	otherwise. When the next user enters a search request, the search request and the user's personal data are combined to form groupings containing key term groupings, key terms and personal data groupings, category and personal data groupings, rating and personal data groupings, etc. Articles associated with these groupings are then retrieved from the index, and their relevancy scores are used or combined to determine their rankings." 7:14-42: "Another embodiment of the present invention keeps track of the full queries, or portions thereof such as key terms groupings, which are entered by users having certain personal data characteristics. In this embodiment, queries or portions thereof such as key term groupings, are stored within an index, preferably along with the personal data and a previous-user relevancy score for each query. The previous-user relevancy score for a particular query or portion thereof can be: the number of times the query was entered by all users; the number of times a query was entered by unique users; the number of times a query was entered by a particular group of unique users sharing a particular personal data characteristic; the product, sum or average of the previous-user relevancy scores of all or some of the articles under the query or portion(s) thereof for all users, unique users or a particular group of users; the product, sum or average of the previous-user relevancy scores of all or some of the articles under the query or portion(s) thereof for all users, unique users or a particular group of users and the number of times a query was entered by all users all users, unique users or a particular group of users: or any combination of these or other indicators of relevancy of the particular query, or portions thereof, to a particular person or group having certain personal data characteristics. These previous-user relevancy scores for the queries, or portions thereof, can be normalized by factors such as time, number of previous users sharing a particular personal data characteristic, or otherwise by dividing or otherwise altering the raw scores by the normalizing factor or factors." *See also*: **MLADENIC** at 3, 8. **REFUAH** at Abstract; 5:34-50; 6:5-15; 8:31-39; 20:31-37. **SCHUETZE** at 10:14-18; 10:32-39; 11:12-17; 28:65 – 29:6; 34:34-37; *See generally* 17:47 – 18:27.

	WASFI at 58, 60, 61. Autonomy Press Release, at 1 Autonomy Technology Whitepaper, at AUT00069-70 **MONTEBELLO** at 3-4.
c) estimating parameters of a learning machine, wherein the parameters define a User Model specific to the user and wherein the parameters are estimated in part from the user-specific data files;	CULLISS, 3:57-65: "A cumulative score can be kept with regard to these occurrences of certain classified key terms, queries or visited URLs to quantify how strongly someone is associated with a particular item of personal data. The score can be normalized over time, frequency or other activity such as the number of searches performed, the amount of time spent online, the amount of time spent browsing on a particular subject, the number of URLs or articles selected for a particular subject, or otherwise." 4:54 – 5:10: "Then, the user can be identified as having the personal data characteristic of being a sports fan and having an interest in finance because there are three queries relating to sports ("sports scores," "football," and "nba") and five queries containing key words relating to finance ("stock quotes," "cnnfn," "junk bonds," "stock quotes," and "dow jones"). This can be done by keeping a cumulative score for a user for search requests or URLs. For example, whenever there is a match (whole or partial) between a search request or URL and an item of personal data, a record for the user can be updated to give a +1 for that item of personal data. A cumulative score can be developed for the user for each item of personal data, called a personal data item score. When the personal data item score of the user reaches a certain threshold, then the item of personal data can be said to be associated with the user. Additionally or alternatively, the strength of the association can be determined by the cumulative personal data item score. The personal data item score for each item of personal data can be normalized by any normalizing factor, such as the number of requests entered, the number of URLs visited, the average personal data item score for other users in that item of personal data, the median personal data item score for other users in that item of personal data or otherwise." *See also*:

	MLADENIC, at 9, 10. **REFUAH** at 2:9-35; 6:49-64; 8:30-58; *See generally* 14:21 – 15:45. **SCHUETZE** at 27:44-64. *See also* 27:65 – 28:14. **WASFI** at 58, 61, 63. Autonomy Press Release, at 1 Autonomy Technology Whitepaper, at AUT00069-70 Autonomy Agentware User Guide, at AUT00002 **MONTEBELLO** at 3, 4.
d) analyzing a document d to identify properties of the document;	CULLISS, 2:25-37: "the present invention maintains an index of key words, terms, data or identifiers in English or other languages, computer code, or encryption which are collectively referred to as key terms . . . The articles can each be associated with one or more of these key terms by any conceivable method of association now known or later developed. A key term score is associated with each article for each of the key terms. Optionally, a key term total score can also be associated with the article." 2:43-46: "Articles can have their key term scores or key term total scores altered according to whether they were displayed to a user, whether they were selected by a user, how much time the user spent with the article, etc." 5:36-48: "The personal data can be used to recall different lists of articles in response to new queries from new users. In this respect, it is possible to simply store all elements of personal data, individually or in key term groupings, within the index separately, with components of the query or otherwise. When the next user enters a search request, the search request and the user's personal data

	are combined to form groupings containing key term groupings, key terms and personal data groupings, category and personal data groupings, rating and personal data groupings, etc. Articles associated with these groupings are then retrieved from the index, and their relevancy scores are used or combined to determine their rankings." 10:47-52: "To present personalized search results to a particular person searching with a particular term or query, the present invention may display a number of articles from a number of the narrower related key term groupings or queries which are ranked by their respective previous-user relevancy scores." *See also*: **MLADENIC**, at 4, 12. **REFUAH** at 7:53 – 8:6; 20:31-37; 21:6-30. **SCHUETZE** at 3:40-44; 5:59-63; 10:20-31; Fig.1. *See generally* 11:42 – 15:19. **WASFI** at 61. Autonomy Press Release, at 1 Autonomy Technology Whitepaper, at AUT00068-69, AUT00071 Autonomy Agentware User Guide, at AUT00002, AUT00117 **MONTEBELLO** at 3, 4.
e) estimating a probability P(u|d) that an unseen document d is of interest to the user u, wherein the	CULLISS, 2:43-51: "Articles can have their key term scores or key term total scores altered according to whether they were displayed to a user, whether they were selected by a user, how much time the user spent with the article, etc. In this application, the phrase previous-user relevancy score, designated by the generic label "PRS," will be used to refer to any of the key term score, key term

probability P(u|d) is estimated by applying the identified properties of the document to the learning machine having the parameters defined by the User Model; and	total score, key term probability score, comparison score, or other ranking score determined by the previous search activity of users." 5:18-21: "When a first user enters a search query, the personal data can be considered part of the request and stored within or added to the index, individually or in groupings with other items of data such as key terms, categories, or ratings." 5:36-62: "The personal data can be used to recall different lists of articles in response to new queries from new users. In this respect, it is possible to simply store all elements of personal data, individually or in key term groupings, within the index separately, with components of the query or otherwise. When the next user enters a search request, the search request and the user's personal data are combined to form groupings containing key term groupings, key terms and personal data groupings, category and personal data groupings, rating and personal data groupings, etc. Articles associated with these groupings are then retrieved from the index, and their relevancy scores are used or combined to determine their rankings. For example, as illustrated in FIG. 1, if a first user enters a search request at step 10 of Alpha and has personal data characteristics of PS1 and PS5, then the request can be combined in step 20 with the personal data to form the following groupings: Alpha-PS1 and Alpha-PS5. In addition, other groupings or permutations such as PS1-PS5 and Alpha-PS1-PS5 are also possible and can be stored within the index. These groupings are stored within the index and the relevancy scores of selected articles are updated at step 30 according to methods described in my previous applications. To initially retrieve articles for presentation to the first user using a conventional search engine, just the key term "Alpha" can be used as a key term to pull articles from within an index." 6:37-63: (showing different relevancy scores for different articles retrieved for "shoes" dependent on whether the user is a man or a woman) 8:45-50: "For example, the query or key term grouping such as 'pump shoes' may have different relevancy scores depending on whether the previous searchers were women or men, whereas the rankings may not be different for a personal data characteristic such as profession." *See also*: **MLADENIC** at 5, 7, 10, 12, Table 2.

	REFUAH at 3:56 – 4:4; 17:32 – 18:4. **SCHUETZE** at 18:62 – 19:10. **WASFI** at 60, 61. Autonomy Press Release, at 1-2 Autonomy Technology Whitepaper, at AUT00069, AUT00071 Autonomy Agentware User Guide, at AUT00119 **MONTEBELLO** at 4.
f) using the estimated probability to provide automatic, personalized information services to the user.	CULLISS, 2:39-51: "As described in my previous applications, the invention can accept a search query from a user and a search engine will identify matched articles and display squibs of the matched articles in accordance with their comparison scores. Articles can have their key term scores or key term total scores altered according to whether they were displayed to a user, whether they were selected by a user, how much time the user spent with the article, etc. In this application, the phrase previous-user relevancy score, designated by the generic label "PRS," will be used to refer to any of the key term score, key term total score, key term probability score, comparison score, or other ranking score determined by the previous search activity of users." 10:53 – 11:10 (describing personalized results for "shoe") *See also*: **MLADENIC** at 2, 12. **REFUAH** at Abstract, 2:63 – 3:11; 3:47-55; 17:21-43; 18:56-65; 23:11-28.

