines, and act to process a user query to generate a new search request that will more effectively retrieve information from the database that is relevant to the query of the user. To this end, in one embodiment the systems can be realized as computer programs that act as front ends to databases. The front ends may include a user interface that is presented to a user and which may prompt the user to enter one or more key phrases that are representative of a user search request. The user interface may collect the key phrases provided by the user and may analyze these key phrases to identify at least one meaning that may be associated with this user search request. The systems may then process the user search request and the identified meaning to generate an expanded search request that may be represented as a compound search string, such as a Boolean search string, or other logical string. This compound search string may then be processed to create one or more expanded user queries that may be presented to a search engine to collect from a search engine information that is relevant to the interest of the user." Stibel 2:42-63. "Referring to FIG. 1 it may be seen that the query engine 14 may employ the knowledge base 16 to identify meanings that may be associated with the user query provided by the user interface 12. To this end, the query engine may employ words and phrases from the user query to query the knowledge base 15 and collect therefrom one or more meanings that may be associated with the user query." Stibel at 10:30-36. "In one particular embodiment, the knowledge base 16 comprises two types of entities: Senses and Words. The knowledge base 16 can store a Sense as a data structure that has associated with it five items. … As shown in FIG. 3, these items can include a generalization member that is representative of the Sense that is the closest term that represents a generalization for the associated Sense; a specific meaning that can be an optional string representative of the meaning to be displayed to the user for the Sense; related terms each being representative of a word form or Sense or Word meaning that is likely to appear on a web page when a user is querying for information associated with this Sense, Sense words which may be representative of a list of words from the knowledge base 16 that have this Sense as one of their meanings; and a flag member that may contain flags associated with this Sense." Stibel at 9:21-43. "In either embodiment, after the user has selected a meaning to be associated with the user query the Query engine 14 may then build an expanded search query by employing the related terms, depicted in FIG. 3, to amend the user query to one more suited for identifying documents associated with the

		interest of the client. For example, if the user enters "java" as the user query and selects "coffee" as its meaning, the knowledge base 16 will retrieve a number of related words (e.g., mocha, espresso) and append these words, along with the chosen meaning, to the original user query." Stibel at 11:56-66. "These related words provide additional keyphrases that may be added, such as by Boolean logic operators or by other logical operators, to the keyphrases of the user query. For example, the original user query "java," can now be expanded to java+coffee+espresso(W1)+beverage(W2)+ (NOT)programming. As expanded, the query now includes terms that are selected to increase the likelihood that an Internet search engine will return a meaningful hit list." Stibel at 12:13-21. For elements found in Claim 1, see, e.g., chart for Wieser (Appendix J). To the extent these references do not teach this claim element, these references in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.

19	The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	See Chart for Claim 2 (above).

Appendix O

By identifying in this chart where a limitation of an asserted claim is found in a reference, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case.

	Claims of the '979 Patent	**Syskill & Webert: Identifying interesting web sites" (AAAI-96 Proceedings, Copyright 1996) ("Syskill")**
3	The method according to claim 2, wherein the number of terms added is limited to a predefined number.	"Since LYCOS cannot accept very long queries, we use the 7 most discriminating words that are found in a higher proportion of hot pages than all pages and the 7 most commonly occurring words as a query." Syskill at p. 56 at col 2:3-6. For elements found in Claims 1 and 2, see, e.g., Appendices A-N and P. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5

		U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.

Appendix P

By identifying in this chart where a limitation of an asserted claim is found in a reference, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case.

	Claims of the '979 Patent	**U.S. Patent No. 6,546,386 ("Black"), U.S. Patent No. 7,225,180 ("Donaldson") and U.S. Patent No. 7,089,236 ("Stibel")**
2	The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	"Once the keywords have been selected and the hook for a body of text has been determined or automatically generated, the searches are created by generating a link for every keyword extracted from the body of text and combining it with the hook in a search that results in a result set that is the logical intersection of the results generated by the hook and the keyword. Basically, each entry in the list of search results must contain both the hook and the keyword and not just one or the other." Black at 3:37-45. "After the hook and the keywords have been established the query is conducted selecting one of the hook-keyword sets." Black at 3:55-57. "Classifying the search term (922) generally includes classifying the received search term among one or more categories, with a first category and a second category being described and shown for illustrative purposes. If several search terms are grouped as a single string, the search terms may be collectively classified as a single string based on the grouping of search terms, or they may be classified individually based on each individual search term." Donaldson at 18:47-54. "Comparing the search terms (step 924) generally includes comparing the search term to first electronic information within a first electronic information store when the search term is classified within the first category. By contrast, comparing the search term (step 926) generally includes comparing the search term to the second electronic information within the second electronic information store to determine whether matches exist when the search term is classified within the second category." Donaldson at 18:55-63; Figs 9b and 9c. "The following describes an example applying the described search methods of FIG. 9b to this