	SCHUETZE at 1:29-33; 7:54-60. **WASFI** at 61. Autonomy Press Release, at 1 Autonomy Technology Whitepaper, at AUT00068 Autonomy Agentware User Guide, at AUT00002, AUT00004, AUT00116-117 **MONTEBELLO** at 4.
Claim 11	
11. The method of claim **1** further comprising estimating a posterior probability P(u|d,q) that the document d is of interest to the user u, given a query q submitted by the user.	CULLISS, 2:39-51: "As described in my previous applications, the invention can accept a search query from a user and a search engine will identify matched articles and display squibs of the matched articles in accordance with their comparison scores. Articles can have their key term scores or key term total scores altered according to whether they were displayed to a user, whether they were selected by a user, how much time the user spent with the article, etc. In this application, the phrase previous-user relevancy score, designated by the generic label "PRS," will be used to refer to any of the key term score, key term total score, key term probability score, comparison score, or other ranking score determined by the previous search activity of users." 5:40-48: "When the next user enters a search request, the search request and the user's personal data are combined to form groupings . . . Articles associated with these groupings are then retrieved from the index, and their relevancy scores are used or combined to determine their rankings." 9:41-49: "As described in my earlier applications, when a new or second user enters a search query containing one or more words, the system can look for related key term groupings or queries that contain the original query or portions thereof and suggest those additional words, groupings, or queries or portions thereof, of the narrower related key term groupings or queries to refine the search. Preferably, the related key term groupings or queries will be narrower related key term groupings or queries, which are more narrow in scope."

	10:1-7: "To present personalized narrower related key term groupings to a user, the system can present the narrower related key term groupings that include not only at least a portion of the original search request, but also at least a portion of the user's personal data. These narrower related key term groupings can be presented in order of superiority according to their previous-user relevancy scores." 10:47-52: "To present personalized search results to a particular person searching with a particular term or query, the present invention may display a number of articles from a number of the narrower related key term groupings or queries which are ranked by their respective previous-user relevancy scores." *See also*: **MLADENIC** at 2. **REFUAH** at 17:21-43. **SCHUETZE** at 21:57 – 22:16, 22:31-48, 30:58 – 31:13.
Claim 22	
22. The method of claim **1** wherein the monitored user interactions include a sequence of interaction times	CULLISS, 2:43-46: "Articles can have their key term scores or key term total scores altered according to whether they were displayed to a user, whether they were selected by a user, how much time the user spent with the particle, etc." 3:30-35: "Personal activity data includes data about past actions of the user, such as reading habits, viewing habits, searching habits, previous articles displayed or selected, previous search requests entered, previous or current site visits, previous key terms utilized within previous search requests, and time or date of any previous activity." *See also*:

	MLADENIC at 2. **REFUAH** at 5:57-58. **WASFI** at Abstract, 57. Autonomy Press Release, at 1 Autonomy Agentware User Guide, at AUT00124
Claim 32	
32. A program storage device accessible by a central computer, tangibly embodying a program of instructions executable by the central computer to perform method steps for providing automatic, personalized information services to a user u, the method steps comprising:	*See* citations for claim 1[preamble].
a) transparently monitoring user interactions with data while the user is engaged in normal use of a client computer in communication with the central computer;	*See* citations for claim 1[a].
b) updating user-specific data files, wherein the user-specific data files comprise the monitored user	*See* citations for claim 1[b].

interactions with the data and a set of documents associated with the user;	
c) estimating parameters of a learning machine, wherein the parameters define a User Model specific to the user and wherein the parameters are estimated in part from the user-specific data files;	*See* citations for claim 1[c].
d) analyzing a document d to identify properties of the document;	*See* citations for claim 1[d].
e) estimating a probability P(u|d) that an unseen document d is of interest to the user u, wherein the probability P(u|d) is estimated by applying the identified properties of the document to the learning machine having the parameters defined by the User Model; and	*See* citations for claim 1[e].
f) using the estimated probability to provide automatic, personalized information services to the user	*See* citations for claim 1[f].
Claim 34	

34. The program storage device of claim **32** wherein analyzing the document d provides for the analysis of documents having multiple distinct media types.	CULLISS, 1:22-28 ("The Internet is an extensive network of computer systems containing hundreds of millions of documents, files, databases, text collections, audio clips, video clips and samples of any other type of information (collectively 'articles'). As explained in my earlier referenced patent applications, search engines are used to locate articles over the Internet."); 2:19-24 ("The Internet is an extensive network of computer systems containing hundreds of millions of documents, files, databases, text collections, audio clips, video clips and samples of any other type of information, collectively referred to as articles and designated herein by the generic labels A1, A2, A3, etc.") 2:19-22: Culliss discloses operating on the Internet, which contains "hundreds of millions of documents, files, databases, text collections, audio clips, video clips and samples of any other type of information." 9:14-17. *See also*: **SCHUETZE** at 4:12-35, 5:48-58, "Summary of the Invention", 5:43 – 8:32 **MLADENIC** at 2. **REFUAH** at 1:63 – 2:2. **WASFI** at 57, 58. Autonomy Press Release, at 2 Autonomy Agentware User Guide, at AUT00004-5
'276 Patent	CULLISS
Claim 1	
A computer-implemented method for providing	*See* citations for '040 Patent, claim 1[preamble].

personalized information services to a user, the method comprising:	
a) transparently monitoring user interactions with data while the user is engaged in normal use of a browser program running on the computer;	*See* citations for '040 Patent, claim 1[a]. 3:52-56: "certain known articles or URLs can be detected in a users searching or browsing habits, such as those relating to CNNfn (www.cnnfn.com) or Quote.com (www.quote.com), and also used to classify the user as someone interested in finance." *See also*: **MLADENIC** at 2. **REFUAH** at 5:57-58. **WASFI** at Abstract, 57.
b) analyzing the monitored data to determine documents of interest to the user;	*See* citations for '040 Patent, claim 1[b].
c) estimating parameters of a user-specific learning machine based at least in part on the documents of interest to the user;	*See* citations for '040 Patent, claim 1[c].
d) receiving a search query from the user;	CULLISS, 2:39-51: "As described in my previous applications, the invention can accept a search query from a user and a search engine will identify matched articles and display squibs of the matched articles in accordance with their comparison scores. Articles can have their key term scores or key term total scores altered according to whether they were displayed to a user, whether they were selected by a user, how much time the user spent with the article, etc. In this application, the phrase previous-user relevancy score, designated by the generic label "PRS," will be used to refer to

	any of the key term score, key term total score, key term probability score, comparison score, or other ranking score determined by the previous search activity of users." *See also*: **JOACHIMS** at 2, Fig. 1. **MLADENIC** at 1, 2. **REFUAH** at 1:63 – 2:2; 3:12-24. **SCHUETZE** at 21:57 – 22:16; 22:31-48. Autonomy Agentware User Guide, at AUT00002 **MONTEBELLO** at 3.
e) retrieving a plurality of documents based on the search query;	CULLISS, 2:39-51: "As described in my previous applications, the invention can accept a search query from a user and a search engine will identify matched articles and display squibs of the matched articles in accordance with their comparison scores. Articles can have their key term scores or key term total scores altered according to whether they were displayed to a user, whether they were selected by a user, how much time the user spent with the article, etc. In this application, the phrase previous-user relevancy score, designated by the generic label "PRS," will be used to refer to any of the key term score, key term total score, key term probability score, comparison score, or other ranking score determined by the previous search activity of users." *See also*: **MLADENIC** at 1. **REFUAH**, 1:63 – 2:2; 3:12-24. **SCHUETZE**, 21:57 – 22:16; 22:31-48.

	MONTEBELLO at 3.
f) for each retrieved document of said plurality of retrieved documents: identifying properties of the retrieved document, and applying the identified properties of the retrieved document to the user-specific learning machine to estimate a probability that the retrieved document is of interest to the user; and	*See* citations for '040 Patent, claim 1[d, e].
g) using the estimated probabilities for the respective plurality of retrieved documents to present at least a portion of the retrieved documents to the user.	*See* citations for '040 Patent, claim 1[f].
Claim 3	
3. The method of claim **1**, wherein transparently monitoring user interactions with data comprises monitoring user interactions with data during multiple different modes of user interaction with network data.	CULLISS, 2:43-46: "Articles can have their key term scores or key term total scores altered according to whether they were displayed to a user, whether they were selected by a user, how much time the user spent with the article, etc." 3:29-35: "Personal activity data includes data about past actions of the user, such as reading habits, viewing habits, searching habits, previous articles displayed or selected, previous search requests entered, previous or current site visits, previous key terms utilized within previous search requests, and time or date of any previous activity."

	3:46-56: "Users can explicitly specify their own personal data, or it can be inferred from a history of their search requests or article viewing habits. In this respect, certain key words or terms, such as those relating to sports (i.e. "football" and "soccer"), can be detected within search requests and used to classify the user as someone interested in sports. Also, certain known articles or URLs can be detected in a users searching or browsing habits, such as those relating to CNNfn (www.cnnfn.com) or Quote.com (www.quote.com), and also used to classify the user as someone interested in finance." *See also*: **MLADENIC** at 3. **REFUAH** at 5:34-50. **SCHUETZE** at 18:11-17. **WASFI** at 58, 61. Autonomy Press Release, at 1
Claim 5	
5. The method of claim **1**, further comprising analyzing the monitored data to determine documents not of interest to the user, and wherein estimating parameters of a user-specific learning machine further comprises estimating parameters of a user-specific learning machine based at least in part on the documents	*See* citations for '040 Patent, claim 1. *See also*: **SCHUETZE** at 18:62 – 19:10 **CULLISS** at 2:43-51; 5:36-62. **MLADENIC** at 5, 7, 10, 12, Table 2. **REFUAH** at 3:56 – 4:4; 17:32 – 18:4.

not of interest to the user.	**WASFI** at 60, 61. Autonomy Press Release, at 1 Autonomy Technology Whitepaper, at AUT00070 Autonomy Agentware User Guide, at AUT00123
Claim 6	
6. The method of claim **1**, wherein monitoring user interactions with data for a document comprises monitoring at least one type of data selected from the group consisting of information about the document, whether the user viewed the document, information about the user's interaction with the document, context information, the user's degree of interest in the document, time spent by the user viewing the document, whether the user followed at least one link contained in the document, and a number of links in the document followed by the user.	CULLISS, 2:43-46: "Articles can have their key term scores or key term total scores altered according to whether they were displayed to a user, whether they were selected by a user, how much time the user spent with the article, etc." 3:27-35: "Personal interest data includes items such as interests, hobbies, sports, profession or employment, areas of skill, areas of expert opinion, areas of deficiency, political orientation, or habits. Personal activity data includes data about past actions of the user, such as reading habits, viewing habits, searching habits, previous articles displayed or selected, previous search requests entered, previous or current site visits, previous key terms utilized within previous search requests, and time or date of any previous activity." 3:46-48: "Users can explicitly specify their own personal data, or it can be inferred from a history of their search requests or article viewing habits." *See also*: **MLADENIC** at 8. **REFUAH** at 5:34-50, 14:54-59. **SCHUETZE** at 5:36-40, 11:12-14, 18:11-17. **WASFI** at Abstract, 60, 61.

Claim 7	
7. The method of claim **1**, wherein said plurality of retrieved documents correspond to a respective plurality of products.	CULLISS, 9:55- 10:13: "For example, the query "High-Heels" can be identified as related to the subject "shoes" as a particular narrower query of the broader request "shoes." Accordingly, the query "High-Heels" should be considered along with other queries that actually contain the word "shoes." One way these queries related to the subject of the original search query, but which do not actually contain portions of the original search query can be identified is by first utilizing a thesaurus database of equivalent terms for terms in the original search query. Narrower queries or narrower key term groupings that contain one or more of these equivalent terms can then be identified as narrower related key term groupings. To present personalized narrower related key term groupings to a user, the system can present the narrower related key term groupings that include not only at least a portion of the original search request, but also at least a portion of the user's personal data. These narrower related key term groupings can be presented in order of superiority according to their previous-user relevancy scores. For example, when a woman enters the search request "shoes," the system can look for narrower related queries or key term groupings which contain or are related to the term "shoes" and which have been entered by previous users having similar personal data, such as that of being a "woman."" *See generally* 9:55 – 11:33. *See also*: **MLADENIC** at 2, 8, Fig. 2. **REFUAH** at 1:63 – 2:2, 3:56 – 4:4, 7:24-32, 18:35-39, 18:40-55. **SCHUETZE** at 35:66 – 36:8.
Claim 21	
21. The method of claim **1**, wherein using the estimated probabilities for the	2:39-46: "the invention can accept a search query from a user and a search engine will identify matched articles and display squibs of the matched articles in accordance with their comparison scores. Articles can have their key term scores or key term total scores altered according to whether

respective plurality of retrieved documents to present at least a portion of the retrieved documents to the user comprises presenting to the user at least said portion of the retrieved documents based on the estimated probability that the retrieved document is of interest to the user and the relevance of the retrieved document to the search query.	they were displayed to a user, whether they were selected by a user, how much time the user spent with the article, etc." *See* citations for claim 1[g].
Claim 22	
22. The method of claim **1**, wherein identifying properties of the retrieved document comprises identifying properties selected from the properties consisting of a topic associated with the retrieved document, at least one product feature extracted from the retrieved document, an author of the retrieved document, an age of the retrieved document, a list of documents linked to the retrieved document, a number of users who have accessed the retrieved document, and a	CULLISS, 2:26-37: "As described in my previous applications, the present invention maintains an index of key words, terms, data or identifiers in English or other languages, computer code, or encryption which are collectively referred to as key terms and represented herein by the generic labels "Alpha," "Beta," "Gamma," "Delta," "Epsilon," etc. The articles can each be associated with one or more of these key terms by any conceivable method of association now known or later developed. A key term score is associated with each article for each of the key terms. Optionally, a key term total score can also be associated with the article." *See also*: **MLADENIC** at 3, 4, 12. *See generally* 3-6. **REFUAH** at 7:53 – 8:6, 9:50-59, 20:19-30, 21:6-30. *See generally* 20:19- 21:36. **SCHUETZE** at 6:58 – 7:15, 10:40-56, Fig. 3. *See generally* 17:47 – 18:27. **WASFI** at 61.

number of users who have saved the retrieved document in a favorite document list.	

EXHIBIT 3 F

Exhibit 3-F

Claim Chart of U.S. Patent No. 7,631,032 to Refuah et al. ("REFUAH")

as prior art to

Asserted Claims of U.S. Patent No. 6,9981,040 ("'040 Patent")
and
Asserted Claims of U.S. Patent No. 7,685,276 ("'276 Patent")

'040 Patent	REFUAH
Claim 1	
A computer-implemented method for providing automatic, personalized information services to a user u, the method comprising:	Refuah describes a persona-based and mood-based method of interacting with the Internet. The use's normal interactions are monitored to determine the user's persona and mood, which can be thought of as a temporary modification to the user's persona. (Refuah at 2:39-40.) The system then guides users to certain sites, affects searches, changes displays and layouts, etc. based on the user's profile. (*Id.* at 13:64 to 14:14.) *See* REFUAH, Abstract: "A method of a user interacting with an Internet, comprising: tracking interactions of the user with an Internet; analyzing said tracked interactions to determine at least one aspect of a user's interaction with the Internet; and modifying future interactions of said user with said Internet, responsive to said determined aspect, wherein said modified interactions comprise site-content related interactions with a plurality of unrelated sites. Preferably, the aspect is adapted in real-time to reflect changes in the tracked interactions." 1:63 – 2:2: "One object of some preferred embodiments of the invention is to provide a method of aiding information search and retrieval on an Internet. In a preferred embodiment of the invention, Internet searching is personalized to a particular user's profile. Alternatively or additionally, matching up of a supplier and a buyer, of a goods and/or a service, is facilitated, based on such personalization."

	2:36-49: "Another aspect of some preferred embodiments of the invention relates to providing such electronic persona with "moods", which define an instantaneous configuration of preferences and/or outlook. Typically, the mood modifies a persona. However, a mood may also operate without a persona. In one example, a particular persona may include a preference for difficult language. However, the persona's mood may be an "easy-going" mood, in which cases WWW sites having a simpler sentence structure and more graphics will be preferred. In another example, a mood may change between a "rush" mood, in which a user does not want to download large images and a leisure mood, where a user is willing to wait for long downloaded ad is willing to view advertisements if this makes his WWW access cheaper."
a) transparently monitoring user interactions with data while the user is engaged in normal use of a computer;	REFUAH, Abstract: "A method of a user interacting with an Internet, comprising: tracking interactions of the user with an Internet; analyzing said tracked interactions to determine at least one aspect of a user's interaction with the Internet; and modifying future interactions of said user with said Internet, responsive to said determined aspect, wherein said modified interactions comprise site-content related interactions with a plurality of unrelated sites. Preferably, the aspect is adapted in real-time to reflect changes in the tracked interactions." 3:3-11: "One aspect of this interaction, developed below, relates to an ability of automatically updating a mood based on actions of a user on the Internet. Thus, actions of a user affect can a style in which an Internet responds. In one example, a harried access to the Internet (not waiting for images to download, short dwell times) will result in the identification/definition of a rushed mood. Thereafter, search engines may steer the user away from sites which require long download times." 5:34-50: "Additionally or alternatively, personality may be updated automatically. In a preferred embodiment of the invention, the mood is updated based on the one or more of the identification of sites visited by a user, the number of site visited, the dwell time at each site, the order in which sites are visited, the contents of the sites, services purchased, information downloaded, actions performed at the sites and/or a predefined or adaptive time-line based function. Alternatively or additionally, a mood, for example a "rush" mood, may be identified by tracking whether a user waits until images are downloaded, whether a user waits for a complete site to download, whether a user follows links and how many links are followed, and/or rate of changing WWW pages and/or sites. These tracked variables may be compared to a standard. Alternatively or additionally, the tracked variables may be