		implementation. A user of a client system enters a search term (step 910). The search term is classified as either being offensive or non-offensive (step 922). If the term is classified as being non-offensive, then only the contents of the first electronic information store are searched (924) and results from the search are communicated for display to the user (step 930). In this example, the first electronic information store only contains contents that previously have been classified as non-offensive. If the search term entered by the user is classified as being offensive, the contents of either the second electronic information store or both the first and second electronic information stores are searched (step 926) and the results are communicated for display to the user (step 930)." Donaldson at 20:3-18. "The described filtering of results between offensive content and non-offensive content based on the classification of the search term may allow a web host to implement a parental type of control in determining what search results are displayed to the user. Because the offensive and non-offensive contents are stored in different electronic information stores, the ability to restrict access is enhanced. For instance, parental control can be exercised by blocking the access of a user to one or more electronic information stores. Other forms of data filtering also are enabled through this process and related techniques." Donaldson at 20:19-29. Stibel, Table 2: TABLE 2 Search Terms Generated by the Relational Knowledgebase for User Query "java" java (required), coffee (meaning), cafe au lait, cafe noir, demitasse, decaffeinated coffee, decaf, espresso, capuccino, coffee capuccino, iced coffee, ice coffee, mocha, Turkish coffee, cafe royale, beverage, coffee royal, Irish coffee

		"The systems and methods described herein include systems that, inter alia, operate as a front end to a database search engine or engines, and act to process a user query to generate a new search request that will more effectively retrieve information from the database that is relevant to the query of the user. To this end, in one embodiment the systems can be realized as computer programs that act as front ends to databases. The front ends may include a user interface that is presented to a user and which may prompt the user to enter one or more key phrases that are representative of a user search request. The user interface may collect the key phrases provided by the user and may analyze these key phrases to identify at least one meaning that may be associated with this user search request. The systems may then process the user search request and the identified meaning to generate an expanded search request that may be represented as a compound search string, such as a Boolean search string, or other logical string. This compound search string may then be processed to create one or more expanded user queries that may be presented to a search engine to collect from a search engine information that is relevant to the interest of the user." Stibel at 2:42-63. "Referring to FIG. 1 it may be seen that the query engine 14 may employ the knowledge base 16 to identify meanings that may be associated with the user query provided by the user interface 12. To this end, the query engine may employ words and phrases from the user query to query the knowledge base 15 and collect therefrom one or more meanings that may be associated with the user query." Stibel at 10:30-36. "In one particular embodiment, the knowledge base 16 comprises two types of entities: Senses and Words. The knowledge base 16 can store a Sense as a data structure that has associated with it five items. … As shown in FIG. 3, these items can include a generalization member that is representative of the Sense that is the closest term that represents a generalization for the associated Sense; a specific meaning that can be an optional string representative of the meaning to be displayed to the user for the Sense; related terms each being representative of a word form or Sense or Word meaning that is likely to appear on a web page when a user is querying for information associated with this Sense, Sense words which may be representative of a list of words from the knowledge base 16 that have this Sense as one of their meanings; and a flag member that may contain flags associated with this Sense." Stibel at 9:21-43. "In either embodiment, after the user has selected a meaning to be associated with the user query the Query engine 14 may then build an expanded search query by employing the related terms, depicted in

		FIG. 3, to amend the user query to one more suited for identifying documents associated with the interest of the client. For example, if the user enters "java" as the user query and selects "coffee" as its meaning, the knowledge base 16 will retrieve a number of related words (e.g., mocha, espresso) and append these words, along with the chosen meaning, to the original user query." Stibel at 11:56-66. "These related words provide additional keyphrases that may be added, such as by Boolean logic operators or by other logical operators, to the keyphrases of the user query. For example, the original user query "java," can now be expanded to java+coffee+espresso(W1)+beverage(W2)+ (NOT)programming. As expanded, the query now includes terms that are selected to increase the likelihood that an Internet search engine will return a meaningful hit list." Stibel at 12:13-21. For elements found in Claim 1, see, e.g., chart for Black and Donaldson (Appendix L). To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410 Pretschner at p. 2.
19	The article of manufacture according to claim 18, wherein the instructions stored in	See Chart for Claim 2 (above).

	the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	

Appendix Q

By identifying in this chart where a limitation of an asserted claim is found in a reference, Yahoo! may address a range of potential claim constructions of such limitation, including constructions with which Yahoo! may disagree, and Yahoo! does not thereby concede or admit that the limitation is found in that reference under all possible constructions of the limitation in question or under the construction that may ultimately be adopted by the Court for purposes of this case.

	Claims of the '979 Patent	**"Ontology Based Personalized Search" by Pretschner and Gauch (Proceedings of the 11th IEEE Int'l Conf. on Tools with Artificial Intelligence, pp. 391-98, Nov. 1999) ("Pretschner")**
1	A method for automatically generating a query from selected document content, comprising:	"Or, the system could navigate through the Web on its own and notify the user if it found a page or site of presumed interest." Page 1. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 1, 2, 4, 15; Fig. 4 U.S. Patent No. 6,546,386 at Abstract, 1:47-54, 1:64-67, 2:10-14, 2:26-29, 2:42-45, 2:57-61, 3:29-35, 3:37-45, 3:55-57, 4:18-20; Fig. 1 PCT Application Pub. No. WO 01/44992 at 2:16-19, 5:8-12 U.S. Patent No. 6,236,768 at 1:56-2:6, 10:42-51, 12:53-57, 13:1-8, 13:15-19, 13:42-47 U.S. Patent No. 6,122,647 at 2:42-51, 3:20-23 Finkelstein at p. 406 U.S. Patent No. 6,473,752 at Abstract; Fig. 2 U.S. Patent No. 6,606,644 at 5:3-8, 5:30-39