	compared to a previously acquired baseline of a user. Thus, relative changes in dwell time are tracked." 14:54-59: "In a preferred embodiment of the invention, the persona server tracks frequency, length and/or content of electronic communications with the friends, for example by e-mail, by chat group, by Internet telephone or by computer-dialer, to evaluate an instantaneous mood and/or to assess the relative effect of these friends." 19:20-22: "In a preferred embodiment of the invention, a persona of a client may be automatically generated by tracking the way a client interacts with the Internet." *See also*: **CULLISS** at 3:46-56; 5:18-21; 7:14-20. **MLADENIC**, at 3. **SCHUETZE** at 5:36-40; 11:12-14; 18:11-17; 28:65 – 29:6. **WASFI** at Abstract, 60. Autonomy Press Release, at 2 Autonomy Technology Whitepaper, AUT00069-70 Autonomy Agentware User Guide, at AUT00119 **MONTEBELLO** at 3.
b) updating user-specific data files, wherein the user-specific data files comprise the monitored user	REFUAH, Abstract: "A method of a user interacting with an Internet, comprising: tracking interactions of the user with an Internet; analyzing said tracked interactions to determine at least one aspect of a user's interaction with the Internet; and modifying future interactions of said user with said Internet, responsive to said determined aspect, wherein said modified interactions comprise site-

interactions with the data and a set of documents associated with the user;	content related interactions with a plurality of unrelated sites. Preferably, the aspect is adapted in real-time to reflect changes in the tracked interactions." 5:34-50: "Additionally or alternatively, personality may be updated automatically. In a preferred embodiment of the invention, the mood is updated based on the one or more of the identification of sites visited by a user, the number of site visited, the dwell time at each site, the order in which sites are visited, the contents of the sites, services purchased, information downloaded, actions performed at the sites and/or a predefined or adaptive time-line based function. Alternatively or additionally, a mood, for example a "rush" mood, may be identified by tracking whether a user waits until images are downloaded, whether a user waits for a complete site to download, whether a user follows links and how many links are followed, and/or rate of changing WWW pages and/or sites. These tracked variables may be compared to a standard. Alternatively or additionally, the tracked variables may be compared to a previously acquired baseline of a user. Thus, relative changes in dwell time are tracked." 6:5-15: "In a preferred embodiment of the invention, a mood and/or a persona may be updated by modifying continuous parameters. Additionally or alternatively, such updating may include modifying discrete parameters. An example of a continuous parameters is "wait time" which indicates how long a user is willing to wait for a site to be downloaded. An example of a discrete parameter is a level of Parental Guidance rating of sites (PG-13, R, X). In a preferred embodiment of the invention, a plurality of personalities are predefined. Modifying a persona and/or a mood may include switching between such predefined personalities." 8:31-39: "In a preferred embodiment of the invention, said analyzing comprises analyzing previously acquired tracking data. Alternatively or additionally, said analyzing comprises analyzing of currently acquired tracking data. Alternatively or additionally, said determined aspect is modeled using a virtual personality, which is a complex of characteristics that distinguishes an electronic person, for the purpose of interacting with an Internet. Preferably, said virtual personality comprises a persona, which is a static aspect of a personality." 20:31-37: "In a preferred embodiment of the invention, a site may be automatically evaluated by tracing the personas and/or moods of clients who visit the site and/or remain at the site for a significant amount of time. In a preferred embodiment of the invention, such tracing is performed by

	the site server. Additionally or alternatively, the tracing is performed by a persona server and/or an atmosphere server." 22:27-33: "Information about each document that the user views is stored in a recently accessed buffer for subsequent analysis. The recently accessed buffer includes information about the document itself and information about the user's interaction with the document. One possible implementation of a buffer is illustrated in FIG. 14; however, any suitable data structure may be used." *See also*: **CULLISS** at 3:13-35; 5:36-48; 7:14-42. **MLADENIC** at 3, 8. **SCHUETZE** at 10:14-18; 10:32-39; 11:12-17; 28:65 – 29:6; 34:34-37; *See generally* 17:47 – 18:27. **WASFI** at 58, 60, 61. Autonomy Press Release, at 1 Autonomy Technology Whitepaper, at AUT00069-70 **MONTEBELLO** at 3-4.
c) estimating parameters of a learning machine, wherein the parameters define a User Model specific to the user and wherein the parameters are estimated in part from the user-specific data files;	REFUAH, 2:9-35: "One aspect of some preferred embodiments of the invention relates to generating an electronic person having a personality profile. The person preferably defines a personality for the Internet to interact with and/or be personalized to. Since the personality does not exist in a non-electronic form, it may be termed a "virtual personality". In a preferred embodiment of the invention, a user may switch between several personalities. In a preferred embodiment of the invention, a personality includes one or more of demographic information, geographic location, marketing information, subjects of interest to the user and/or other information, such as entertainment habits

	and ownership of a car. Each of the above subjects may include many sub-elements, for example, subjects of interest may include chess, checkers, baseball and swimming. Marketing information may include price preferences and buying habits. Geographical location may include one or more home locations, one or more shopping locations, one or more work locations and/or one or more vacation locations. In a preferred embodiment of the invention, the elements are hierarchically defined. In one example, a geographical location may include one or more of a country, city, neighborhood, street and house number levels. In another example, the area of interest "basketball" divides into a plurality of subject teams, and each team may be further subdivided into a plurality of players of interest. In a preferred embodiment of the invention, the elements include a relative weighting." 6:29-48: "In one preferred embodiment of the invention, a persona is defined as a set of parameters with values associated with each parameter. Additionally or alternatively, the parameters may be organized, for example by subject and/or in a hierarchical manner. In a preferred embodiment of the invention, the persona is organized in an object oriented manner. In a preferred embodiment of the invention, not all persona have the same parameters. In a preferred embodiment of the invention, two types of parameters are used, local and global. Local parameters affect only a small part of the interaction with the Internet. For example "subject of interest=baseball" does not affect browsing of business sites, except perhaps advertisements. However, "image download tolerance time=3 sec" affects the browsing of any site having images. Also "Color scheme=garish" will affect the search results of diverse searches. In contrast to such site-general parameters, a persona may also include site specific parameters, for example, "CNN subscriber number=123456", which affect substantially only interaction with the CNN web site." 6:49-64: "In a preferred embodiment of the invention, parameters may include information, such as "subject of interest=chess". A parameter may also be negative, for example, a blacklisting: "reject=pornography" or "reject if pornography level >3". Additionally or alternatively, a persona may include weighing information, such as relative preference of subjects of interest, for example "baseball=5, basketball=3". Additionally or alternatively, a persona may include functional information, such as how to evaluate a particular parameter, the affect of a parameter and/or evaluate a grade for a particular site, in view of a parameter. Additionally or alternatively, a parameter may be reflexive towards the persona, for example defining how to modify the persona and/or a mood based on user activities. Additionally or alternatively, a parameter may define the traits which should be

	evaluated when determining a suitability of a site to a persona." 7:33-38: "In a preferred embodiment of the invention, a mood is defined as parameters with values that affect a persona. Such values may be, inter alia, single values, ranges of allowed values, scripts, continuous values and/or discrete values. A mood may replace certain parameter values, affect their value and/or affect their relative weighting." 8:30-58: "n a preferred embodiment of the invention, said analyzing comprises analyzing previously acquired tracking data. Alternatively or additionally, said analyzing comprises analyzing of currently acquired tracking data. Alternatively or additionally, said determined aspect is modeled using a virtual personality, which is a complex of characteristics that distinguishes an electronic person, for the purpose of interacting with an Internet. Preferably, said virtual personality comprises a persona, which is a static aspect of a personality. Alternatively or additionally, said virtual personality comprises a mood, which is a dynamic aspect of a personality. Preferably, said mood comprises a rush mood, which favors fast responses. Alternatively or additionally, said persona comprises a meticulous persona, which favors complete responses. In a preferred embodiment of the invention, said personality comprises geographical information. Alternatively or additionally, said personality comprises demographic information. Alternatively or additionally, said personality comprises interests and preference information. Alternatively or additionally, said personality comprises marketing information. Alternatively or additionally, said personality comprises identification and contact information. Alternatively or additionally, said personality comprises relational information, which defines relations between various aspects of the personality. Alternatively or additionally, said personality comprises reflective information, which defines how a personality changes and/or interacts with other electronic entities." *See generally* 14:21 – 15:45. *See also*: **CULLISS** at 3:57-65; 4:54 – 5:10. **MLADENIC**, at 9, 10.

	SCHUETZE at 27:44-64. *See also* 27:65 – 28:14. **WASFI** at 58, 61, 63. Autonomy Press Release, at 1 Autonomy Technology Whitepaper, at AUT00069-70 Autonomy Agentware User Guide, at AUT00002 **MONTEBELLO** at 3, 4.
d) analyzing a document d to identify properties of the document;	REFUAH, 7:53 – 8:6: "Another aspect of some preferred embodiments of the invention relates to evaluating an atmosphere and/or other traits of a site. In a preferred embodiment of the invention, depending on a persona, several characteristics of a site may be defined, which may be used in filtering out and/or prioritizing such a site. Alternatively or additionally, such information may be used database of sites with their associated values is maintained, so that such characteristics do not need to be re-evaluated very often. Another aspect of some preferred embodiments of the invention relates to associating traits and/or an atmosphere with a WWW site. The associations may be stored at a central location. Additionally or alternatively, the associations and/or trait-related keywords and/or values may be associated with each site. Additionally or alternatively, a site may include an identification number, which when used with a proper trait server, provides information about the sites traits and/or a match and/or grade with a particular personality. Additionally or alternatively, such associations may be stored in search indexes, preferably in a manner similar to the storage of key words." 9:50-56: "Preferably, analyzing a content, comprises determining at least one trait of said site. Alternatively or additionally, analyzing a content comprises determining an ambiance of said site. Alternatively or additionally, analyzing comprises analyzing lexicographical characteristics of said site. Alternatively or additionally, analyzing comprises analyzing graphical characteristics of said

	site." 20:31-37: "In a preferred embodiment of the invention, a site may be automatically evaluated by tracing the personas and/or moods of clients who visit the site and/or remain at the site for a significant amount of time. In a preferred embodiment of the invention, such tracing is performed by the site server. Additionally or alternatively, the tracing is performed by a persona server and/or an atmosphere server." 21:6-30:"In a preferred embodiment of the invention, an atmosphere of a site may be automatically evaluated by analyzing the content of a site, in addition to or instead of utilizing a client's reaction to the site or statistics of accessing the site. Various characteristics of a site may be automatically determined. Each of these characteristics and/or combinations thereof may be used to estimate values for traits and/or atmosphere. The characteristics preferably include one or more of: (a) word length; (b) whether certain words and/or phrases used by or associated with the site belong to certain groups, such as "academic words", "swear words", "adult words", "new-age words", "sports words", "baseball words"; (c) sentence complexity; (d) density of displayed text; (e) ratio between images and text; (f) size of text; (g) distribution of colors in image sand in background; (h) number of links; number of links visited, date of last visit, by the client, by the persona, by the mood and/or by other moods, personas and/or clients; (i) size of site; (j) key-words presented by the site; and/or (k) number of images; and/or (l) number and/or type of multimedia files." *See also*: **CULLISS** at 2:43-46; 5:36-48; 10:47-52.

	MLADENIC, at 4, 12. **SCHUETZE** at 3:40-44; 5:59-63; 10:20-31; Fig.1. *See generally* 11:42 – 15:19. **WASFI** at 61. Autonomy Press Release, at 1 Autonomy Technology Whitepaper, at AUT00068-69, AUT00071 Autonomy Agentware User Guide, at AUT00002, AUT00117 **MONTEBELLO** at 3, 4.
e) estimating a probability P(u|d) that an unseen document d is of interest to the user u, wherein the probability P(u|d) is estimated by applying the identified properties of the document to the learning machine having the parameters defined by the User Model; and	REFUAH, 3:56 – 4:4: "Another aspect of some preferred embodiments of the invention relates to personalizing advertisements responsive to a mood and/or a persona. This personalization of advertisements may be in addition to or alternatively to personalization responsive to a particular search and/or other actions performed by a user at a site. In a preferred embodiment of the invention, a site obtains information on a persona and/or a mood of the accessing user and then tailors services and/or advertisements based on the mood or persona. In a preferred embodiment of the invention, when a user enters a book-seller's web site, even if the user has never been at the book-seller, he may be offered books which match his persona and/or mood. It should be appreciated that, in some preferred embodiments of the invention, such a personality is not generally created and/or maintained by the site which uses the information for personalization." 17:32 – 18:4: "The information retrieval may be filtered and/or sorted based on the persona. Additionally, the persona may decide how a particular search word is interpreted. Various combinations of thresholding, grading and sorting may be applied on search results, by comparing them to a persona. One or both of the following two techniques are preferably used to match a persona to a search result, namely key-words and evaluation. In the key-words technique, a search index includes a classification and/or key-words which match parameters such as those described above for a persona. For example, a site may be indexed as being related to a particular type of

	music, which type may match a preference of a client. In the evaluation technique, a site is evaluated for suitability and/or for qualities which are preferred and/or match a particular persona. Example include: number of images in the site, expected download time and/or number of links from the site. In a preferred embodiment of the invention, the presentation of search results may also be parameters of the persona. In one example, the persona can dictate whether or not to grade sites or information files and whether or not to limit the results using criteria such as geographical criteria. Thus, in one case, a strong match will be shown even if its associated geographical location is 1000 miles away. In the other case, only hits having an associated geographical location within 50 miles are shown. Additionally or alternatively, a mix between near and far results may be defined. One or more parameters of a persona may define matching requirements, for example exactness of match and allowed error. These parameters may depend on the preference being matched. In some cases, there is no way to match a parameter of a persona (e.g., if no geographical location is associated with the site). Default behaviour in such cases may also be a parameter of the persona. In one example, entering searching for a pizza store will generate a different web site connection, based on where the connection is from. For example, a user in Brighton, Mass. will be directed to a different pizza store from a user in downtown Boston, even if both stores belong to the same franchise. Additionally or alternatively, a persona may define multiple response sets. In one example, one set includes low-cost book stores and a second set-includes high-cost book stores. This division may be the result of a preference for differentiating between high-cost and low-cost suppliers." *See also*: **CULLISS** at 2:43-51; 5:36-62. **MLADENIC** at 5, 7, 10, 12, Table 2. **SCHUETZE** at 18:62 – 19:10.

	WASFI at 60, 61. Autonomy Press Release, at 1-2 Autonomy Technology Whitepaper, at AUT00069, AUT00071 Autonomy Agentware User Guide, at AUT00119 **MONTEBELLO** at 4.
f) using the estimated probability to provide automatic, personalized information services to the user.	REFUAH, Abstract: "A method of a user interacting with an Internet, comprising: tracking interactions of the user with an Internet; analyzing said tracked interactions to determine at least one aspect of a user's interaction with the Internet; and modifying future interactions of said user with said Internet, responsive to said determined aspect, wherein said modified interactions comprise site-content related interactions with a plurality of unrelated sites. Preferably, the aspect is adapted in real-time to reflect changes in the tracked interactions." 2:63 – 3:11: "Another aspect of some preferred embodiments of the invention relates to using "persona" and/or "mood" (hereafter referred to together as "personality") to define a view of an Internet. In a preferred embodiment of the invention, one or more aspects of browsing and/or using the Internet may be affected by personality. In a preferred embodiment of the invention, the personality affects which data is displayed by the Internet. One aspect of this interaction, developed below, relates to an ability of automatically updating a mood based on actions of a user on the Internet. Thus, actions of a user affect can a style in which an Internet responds. In one example, a harried access to the Internet (not waiting for images to download, short dwell times) will result in the identification/definition of a rushed mood. Thereafter, search engines may steer the user away from sites which require long download times." 3:47-55: "Additionally or alternatively, the personality may be used when entering any WWW site to provide personally tailored service. In one example, a news site will provide happy news for an "up-beat" persona and depressing news to a "pessimistic" persona. In another example, when entering a book-store or a library site, the site can tailor searches performed to the personality, for example, the

	regular interests of the user. In another example, a business mood will be greeted mainly with business news (and business related advertisements)." 13:64- 14:15: "In a preferred embodiment of the invention, a persona and/or a mood may be used to have one or more of the following effects on the interaction between client 10 and Internet 16: (a) preferentially guide client 10 to certain sites; (b) affect the way searches are performed for information and/or web sites; (c) affect the way a particular web site responds to a client's request; (d) affect the display of information; (e) affect the format and/or layout of a site on the client's terminal; (f) affect the interpretation of a client's actions and/or data entry; (g) target promotions and/or advertisements to a client; and/or (h) protect a client from unwanted influences on Internet 16." 17:21-43: "In a preferred embodiment of the invention, a persona is used to personalize information retrieval. Such personalization can affect many methods of information retrieval, including search engines, name servers, intelligent agents, yellow pages, white pages, and searching inside a WWW site, such as searching for articles on Microsoft products inside the Microsoft WWW site. It should be noted in this context that search engines return matches for a particular query, while personality and mood are designed to affect the results of substantially any query, even though a personality does not specifically point out a desired piece of information. The information retrieval may be filtered and/or sorted based on the persona. Additionally, the persona may decide how a particular search word is interpreted. Various combinations of thresholding, grading and sorting may be applied on search results, by comparing them to a persona. One or both of the following two techniques are preferably used to match a persona to a search result, namely key-words and evaluation. In the key-words technique, a search index includes a classification and/or key-words which match parameters such as those described above for a persona. For example, a site may be indexed as being related to a particular type of music, which type may match a preference of a client." 18:56-65: "In a preferred embodiment of the invention, a WWW site may tailor its reactions to the client based on the persona. In one example, the content of links on a page may depend on the

	persona. Additionally or alternatively, the effect of a button may depend on the persona, for example a persona's geographical location. Additionally or alternatively, a Java applet and/or a JavaScript script may utilize persona information in their execution. For example, prior to rendering a button, a Java applet may check if to use a garish background for an upbeat persona or a muted background for a somber persona." 23:11-28: "In an example of utilizing personas for Internet commerce, a client may be a married business man, having two children, on teenage and one a toddler, a dog, an company office in London and family living in London. Also, the client is an exercise freak. All of the above information is preferably part of the client's profile. The client is looking for a hotel in London for business meeting. When such a client connects to WWW sites of hotel chains, he is not required to reenter personal information. Each site can offer an hotel which best suites his needs (near the office, shopping and family). Alternatively or additionally, each site can personalize its response to his query, for example, informing that it does or does not have an exercise spa. Alternatively or additionally, each site can personalize its promotions, for example, offer a low-rate accommodation for an accompanying teenage family member. Alternatively or additionally, the site can personalize advertisements, for example display advertisements for quality dog food available in the London region and/or dog sitters." *See also*: **CULLISS** at 2:39-51. **MLADENIC** at 2, 12. **SCHUETZE** at 1:29-33; 7:54-60. **WASFI** at 61. Autonomy Press Release, at 1 Autonomy Technology Whitepaper, at AUT00068