		U.S. Patent No. 6,829,780 at Abstract U.S. Patent No. 7,225,142 at 6:60-67 U.S. Patent No. 7,451,099 at 2:42-49, 44:8-19
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	"Browsing is usually done by clicking through a hierarchy of subjects until the area of interest has been reached. The corresponding node then provides the user with links to related websites." Page 1. "User interests are inferred by analyzing the web pages the user visits. For this purpose, it is necessary to determine the content, or characterize, these surfed pages. This is done by using a hierarchy of concepts, or rather ontology. This ontology is based on a publicly accessible browsing hierarchy. For this paper, the Magellan hierarchy, which is comprised of approximately 4,400 nodes, has been mirrored (magellan.excite.com). The nodes of the ontology are labelled with the names of the nodes in the browsing hierarchy. The semantics of the edges of this hierarchy is not formally specified; in most cases, they correspond to a specialization relation (super-/subconcept)." Page 2. "For each of the surfed pages a keyword vector is calculated. This page vector is compared with the keyword vectors associated with every node to calculate similarities. The nodes with the top matching vectors are assumed to be most related to the content of the surfed page. The accuracy of this text categorization algorithm was validated in [30]." Page 2. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 14, 15 Mase at Abstract; Fig. 1; p. 377-378 PCT Application Pub. No. WO 01/44992 at 15:8-11, 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 2:15-24, 4:20-27, 4:45-55

		U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19, 19:51-56; Fig. 8c U.S. Patent No. 6,122,647 at 3:37-39, 5:60-62, 6:28-30, 11:8-15, 5:49-56; Fig. 5 Weiss at Fig. 3; pgs. 181, 184 Finkelstein at p. 410 U.S. Patent No. 6,473,752 at 1:44-53, 2:12-26, 6:24-44, 6:45-60 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 5:55-6:10, 11:16-40 U.S. Patent No. 7,225,142 at 9:33-45 U.S. Patent No. 7,451,099 at 9:20-25, 13:25-55, 14:4-9, 27:14-29, 29:25-29; Figs 3, 8, 23, 25C Oracle Text White Paper, p. 11, 18, 19.
	automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;	"Documents as well as superdocuments are represented as weighted keyword vectors using the vector space model [23]. The weights are based on term frequencies and inverted document frequencies: It is assumed that multiple occurrences of a word indicate that its meaning contributes to the content of the document more than less frequent terms. However, words that occur with a very high overall frequency (i.e., in the collection of documents in question) do not discriminate between documents within this collection." Page 2. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶¶ 13, 15 PCT Application Pub. No. WO 01/44992 at 14:19 to 15:5

		U.S. Patent No. 6,236,768 at 1:56-2:6, 2:15-54, 13:19-34 U.S. Patent No. 6,546,386 at 2:26-29, 2:42-45, 3:29-35 U.S. Patent No. 6,122,647 at 6:49-55, 7:46-49, 9:27-28, 12:27-32; Figs. 4c-4d, 5, 7 Finkelstein at pgs. 410, 408 U.S. Patent No. 6,473,752 at 3:5-8, 3:14-23, 5:62-6:7, 9:26-29; Figs. 2, 5, 8, 10B U.S. Patent No. 6,606,644 at 11:4-11, 12:52-63 U.S. Patent No. 6,829,780 at 5:55-6:14, 9:34-36 U.S. Patent No. 7,225,142 at 9:24-32 U.S. Patent No. 7,451,099 at 4:56-64, 7:22-25, 27:14-29, 27:65-28:5; Figs. 16A, 16B Oracle Text White Paper, p. 8, 9.
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	"The goal of OBIWAN is to investigate a novel content-based approach to distributed information retrieval. Websites are clustered into regions. Examples for clustering criteria include but are not restricted to content, geographic location, and association with a specific company. Regions are clustered into super regions, super regions into hyper regions, etc., thus forming a hierarchy of regions. A node of this hierarchy can be browsed by simultaneously browsing its child nodes." Page 2. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 Mase at p. 377-379; Fig. 1

		PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:16-22 U.S. Patent No. 6,236,768 at Table 2; 5:12-28 U.S. Patent No. 6,546,386 at 2:10-18, 2:57-61 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-18, 20:19-29 U.S. Patent No. 6,122,647 at 3:25-35, 7:66-8:5; Figs. 6-8 Weiss at pgs. 181, 182, 185, 186, 191 Finkelstein at pgs. 408, 410 U.S. Patent No. 6,473,752 at 1:6-9, 3:5-8, 9:26-29, 9:65-10:2, 12:18-24 U.S. Patent No. 6,606,644 at 11:4-17 U.S. Patent No. 6,829,780 at 11:16-40 U.S. Patent No. 7,225,142 at 9:33-37 U.S. Patent No. 7,451,099 at 27:14-34, 27:46-51; Fig 16A Oracle Text White Paper, p. 11, 18, 19.
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	"Or, the system could navigate through the Web on its own and notify the user if it found a page or site of presumed interest." Page 1. "In terms of searching, queries are brokered within one node by deciding which child nodes are the most promising candidates for the retrieval process. This is done by determining the content of the query and using a sitemap containing information about the content of every node in the (sub)hierarchy: the query is brokered to those nodes with a content that best matches the content of the