	Autonomy Agentware User Guide, at AUT00002, AUT00004, AUT00116-117 **MONTEBELLO** at 4.
Claim 11	
11. The method of claim **1** further comprising estimating a posterior probability P(u|d,q) that the document d is of interest to the user u, given a query q submitted by the user.	REFUAH, 17:21-43: "In a preferred embodiment of the invention, a persona is used to personalize information retrieval. Such personalization can affect many methods of information retrieval, including search engines, name servers, intelligent agents, yellow pages, white pages, and searching inside a WWW site, such as searching for articles on Microsoft products inside the Microsoft WWW site. It should be noted in this context that search engines return matches for a particular query, while personality and mood are designed to affect the results of substantially any query, even though a personality does not specifically point out a desired piece of information. The information retrieval may be filtered and/or sorted based on the persona. Additionally, the persona may decide how a particular search word is interpreted. Various combinations of thresholding, grading and sorting may be applied on search results, by comparing them to a persona. One or both of the following two techniques are preferably used to match a persona to a search result, namely key-words and evaluation. In the key-words technique, a search index includes a classification and/or key-words which match parameters such as those described above for a persona. For example, a site may be indexed as being related to a particular type of music, which type may match a preference of a client." 17:49-65: "the presentation of search results may also be parameters of the persona.... [o]ne or more parameters of a persona may define matching requirements, for example exactness of match and allowed error." *See also*: **CULLISS** at 2:39-51, 9:41-49, 10:1-7, 10:47-52. **MLADENIC** at 2. **PAYNE** at 1, 4.

	SCHUETZE at 21:57 – 22:16, 22:31-48, 30:58 – 31:13.
Claim 22	
22. The method of claim **1** wherein the monitored user interactions include a sequence of interaction times	REFUAH, 3:3-11: "One aspect of this interaction, developed below, relates to an ability of automatically updating a mood based on actions of a user on the Internet. Thus, actions of a user affect can a style in which an Internet responds. In one example, a harried access to the Internet (not waiting for images to download, short dwell times) will result in the identification/definition of a rushed mood. Thereafter, search engines may steer the user away from sites which require long download times." 5:34-50: "Additionally or alternatively, personality may be updated automatically. In a preferred embodiment of the invention, the mood is updated based on the one or more of the identification of sites visited by a user, the number of site visited, the dwell time at each site, the order in which sites are visited, the contents of the sites, services purchased, information downloaded, actions performed at the sites and/or a predefined or adaptive time-line based function. Alternatively or additionally, a mood, for example a "rush" mood, may be identified by tracking whether a user waits until images are downloaded, whether a user waits for a complete site to download, whether a user follows links and how many links are followed, and/or rate of changing WWW pages and/or sites. These tracked variables may be compared to a standard. Alternatively or additionally, the tracked variables may be compared to a previously acquired baseline of a user. Thus, relative changes in dwell time are tracked." 14:54-59: "In a preferred embodiment of the invention, the persona server tracks frequency, length and/or content of electronic communications with the friends, for example by e-mail, by chat group, by Internet telephone or by computer-dialer, to evaluate an instantaneous mood and/or to assess the relative effect of these friends." *See also*: **MLADENIC** at 2.

	REFUAH at 5:57-58. **WASFI** at Abstract, 57. Autonomy Press Release, at 1 Autonomy Agentware User Guide, at AUT00124
Claim 32	
32. A program storage device accessible by a central computer, tangibly embodying a program of instructions executable by the central computer to perform method steps for providing automatic, personalized information services to a user u, the method steps comprising:	*See* citations for claim 1[preamble].
a) transparently monitoring user interactions with data while the user is engaged in normal use of a client computer in communication with the central computer;	*See* citations for claim 1[a].
b) updating user-specific data files, wherein the user-specific data files comprise the monitored user interactions with the data and	*See* citations for claim 1[b].

a set of documents associated with the user;	
c) estimating parameters of a learning machine, wherein the parameters define a User Model specific to the user and wherein the parameters are estimated in part from the user-specific data files;	*See* citations for claim 1[c].
d) analyzing a document d to identify properties of the document;	*See* citations for claim 1[d].
e) estimating a probability P(u|d) that an unseen document d is of interest to the user u, wherein the probability P(u|d) is estimated by applying the identified properties of the document to the learning machine having the parameters defined by the User Model; and	*See* citations for claim 1[e].
f) using the estimated probability to provide automatic, personalized information services to the user	*See* citations for claim 1[f].
Claim 34	
34. The program storage	REFUAH, 1:63 – 2:2: "One object of some preferred embodiments of the invention is to provide a

device of claim **32** wherein analyzing the document d provides for the analysis of documents having multiple distinct media types.	method of aiding information search and retrieval on an Internet. In a preferred embodiment of the invention, Internet searching is personalized to a particular user's profile. Alternatively or additionally, matching up of a supplier and a buyer, of a goods and/or a service, is facilitated, based on such personalization" 9:50-56: "Preferably, analyzing a content, comprises determining at least one trait of said site. Alternatively or additionally, analyzing a content comprises determining an ambiance of said site. Alternatively or additionally, analyzing comprises analyzing lexicographical characteristics of said site. Alternatively or additionally, analyzing comprises analyzing graphical characteristics of said site." *See also*: **CULLISS** at 1:22-28, 2:19-24. **MLADENIC** at 2. **SCHUETZE** at 4:12-35, 5:48-58. *See generally* "Summary of the Invention", 5:43 – 8:32. **WASFI** at 57, 58. **TAN** at 4.1. Autonomy Press Release, at 2 Autonomy Agentware User Guide, at AUT00004-5
'276 Patent	REFUAH
Claim 1	
A computer-implemented method for providing personalized information	*See* citations for '040 Patent, claim 1[preamble].

services to a user, the method comprising:	
a) transparently monitoring user interactions with data while the user is engaged in normal use of a browser program running on the computer;	*See* citations for '040 Patent, claim 1[a]. 5:57-65: "The automatic updating may be performed at the user site, for example by tracking the activity of an Internet browser. Such tracking is preferably achieved using a standalone program which monitors the browser and/or TCP/IP connections. Alternatively or additionally, a dedicated TCP/IP stack and/or driver is used. Additionally or alternatively, the updating may be performed by a server, such as a proxy server, through which a significant portion of a user's requests and/or traffic, pass" *See also*: **MLADENIC** at 2. **WASFI** at Abstract, 57.
b) analyzing the monitored data to determine documents of interest to the user;	*See* citations for '040 Patent, claim 1[b].
c) estimating parameters of a user-specific learning machine based at least in part on the documents of interest to the user;	*See* citations for '040 Patent, claim 1[c].
d) receiving a search query from the user;	REFUAH, 1:63 – 2:2: "One object of some preferred embodiments of the invention is to provide a method of aiding information search and retrieval on an Internet. In a preferred embodiment of the invention, Internet searching is personalized to a particular user's profile. Alternatively or additionally, matching up of a supplier and a buyer, of a goods and/or a service, is facilitated, based on such personalization."

	3:12-24: "In one example, a search mechanism, such as yellow pages, white pages, indexes, search engines, intelligent agents and/or registry search, may filter and/or sort search results responsive to personality. In one example, a search may be limited to sites having a minimum percentage of graphics. In another example, a search result may be ordered by their average word length. Alternatively or additionally, the level of detail of the search results may be dependent on the mood, for example, in a rushed mood, only a line will be displayed for each search result. It should be appreciated that a persona is an indirect method of defining search criteria, as opposed to a usual method of defining search criteria, which precisely defines what type of information is desired in a site." 17:21-43: "a persona is used to personalize information retrieval.... including search engines, name servers, intelligent agents, yellow pages, white pages, and searching inside a WWW site, such as searching for articles on Microsoft products inside the Microsoft WWW site." 17:49-65: "the presentation of search results may also be parameters of the persona.... [o]ne or more parameters of a persona may define matching requirements, for example exactness of match and allowed error." *See also*: **CULLISS** at 2:39-51. **MLADENIC** at 1, 2. **SCHUETZE** at 21:57 – 22:16; 22:31-48. Autonomy Agentware User Guide, at AUT00002 **MONTEBELLO** at 3.
e) retrieving a plurality of documents based on the search query;	REFUAH, 1:63 – 2:2: "One object of some preferred embodiments of the invention is to provide a method of aiding information search and retrieval on an Internet. In a preferred embodiment of the invention, Internet searching is personalized to a particular user's profile. Alternatively or

	additionally, matching up of a supplier and a buyer, of a goods and/or a service, is facilitated, based on such personalization." 3:12-24: "In one example, a search mechanism, such as yellow pages, white pages, indexes, search engines, intelligent agents and/or registry search, may filter and/or sort search results responsive to personality. In one example, a search may be limited to sites having a minimum percentage of graphics. In another example, a search result may be ordered by their average word length. Alternatively or additionally, the level of detail of the search results may be dependent on the mood, for example, in a rushed mood, only a line will be displayed for each search result. It should be appreciated that a persona is an indirect method of defining search criteria, as opposed to a usual method of defining search criteria, which precisely defines what type of information is desired in a site." 17:21-43: "a persona is used to personalize information retrieval…. including search engines, name servers, intelligent agents, yellow pages, white pages, and searching inside a WWW site, such as searching for articles on Microsoft products inside the Microsoft WWW site." 17:49-65: "the presentation of search results may also be parameters of the persona…. [o]ne or more parameters of a persona may define matching requirements, for example exactness of match and allowed error." *See also*: **CULLISS** at 2:39-51. **MLADENIC** at 1. **SCHUETZE**, 21:57 – 22:16; 22:31-48. **MONTEBELLO** at 3.
f) for each retrieved document of said plurality of	*See* citations for '040 Patent, claim 1[d, e].