		query. The results of the child nodes are then merged and returned to the parent node or, eventually, to the initiator of the query." Page 2. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15; Fig. 4 PCT Application Pub. No. WO 01/44992 at 17:15-32, 18:14-22 U.S. Patent No. 6,236,768 at 5:12-28 U.S. Patent No. 6,546,386 at 3:37-45, 3:55-57 U.S. Patent No. 7,225,180 at Figs. 9b and 9c; 18:47-63, 20:3-29 U.S. Patent No. 6,122,647 at 3:64-4:8, 6:56-57, 8:1-7, 8:20-24, 10:13-18; Figs. 5-6, 8 Weiss at pgs. 181, 182, 184, 185, 186 Finkelstein at p. 410 U.S. Patent No. 6,473,752 at 4:23-29, 7:51-8:4, 10:34-38, 15:19-23; Figs. 2, 7, 10B U.S. Patent No. 6,606,644 at 11:4-17, 12:2-7 U.S. Patent No. 6,829,780 at 13:40-50, 13:64-67, 14:43-52 U.S. Patent No. 7,225,142 at 4:20-22, 9:38-45 U.S. Patent No. 7,451,099 at 27:46-28:5, 28:30-44.
2	The method according to claim 1, further	"In terms of searching, queries are brokered within one node by deciding which child nodes are the

	comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	most promising candidates for the retrieval process. This is done by determining the content of the query and using a sitemap containing information about the content of every node in the (sub)hierarchy: the query is brokered to those nodes with a content that best matches the content of the query. The results of the child nodes are then merged and returned to the parent node or, eventually, to the initiator of the query." Page 2. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2 U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410.
3	The method according to claim 2, wherein the number of terms added is limited to a predefined number	To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent No. 7,089,236 at 2:42-63, 10:30-36, 9:21-43, 11:56-66, 12:13-21; Table 2

		U.S. Patent No. 6,236,768 at 13:19-34, 11:10-31, 10:42-51, 12:41-49, 13:15-19, 13:42-47 U.S. Patent No. 6,546,386 at 1:64-2:9, 3:37-45, 3:37-45 U.S. Patent No. 6,122,647 at 7:19-25, 7:46-49 Weiss at p. 186 U.S. Patent No. 6,473,752 at 6:45-46, 12:11-14 U.S. Patent No. 6,829,780 at 10:32-36, 13:2-40, 12:13-21, 13:40-50 U.S. Patent No. 7,451,099 at 27:65-28:5 U.S. Patent No. 7,225,180 at 18:47-63, 20:3-29, Figs. 9b and 9c Finkelstein at pgs. 408, 410.
5	The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.	"The goal of OBIWAN is to investigate a novel content-based approach to distributed information retrieval. Websites are clustered into regions. Examples for clustering criteria include but are not restricted to content, geographic location, and association with a specific company. Regions are clustered into super regions, super regions into hyper regions, etc., thus forming a hierarchy of regions. A node of this hierarchy can be browsed by simultaneously browsing its child nodes." Page 2. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g. U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 15:38-46, 17:17-19; Fig. 8c U.S. Patent No. 6,122,647 at 5:65-6:4

		Weiss at pgs. 180, 184; Fig 3 U.S. Patent No. 6,236,768 at 12:41-49, 11:62-12:4, 13:1-8, 13:15-19 U.S. Patent No. 7,451,099 at 3:1-3, 14:4-9, 14:25-36, 27:30-34; Figs. 3, 23 Finkelstein at p. 410 Oracle Text White Paper, p. 19.
10	The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary.	"Documents as well as superdocuments are represented as weighted keyword vectors using the vector space model [23]. The weights are based on term frequencies and inverted document frequencies: It is assumed that multiple occurrences of a word indicate that its meaning contributes to the content of the document more than less frequent terms. However, words that occur with a very high overall frequency (i.e., in the collection of documents in question) do not discriminate between documents within this collection... For each of the surfed pages a keyword vector is calculated. This page vector is compared with the keyword vectors associated with every node to calculate similarities. The nodes with the top matching vectors are assumed to be most related to the content of the surfed page. The accuracy of this text categorization algorithm was validated in [30]." Page 2. To the extent this reference does not teach this claim element, this reference in combination with the knowledge of one of ordinary skill in the art renders this claim element obvious. See, e.g.: U.S. Patent Application Pub. No. 2002/0147738 ¶ 15 PCT Application Pub. No. WO 01/44992 at 18:24-19:1, 22:3-15; Fig 15 U.S. Patent No. 7,225,180 at 16:5-9 U.S. Patent No. 6,122,647 at 3:37-39, 5:50-55 Weiss at p. 185

		Mase at pgs. 377, 379 U.S. Patent No. 6,236,768 at 4:45-55 U.S. Patent No. 6,473,752 at 6:45-60, 12:3-11; Fig. 9 Oracle Text White Paper, p. 19
18	An article of manufacture for use in a computer system, comprising:	"Future work includes the integration of the system into a web browser (right now, cache folders are analyzed) which will allow for more accurate interest detection if other interactions such as scrolling behavior are monitored." Page 7. "In terms of privacy, the existing system stores the profile on the user's machine." Page 7.
	a memory;	"Future work includes the integration of the system into a web browser (right now, cache folders are analyzed) which will allow for more accurate interest detection if other interactions such as scrolling behavior are monitored." Page 7. "In terms of privacy, the existing system stores the profile on the user's machine." Page 7.
	instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:	See Chart for Claim 1 (above).
	defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;	See Chart for Claim 1 (above).
	automatically identifying a set of entities in the selected document content for searching information related thereto using the	See Chart for Claim 1 (above).

	information retrieval system;	
	automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and	See Chart for Claim 1 (above).
	automatically formulating the query to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.	See Chart for Claim 1 (above).
19	The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.	See Chart for Claim 2 (above).