retrieved documents: identifying properties of the retrieved document, and applying the identified properties of the retrieved document to the user-specific learning machine to estimate a probability that the retrieved document is of interest to the user; and	
g) using the estimated probabilities for the respective plurality of retrieved documents to present at least a portion of the retrieved documents to the user.	*See* citations for '040 Patent, claim 1[f].
Claim 3	
3. The method of claim **1**, wherein transparently monitoring user interactions with data comprises monitoring user interactions with data during multiple different modes of user interaction with network data.	REFUAH, 5:34-50: "Additionally or alternatively, personality may be updated automatically. In a preferred embodiment of the invention, the mood is updated based on the one or more of the identification of sites visited by a user, the number of site visited, the dwell time at each site, the order in which sites are visited, the contents of the sites, services purchased, information downloaded, actions performed at the sites and/or a predefined or adaptive time-line based function. Alternatively or additionally, a mood, for example a "rush" mood, may be identified by tracking whether a user waits until images are downloaded, whether a user waits for a complete site to download, whether a user follows links and how many links are followed, and/or rate of changing WWW pages and/or sites. These tracked variables may be compared to a standard. Alternatively or additionally, the tracked variables may be compared to a previously acquired baseline of a user. Thus, relative changes in dwell time are tracked." *See also*:

	CULLISS at 2:43-46, 3:29-35, 3:46-56. **MLADENIC** at 3. **SCHUETZE** at 18:11-17. **WASFI** at 58, 61. Autonomy Press Release, at 1
Claim 5	
5. The method of claim **1**, further comprising analyzing the monitored data to determine documents not of interest to the user, and wherein estimating parameters of a user-specific learning machine further comprises estimating parameters of a user-specific learning machine based at least in part on the documents not of interest to the user.	REFUAH, 6:51-53: "A parameter may also be negative, for example, a blacklisting: 'reject=pornography' or 'reject if pornography level >3" 22:6-14: "In a preferred embodiment of the invention, a personality are updated responsive to one or more of the types and/or contents of sites that a client accesses, the time spent at each site (preferably with a deduction for connection time), activities performed at the sites and/or data downloaded from the sites. In a preferred embodiment of the invention, a client can indicate to the persona server if he is pleased with a particular site and/or displeased. Such an indication may also be used to modify the personality." *See also*: **CULLISS** at 3:57-65, 4:61-64, 5:4-10. **MLADENIC** at 8. *See generally* 10-11. **SCHUETZE** at 17:47-67. **WASFI** at 58. Autonomy Press Release, at 1

	Autonomy Technology Whitepaper, at AUT00070 Autonomy Agentware User Guide, at AUT00123
Claim 6	
6. The method of claim **1**, wherein monitoring user interactions with data for a document comprises monitoring at least one type of data selected from the group consisting of information about the document, whether the user viewed the document, information about the user's interaction with the document, context information, the user's degree of interest in the document, time spent by the user viewing the document, whether the user followed at least one link contained in the document, and a number of links in the document followed by the user.	REFUAH, 5:34-50: "Additionally or alternatively, personality may be updated automatically. In a preferred embodiment of the invention, the mood is updated based on the one or more of the identification of sites visited by a user, the number of site visited, the dwell time at each site, the order in which sites are visited, the contents of the sites, services purchased, information downloaded, actions performed at the sites and/or a predefined or adaptive time-line based function. Alternatively or additionally, a mood, for example a "rush" mood, may be identified by tracking whether a user waits until images are downloaded, whether a user waits for a complete site to download, whether a user follows links and how many links are followed, and/or rate of changing WWW pages and/or sites. These tracked variables may be compared to a standard. Alternatively or additionally, the tracked variables may be compared to a previously acquired baseline of a user. Thus, relative changes in dwell time are tracked." 14:54-59: "In a preferred embodiment of the invention, the persona server tracks frequency, length and/or content of electronic communications with the friends, for example by e-mail, by chat group, by Internet telephone or by computer-dialer, to evaluate an instantaneous mood and/or to assess the relative effect of these friends." *See also*: **CULLISS** at 2:43-46, 3:27-35. **MLADENIC** at 8. **SCHUETZE** at 5:36-40, 11:12-14, 18:11-17. **WASFI** at Abstract, 60, 61.

Claim 7	
7. The method of claim **1**, wherein said plurality of retrieved documents correspond to a respective plurality of products.	REFUAH, 1:63 – 2:2: "One object of some preferred embodiments of the invention is to provide a method of aiding information search and retrieval on an Internet. In a preferred embodiment of the invention, Internet searching is personalized to a particular user's profile. Alternatively or additionally, matching up of a supplier and a buyer, of a goods and/or a service, is facilitated, based on such personalization." 3:56 – 4:4: "Another aspect of some preferred embodiments of the invention relates to personalizing advertisements responsive to a mood and/or a persona. This personalization of advertisements may be in addition to or alternatively to personalization responsive to a particular search and/or other actions performed by a user at a site. In a preferred embodiment of the invention, a site obtains information on a persona and/or a mood of the accessing user and then tailors services and/or advertisements based on the mood or persona. In a preferred embodiment of the invention, when a user enters a book-seller's web site, even if the user has never been at the book-seller, he may be offered books which match his persona and/or mood. It should be appreciated that, in some preferred embodiments of the invention, such a personality is not generally created and/or maintained by the site which uses the information for personalization." 7:24-32: "In a preferred embodiment of the invention, a mood or a persona may be provided by an outside entity. In one example, an advertise may provide a persona and/or mood tailored for a particular product or group of products. This type of persona could be configured to receive advertisement, promotions and/or search results geared towards the product. Typically, using such a mood may result in a rebate on purchases, Internet fees and/or may involve a promotion including the product, for example a free sample." 18:35-39: "When white and/or yellow pages are used, the display of information from a database may be determined by the personality, for example, the display of listings of baby-sitters, handymen and car garages may all depend on a geographical distance." 18:40-55: "In a preferred embodiment of the invention, the personality may be used to target advertisements to the client. Such targeting may also take into account previous advertisements

	viewed by the client. In one example, advertisements are matched to professed subjects of interest. Additionally or alternatively, advertisements are matched to an outlook, for example morbid or sunny. Alternatively or additionally, advertisements are selected from a set of suitable advertisements to match a persona. For example, if two soft drink advertisements are available, one which includes animals and one which includes cars, the "animal" advertisement will be selected for a persona which likes animals. Similarly, some advertisements are garish, while some are reasoned out. A somber mood will preferably be targeted with the reasoned out advertisement, since a garish advertisement might antagonize the client." *See also*: **CULLISS** at 9:55- 10:13. *See generally* 9:55 – 11:33. **MLADENIC** at 2, 8, Fig. 2. **SCHUETZE** at 35:66 – 36:8.
Claim 21	
21. The method of claim **1**, wherein using the estimated probabilities for the respective plurality of retrieved documents to present at least a portion of the retrieved documents to the user comprises presenting to the user at least said portion of the retrieved documents based on the estimated probability that the retrieved document is of interest to the user and the	17:49-65: "the presentation of search results may also be parameters of the persona…. [o]ne or more parameters of a persona may define matching requirements, for example exactness of match and allowed error." *See* citations for claim 1[g].