EXHIBIT 4

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

XEROX CORPORATION, Plaintiff, v. GOOGLE, INC., YAHOO! INC., RIGHT MEDIA INC., RIGHT MEDIA LLC, YOUTUBE, INC., and YOUTUBE, LLC, Defendants.	)	C.A. No. 10-136 (JJF) (MPT) JURY TRIAL DEMANDED

YAHOO! INC. AND RIGHT MEDIA LLC'S SUPPLEMENTAL RESPONSES TO INTERROGATORY NOS. 7 THROUGH 9 OF XEROX'S FIRST SET OF INTERROGATORIES

Defendants and Counter-Claim Plaintiffs Yahoo! Inc. and Right Media LLC (collectively, "Yahoo") provide the following supplemental responses to Plaintiff Xerox Corporation's ("Xerox") April 23, 2010 First Set of Interrogatories.[1] These supplemental responses are made in light of ongoing discovery and are based on information presently known to Yahoo, which reserves the right to supplement or modify these supplemental responses based on the discovery of additional or different information and/or in light of expert opinion and/or the Court's claim construction. These supplemental responses are provided without the benefit of the Court's claim construction or knowledge of Xerox's claim construction positions, and with the understanding that a range of claim construction positions may potentially be advanced by the parties and/or adopted by the Court. These supplemented responses therefore should not be

[1] Right Media LLC responds on its own behalf and as the successor in interest to Right Media Inc., which no longer exists.

deemed to admit the correctness or incorrectness of any construction of any limitation of any asserted patents claim.

SUPPLEMENTAL RESPONSES TO INTERROGATORIES

INTERROGATORY NO. 7:

If you contend that any claim of the Patents in Suit is invalid and/or unenforceable, specify each claim that you contend is invalid and/or unenforceable and describe in full for each such claim the basis for your contention, identifying all prior art, all documents and all facts that you believe support your contention.

RESPONSE TO INTERROGATORY NO. 7:

Yahoo maintains and fully incorporates herein each of the general objections and specific objections to this interrogatory listed in Yahoo's May 27, 2010 Objections and Responses to Xerox's First Set of Interrogatories. Subject to and without waiving these objections, Yahoo responds that the asserted claims of the Patents in Suit are invalid for at least the following reasons.

'979 Patent

Invalidity Under 35 U.S.C. § 101

Claims 1 and 18 of the '979 Patent are invalid under 35 U.S.C. § 101 because they claim unpatentable abstract ideas. Moreover, both claims fail the "machine-or-transformation" test indicative of § 101 patent eligible subject matter. Under the machine-or-transformation test, a claimed method is not patentable unless it (1) is tied to a particular machine or apparatus, or (2) transforms a particular article into a different state or thing. *See Bilski v. Kappos*, 561 U.S. __, slip op. at 3 (2010). The method of Claim 1, which analyzes document content to generate abstract queries, recites only general purpose computing equipment and does not meet the statutory requirements for patentable subject matter. Claim 18 is also invalid under 35 U.S.C. §

101 because the mere recitation of general purpose computer and software components does not transform unpatentable method steps into patent-eligible subject matter and does not constitute recitation of a "particular machine."

Invalidity Under 35 U.S.C. §§ 102 and 103

Claims 1 and 18 of the '979 Patent are invalid under 35 U.S.C. §§ 102 and/or 103 in view of the prior art, including that identified below.

The following patents and patent applications are prior art under at least 35 U.S.C § 102(e): U.S. Patent Application 2002/0147738; U.S. Patent 6,122,647; U.S. Patent 6,473,752; U.S. Patent 6,606,644; U.S. Patent 6,829,780; U.S. Patent 7,076,443; U.S. Patent 7,225,142; U.S. Patent 7,418,657; and U.S. Patent 7,451,099 (collectively the "'979 Prior Art Patents").

The following systems were in public use prior to the invention date of the '979 Patent and are prior art under 35 U.S.C. §§ 102(a) and/or 102(b), and are 102(g)(2) prior inventions: eZula, YellowBrix IntelliClix, WebACE, IntelliZap / Zapper, and SemioMap Discovery Search (collectively the "'979 Prior Art Systems").

The following publications are prior art under 35 U.S.C. §§ 102(a) and/or 102(b): Finkelstein et al., *Placing Search in Context: The Concept Revisited*, Proc. of the 10th International World Wide Web Conference (May 1-5, 2001); Han et al., *WebACE: A Web Agent for Document Categorization and Exploration*, Proc. of the 2nd International Conference on Autonomous Agents (May 1998) (the "'979 Prior Art Publications"); and Wiesner et al., Context Matching System and Method, WO/2001/044992 (June 21, 2001).

The '979 Prior Art Patents, the '979 Prior Art Systems and the '979 Prior Art Publications are collectively referred to as the "'979 Prior Art." The status of certain pieces of the '979 Prior Art may be affected by the Court's claim construction. In addition, some items of

art are presently believed to disclose certain elements of the asserted claims inherently. To the extent it is found that such elements are not inherently disclosed, it may be that the relevant claims are alternatively rendered obvious by the asserted reference and the knowledge of a person of ordinary skill in the art alone, or by various other art in combination with the asserted reference. A person of ordinary skill in the art would have found it obvious to combine each '979 Prior Art reference with one or more of the other '979 Prior Art references at least because each such reference relates to analyzing document content, and primarily the content of webpages, to find related information and/or relevant advertisements.

Yahoo incorporates by reference herein the identification by other defendants of any Prior Art as invalidating claims 1 and/or 18 of the '979 Patent under sections 102 and/or 103, to the extent such Prior Art is not specifically identified above. Yahoo reserves the right to use any of the identified references in support of an argument based on a disclosed system in prior use.

Invalidity Under 35 U.S.C. § 102(f) and 116

Pending further investigation, claims 1 and/or 18 '979 Patent may be invalid under 35 U.S.C. §§ 102(f) and 116 for failing to include all inventors of the claimed subject matter.

'994 Patent

Invalidity Under 35 U.S.C. § 101

Claim 9 of the '994 Patent is invalid under 35 U.S.C. § 101 because it claims an unpatentable abstract idea. Moreover, it fails the "machine-or-transformation" test indicative of § 101 patent eligible subject matter. The method of Claim 9, which relates to generic approaches to integrating abstract data and results of analyses thereof with abstract electronic documents, recites only general purpose computing and database equipment and does not meet the statutory requirements for patentable subject matter.

Invalidity Under 35 U.S.C. §§ 102 and 103

Claim 9 of the '994 Patent is invalid under 35 U.S.C. §§ 102 and/or 103 in view of the prior art, including that identified below.

The following patents are prior art under at least 35 U.S.C § 102(e): U.S. Patent 5,564,044; U.S. Patent 5,630,126; U.S. Patent 5,694,192; U.S. Patent 5,659,676; U.S. Patent 5,913,032; U.S. Patent 6,094,684 (collectively the "'994 Prior Art Patents").

The following systems were in public use prior to the invention date of the '994 Patent and are prior art under 35 U.S.C. §§ 102(a) and/or 102(b), and are 102(g)(2) prior inventions: Amazon.com Product Listings, the Internet Movie Database (IMDB), and Crystal Reports (collectively the "'994 Prior Art Systems").

The following publications are prior art under 35 U.S.C. §§ 102(a) and/or 102(b): Silvano Pozzi, et al., *ALIVE: A Distributed Live-link Documentation System*, Electronic Publishing, Vol. 5(3) (Sept. 1992), 131-142 and Premysl Brada, et al., *Dynamic Information Access Using WWW*, Proceedings of the 4th Conference on Interdisciplinary Information Management (1996), 97-102 (the "'994 Prior Art Publications").

The '994 Prior Art Patents, the '994 Prior Art Systems and the '994 Prior Art Publications are collectively referred to as the "'994 Prior Art." The status of certain pieces of the '994 Prior Art may be affected by the Court's claim construction. In addition, some items of art are presently believed to disclose certain elements of the asserted claims inherently. To the extent it is found that such elements are not inherently disclosed, it may be that the relevant claims are alternatively rendered obvious by the asserted reference and the knowledge of a person of ordinary skill in the art alone, or by various other art in combination with the asserted reference. A person of ordinary skill in the art would have found it obvious to combine each

‘994 Prior Art reference with one or more of the other ‘994 Prior Art references at least because each such reference relates to managing relationships between documents and data and analysis results.

Yahoo incorporates by reference herein the identification by other defendants of Prior Art as invalidating claim 9 of the ‘994 Patents under sections 102 and/or 103, to the extent such art is not specifically identified above. Yahoo reserves the right to use any of the identified references in support of an argument based on a disclosed system in prior use.

Invalidity Under 35 U.S.C. § 112

Claim 9 of the ‘994 Patent is invalid under at least plaintiff’s apparent construction (to the extent discernable, if at all, from its response to plaintiff’s interrogatory responses) for at least the following reasons: The claim limitations “storing knowledge,” “validating the accuracy of the knowledge”, “managing the flow of information between the first database and the document database to enable the integration of the data and analysis results with the documents and to automatically update the documents upon the occurrence of a change in the data or analysis results” and “to generate data and analysis results” do not meet the written description and enablement requirements of 35 U.S.C. § 112. To the extent that the term “knowledge” and the phrases “validating the accuracy of the knowledge”, “managing the flow of information”, “data and analysis results” and “data or analysis results” are insolubly ambiguous, claim 9 is indefinite.

Yahoo reserves the right to supplement, revise or render more specific its response to Interrogatory No. 7.

INTERROGATORY NO. 8:

If you contend that any of your '979 Accused Products do not infringe any claim of the '979 Patent, specify, separately for each '979 Accused Product, each claim that you contend is not infringed and describe in full for each such claim the basis for your contention, identifying all documents and all facts that you believe support your contention.