relevance of the retrieved document to the search query.	
Claim 22	
22. The method of claim **1**, wherein identifying properties of the retrieved document comprises identifying properties selected from the properties consisting of a topic associated with the retrieved document, at least one product feature extracted from the retrieved document, an author of the retrieved document, an age of the retrieved document, a list of documents linked to the retrieved document, a number of users who have accessed the retrieved document, and a number of users who have saved the retrieved document in a favorite document list.	REFUAH, 7:53 – 8:6: Another aspect of some preferred embodiments of the invention relates to evaluating an atmosphere and/or other traits of a site. In a preferred embodiment of the invention, depending on a persona, several characteristics of a site may be defined, which may be used in filtering out and/or prioritizing such a site. Alternatively or additionally, such information may be used database of sites with their associated values is maintained, so that such characteristics do not need to be re-evaluated very often. Another aspect of some preferred embodiments of the invention relates to associating traits and/or an atmosphere with a WWW site. The associations may be stored at a central location. Additionally or alternatively, the associations and/or trait-related keywords and/or values may be associated with each site. Additionally or alternatively, a site may include an identification number, which when used with a proper trait server, provides information about the sites traits and/or a match and/or grade with a particular personality. Additionally or alternatively, such associations may be stored in search indexes, preferably in a manner similar to the storage of key words." 9:50-59: "Preferably, analyzing a content, comprises determining at least one trait of said site. Alternatively or additionally, analyzing a content comprises determining an ambiance of said site. Alternatively or additionally, analyzing comprises analyzing lexicographical characteristics of said site. Alternatively or additionally, analyzing comprises analyzing graphical characteristics of said site. Alternatively or additionally, identifying at least one site comprises identifying a plurality of sites. Preferably, identifying comprises searching using an Internet search engine." 20:19-30: "An atmosphere may include a plurality of traits, for example, political slant, garishness, reading grade level, subjects of interest and in general the complements of parameters of a persona. The atmosphere of a site may be evaluated on the fly using various methods described below. Additionally or alternatively, a client may grade a site. Such grading may become publicly available or it may be limited to the client or a group of clients. Additionally or alternatively, the persona server or the atmosphere server may evaluate a site. Additionally or alternatively, a separate server may provide a site evaluation service."

	21:6-30: "In a preferred embodiment of the invention, an atmosphere of a site may be automatically evaluated by analyzing the content of a site, in addition to or instead of utilizing a client's reaction to the site or statistics of accessing the site. Various characteristics of a site may be automatically determined. Each of these characteristics and/or combinations thereof may be used to estimate values for traits and/or atmosphere. The characteristics preferably include one or more of: (a) word length; (b) whether certain words and/or phrases used by or associated with the site belong to certain groups, such as "academic words", "swear words", "adult words", "new-age words", "sports words", "baseball words"; (c) sentence complexity; (d) density of displayed text; (e) ratio between images and text; (f) size of text; (g) distribution of colors in image sand in background; (h) number of links; number of links visited, date of last visit, by the client, by the persona, by the mood and/or by other moods, personas and/or clients; (i) size of site; (j) key-words presented by the site; and/or (k) number of images; and/or (l) number and/or type of multimedia files." *See generally* 20:19- 21:36. *See also*: **CULLISS** at 2:26-37. **MLADENIC** at 3, 4, 12. *See generally* 3-6. **SCHUETZE** at 6:58 – 7:15, 10:40-56, Fig. 3. *See generally* 17:47 – 18:27. **WASFI** at 61.

EXHIBIT 3 G

Exhibit 3-G

Claim Chart of Joachims et al., "*WebWatcher: A Tour Guide for the World Wide Web*," (1997)

as prior art to

Asserted Claims of U.S. Patent No. 6,9981,040 ("'040 Patent")
and
Asserted Claims of U.S. Patent No. 7,685,276 ("'276 Patent")

'040 Patent	JOACHIMS
Claim 1	
A computer-implemented method for providing automatic, personalized information services to a user u, the method comprising:	WebWacher was a system developed under the umbrella of Carnegie-Mellon University's Text Learning Group. Users can begin by entering a phrase that describes their current interest, such as "intelligent agents." (Joachims at 2). WebWatcher henceforth accompanies users as they navigate web pages. "Each time the user selects a hyperlink, WebWatcher accompanies the user to the next page, and logs this hyperlink selection as a training example for learning to improve future advice." (*Id.*) That advice comes in the form of highlighting selected hyperlinks by inserting "eyeball" icons around the link. *See also* JOACHIMS at Abstract: "This paper describes a simple but operational tour guide, called Web-Watcher, which has given over 5000 tours to people browsing CMU's School of Computer Science Web pages. WebWatcher accompanies users from page to page, suggests appropriate hyperlinks, and learns from experience to improve its advice-giving skills." 2: "WebWatcher was in operation from August, 1995, to February, 1997."

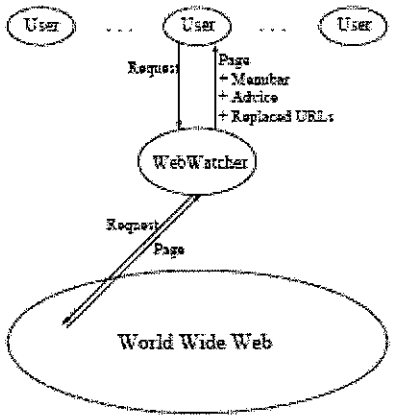

Figure 2: WebWatcher is an interface agent between the user and the World Wide Web.

See also:

CULLISS at 1:16-19; 1:38-40; 1:41-43.

MLADENIC, at 2, 8, Fig. 2.

REFUAH at Abstract; 1:63 – 2:2; 2:36-49.

SCHUETZE at 1:29-33.

WASFI at Abstract.

Autonomy Press Release, at 1-2

AUT00002, AUT00004, AUT00068, AUT00078

MONTEBELLO at Abstract, 1.

a) transparently monitoring user interactions with data while the user is engaged in normal use of a computer;	JOACHIMS at 1: "WebWatcher acts as a learning apprentice [Mitchell et al., 1994], observing and learning from its users' actions. Over time WebWatcher learns to acquire greater expertise for the parts of the World Wide Web that it has visited in the past, and for the types of topics in which previous visitors have had an interest." 2: "We can tell that WebWatcher accompanies us from the additions it makes to the original page." 2: "In general, the user may click on any hyperlink, recommended or not. Each time the user selects a hyperlink, WebWatcher accompanies the user to the next page, and logs this hyperlink selection as a training example for learning to improve future advice." 2: "WebWatcher accompanies the user along any hyperlink anywhere on the World Wide Web. To end the tour, the user clicks on one of two options in the command list: 'Exit: Goal reached' or 'Exit: Goal not found.' This exit provides the user with a way of giving final feedback to WebWatcher." *See also*: **CULLISS** at 3:46-56; 5:18-21; 7:14-20. **MLADENIC**, at 3. **REFUAH** at Abstract; 3:3-11; 5:34-50; 19:20-22. **SCHUETZE** at 5:36-40; 11:12-14; 18:11-17; 28:65 – 29:6. **WASFI** at Abstract, 60. Autonomy Press Release, at 2 Autonomy Technology Whitepaper, AUT00069-70

	Autonomy Agentware User Guide, at AUT00119 **MONTEBELLO** at 3.
b) updating user-specific data files, wherein the user-specific data files comprise the monitored user interactions with the data and a set of documents associated with the user;	JOACHIMS at 2: "In general, the user may click on any hyperlink, recommended or not. Each time the user selects a hyperlink, WebWatcher accompanies the user to the next page, and logs this hyperlink selection as a training example for learning to improve future advice." 3: "What is the form of the knowledge required by WebWatcher? In general, its task is to suggest an appropriate link given an interest and Web page. In other words, it requires knowledge of the following target function: *LinkQuality : Page x Interest x Link* → [0; 1] The value of LinkQuality is interpreted as the probability that a user will select Link given the current Page and Interest. In the following we present three approaches to learning this target function from experience. The first approach uses previously given tours as a source of information to augment the internal representation of each selected hyperlink. The second approach is based on reinforcement learning. The idea is to find tours through the Web so that the amount of relevant information encountered over the trajectory is maximized. The third approach is the combined method that includes both of the first two approaches." *See also*: **CULLISS** at 3:13-35; 5:36-48; 7:14-42. **MLADENIC** at 3, 8. **REFUAH** at Abstract; 5:34-50; 6:5-15; 8:31-39; 20:31-37. **SCHUETZE** at 10:14-18; 10:32-39; 11:12-17; 28:65 – 29:6; 34:34-37; *See generally* 17:47 – 18:27. **WASFI** at 58, 60, 61.

	Autonomy Press Release, at 1 Autonomy Technology Whitepaper, at AUT00069-70 **MONTEBELLO** at 3-4.
c) estimating parameters of a learning machine, wherein the parameters define a User Model specific to the user and wherein the parameters are estimated in part from the user-specific data files;	JOACHIMS at 3: "What is the form of the knowledge required by WebWatcher? In general, its task is to suggest an appropriate link given an interest and Web page. In other words, it requires knowledge of the following target function: *LinkQuality : Page x Interest x Link* → [0; 1] The value of LinkQuality is interpreted as the probability that a user will select Link given the current Page and Interest. In the following we present three approaches to learning this target function from experience. The first approach uses previously given tours as a source of information to augment the internal representation of each selected hyperlink. The second approach is based on reinforcement learning. The idea is to find tours through the Web so that the amount of relevant information encountered over the trajectory is maximized. The third approach is the combined method that includes both of the first two approaches." *See generally* p. 4. *See also*: **CULLISS** at 3:57-65; 4:54 – 5:10. **MLADENIC**, at 9, 10. **REFUAH** at 2:9-35; 6:49-64; 8:30-58; *See generally* 14:21 – 15:45. **SCHUETZE** at 27:44-64. *See also* 27:65 – 28:14. **WASFI** at 58, 61, 63.