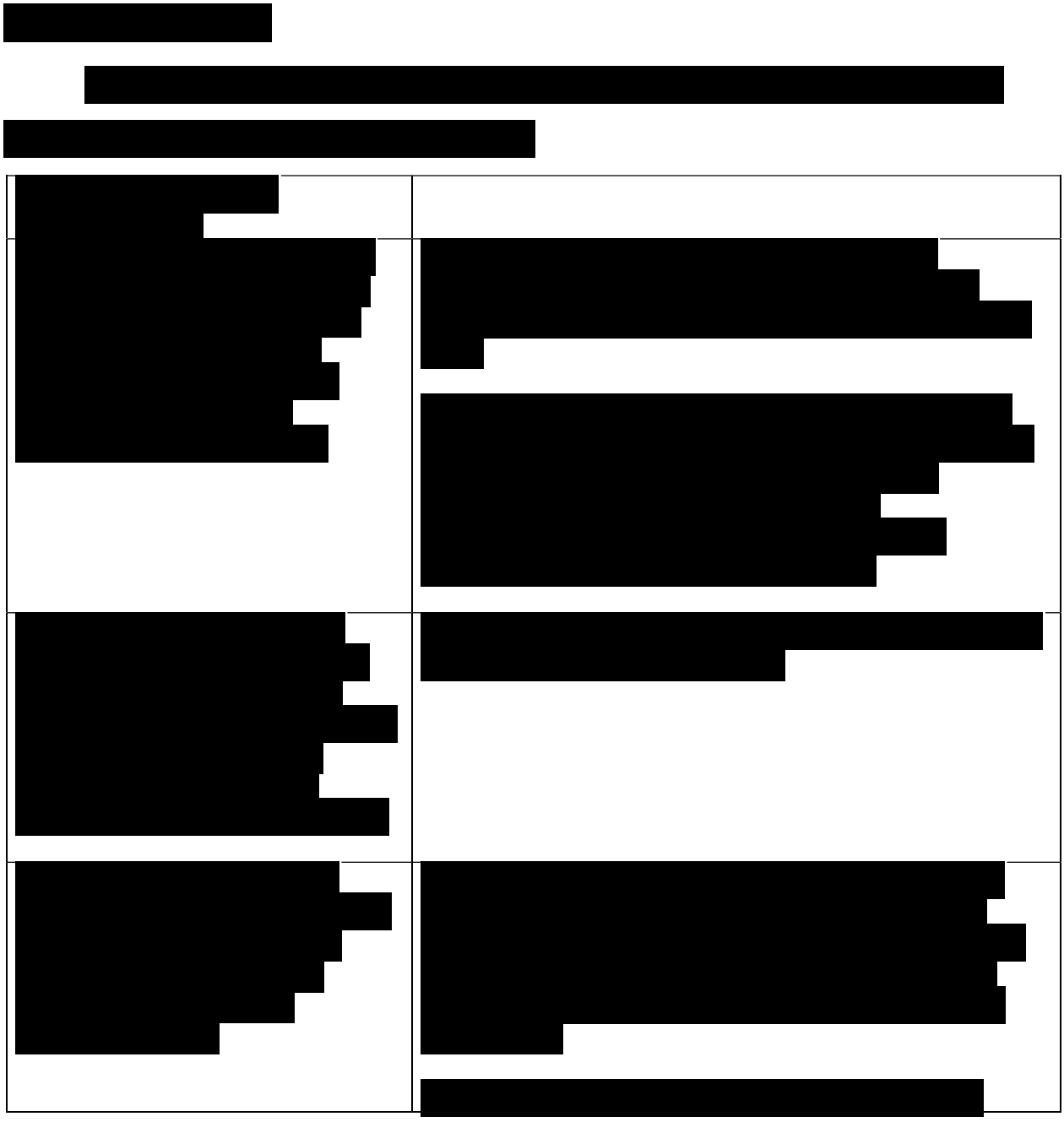
RESPONSE TO INTERROGATORY NO. 8:

Yahoo provides its supplemental response to this interrogatory in view of Xerox's May 27, 2010 response to Yahoo and Right Media's first interrogatory. While Xerox clearly identified the claims of the Patents in Suit that it is asserting in this action, Xerox's response lacks meaningful detail in setting forth the factual bases for its infringement contentions and does not provide sufficient information for Yahoo to discern the nature of Xerox's infringement allegations. Yahoo reserves the right to supplement its response to this interrogatory if and when Xerox provides substantive responses to Yahoo's interrogatories.

Xerox's response is also unclear in identifying the accused products, in particular with respect to its contention that "Yahoo! Search Marketing" and "Yahoo! Publisher Network" infringe claims 1 and 18 of the '979 Patent. Based on correspondence with counsel for Xerox, Yahoo understands that Xerox's references to Yahoo! Search Marketing and Yahoo! Publisher Network are intended in substance to accuse certain Yahoo! Content Match processes to the extent (if any) offered in association with these names. Accordingly, Yahoo provides this supplemental response with the understanding that the instrumentalities that Xerox accuses of infringing the '979 Patent are Yahoo! Content Match, Y!Q Contextual Search, and Right Media Exchange (collectively the "'979 Accused Instrumentalities"). In addition, Yahoo maintains and fully incorporates herein each of the general objections and specific objections to this

interrogatory listed in Yahoo's May 27, 2010 Objections and Responses to Xerox's First Set of Interrogatories. Subject to and without waiving these objections, Yahoo responds that the '979 Accused Instrumentalities do not infringe the asserted claims of the Patents in Suit for at least the following reasons.

REMAINDER OF RESPONSE DESIGNATED CONFIDENTIAL OUTSIDE COUNSEL ONLY



[redacted]

[redacted]

[redacted]	
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

[redacted]

<u>INTERROGATORY NO. 9</u>:

[redacted]

RESPONSE DESIGNATED CONFIDENTIAL OUTSIDE COUNSEL ONLY

[redacted]

RESPONSE TO INTERROGATORY NO. 9:

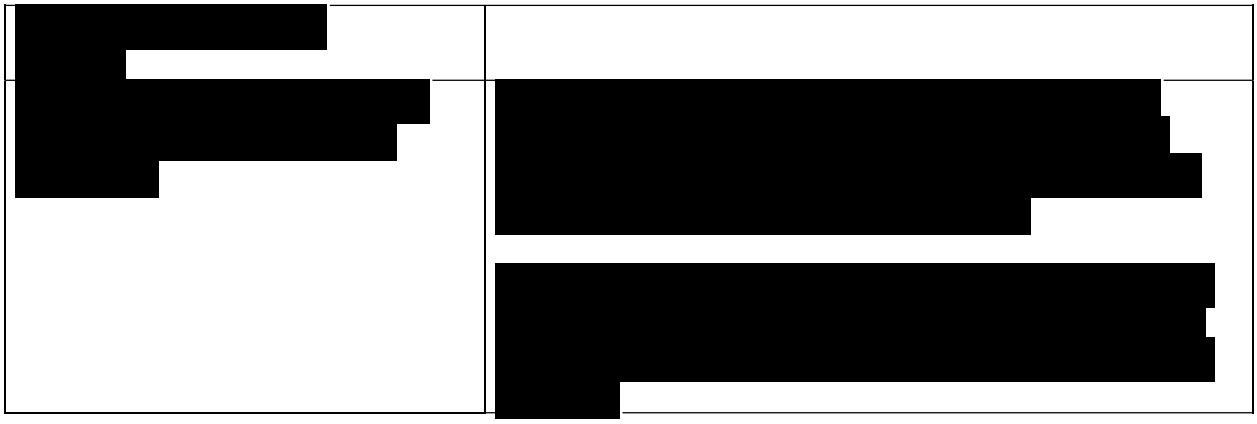
Yahoo provides its supplemental response to this interrogatory in view of Xerox's May 27, 2010 response to Yahoo and Right Media's first interrogatory. While Xerox clearly identified the claims of the Patents in Suit that it is asserting in this action, Xerox's response lacks meaningful detail in setting forth the factual bases for its infringement contentions and does not provide sufficient information for Yahoo to discern the nature of Xerox's infringement allegations. Yahoo reserves the right to supplement its response to this interrogatory if and when Xerox provides substantive responses to Yahoo's interrogatories. In addition, Yahoo maintains and fully incorporates herein each of the general objections and specific objections to this interrogatory listed in Yahoo's May 27, 2010 Objections and Responses to Xerox's First Set of Interrogatories.

Subject to and without waiving these objections, Yahoo responds that Yahoo's '994 Accused Products do not infringe the asserted claims of the Patents in Suit for at least the following reasons.

REMAINDER OF RESPONSE DESIGNATED CONFIDENTIAL OUTSIDE COUNSEL ONLY

[redacted]

[redacted]	



By: /s/ Jesse Dyer

Matthew B. Lehr (#2370)
Anthony I. Fenwick (admitted *pro hac vice*)
Jill Zimmerman (admitted *pro hac vice*)
Jesse Dyer (admitted *pro hac vice*)
DAVIS POLK & WARDWELL LLP
1600 El Camino Real
Menlo Park, CA 94025
650-752-2000
650-752-2111 (fax)

Attorneys For Defendants Yahoo! Inc. and Right Media LLC

July 9, 2010

CERTIFICATE OF SERVICE

I, hereby certify that on July 9, 2010, copies of the foregoing were caused to be served upon the following in the manner indicated:

Lawrence C. Ashby, Esquire — *VIA ELECTRONIC MAIL*
John G. Day, Esquire
Lauren E. Maguire, Esquire
ASHBY & GEDDES
500 Delaware Avenue, 8th Floor
Wilmington, DE 19801

Richard J. Stark, Esquire — *VIA ELECTRONIC MAIL*
Andrei Harasymiak, Esquire
Peter A. Emmi, Esquire
Scott A. Leslie, Esquire
Allison M. Snyder, Esquire
CRAVATH, SWAINE & MOORE LLP
825 Eighth Avenue
New York, NY 10019

Richard L. Horwitz, Esquire — *VIA ELECTRONIC MAIL*
David E. Moore, Esquire
POTTER ANDERSON & CORROON LLP
Hercules Plaza – 6th Floor
1313 North Market Street
Wilmington, DE 19801

Charles K. Verhoeven, Esquire — *VIA ELECTRONIC MAIL*
David A. Perlson, Esquire
Brian C. Cannon, Esquire
QUINN EMANUEL URQUHART
& SULLIVAN, LLP
50 California Street
22nd Floor
San Francisco, CA 94111

/s/ Jesse Dyer
Jesse Dyer

EXHIBIT 5

REDACTED

IN ITS

ENTIRETY

EXHIBIT 6

REDACTED

IN ITS

ENTIRETY