MCALD 400-621 499 Figueira Decl. Tab 307

Redacted Pursuant to

Request of Defendants

Protective Order at

To:

"Matthew Rizzo" <mrizzo@youtube.com> "David Chandler" <chandler@google.com>

From: Cc:

"Cuong Do" <cdo@youtube.com>, "Jeremy Doig" <jeremydo@google.com>, "Joe

White white | @google.com>, "David Stoutamire" < dps@google.com>

Bcc:

Received Date.

2006-11-16 00:39:01 GMT

Subject:

Re: Dear YouTube Search Engine Keepers

Matthew and I spoke. To get to Milestone 1 it's perhaps just a matter of modifying the existing Python script that handles document additions/deletions/updates to use the google Python APIs that vsp_insert.py once used (when Google Video was using mysql + python) to tell my mustang index of updates. Eventually the youtube compython frontend must be modified to speak to our mustang index, but that's a later milestone and should be straightforward since there's a google python API for that part of things too. Yay SWIG.

I'm going to fly up to Mountain View/San Bruno Monday at noon and stay the rest of the 3-day work week. I'm not attached to actually getting anything done, but it will be nice to meet Matthew in person and I'll be available to anyone with questions if I can find a chair and network access. If not I'll hang out at the googleplex and enjoy gourmet food and gourmet network access.

-David

On 11/15/06, Matthew Rizzo <mrizzo@voutube.com> wrote:

- > "Related videos" on the right hand side of the flash player match one
- > or more of the keywords of the video you are watching. The videos in
- > the flash player are from a recommendation engine and does not come
- > from PyLucene. We served the watch page times on saturday
- > so that is about per second queries from just that. We cache
- > some of the results in memcache, but I think that is pretty much in
- > the ballpark.

- > Talk to you tomorrow,
- > Rizzo
- > On Nov 14, 2006, at 11:34 PM, David Chandler wrote:
- >> 14:00 works. When a video gets watched, I'm curious what query is
- >> generated. Is this PyLucene index responsible for the related videos
- > > (in the flash player at the end and outside of it)?
- > >
- >> David
- >>
- > > On 11/14/06, Matthew Rizzo <mrizzo@youtube.com> wrote:
- > >> We are maintaining 4 indexes; one each for video, user, playlists and
- >>> groups. Video is used most white the others get a comparatively
- > >> infinitesimal amount of traffic. I will focus on video since that is
- > >> our primary business and is a larger problem than the other 3 indexes.
- >>> combined. We currently get about queries per second on our
- >>> video search during peak times. We issue 3 different types of queries
- >>> and allow 5 ways to sort.
- >>>
- >>> We issue a search when:

```
* video gets watched
> >>
          * a user requests
> >>
          * the browse by categories page gets hit
>>>
          * a user is subscribed to a tag and hits the index or
> >>
>>> subscription
> >> center page
> >>
>>> We currently index
                                 videos (processed videos that are
>>> public) via PyLucene. We update the index per day, which
> >> constitutes a 25% change in the index. This is done via a batch tool
> >> that selects videos from a replica of our master database if a video
> >> has changed since the tool's last execution. This takes about
> >> hours to apply the changes to the index and hours to replicate the
> >> index to the search machines. The number of records in our video
> >> index
                   every four months and the index's size is
>>>
>>> The following are the fields used to render the results which are
>>> stored in the index:
> >>
> >>
> >>
> >>
> >>
> >>
> >>
> >>
> >>
> >>
> >>
> >>
> >>
> >> I will give you a call tomorrow at 14:00 PST if that works for
>>> mail me with an alternative time if this time is not convenient
> >> for you.
> >>
>>> On Nov 14, 2006, at 7:54 PM, David Chandler wrote:
>>> > How about a phone call (
                                                                ), when
> >> you've got
>>> > 20 minutes, to talk about the best way to get to Milestone 1? I'll
>>> > define Milestone 1 as a search index in Google's prod
>>> datacenters that
>>> has both google video and youtube docs (and the ability to restrict
>>> searches to just one or the other), all such docs as of today, say.
>>> This will be a testing index for a demo, but it will use all of
>>> google's link information to rank things pretty well.
>>>>
>>> The real issue, then, is to communicate where your data lives
> >> now so I
>>> can figure out how to pull it out where I need it.
>>> Any documentation you have would be useful; I'm especially
> >> interested
>>> > in what data you store per video -
                etc. -- and how often that's updated since the perfect
> >> >
>>> system for that may or may not be the system we've got right now.
```

Highly Confidential GO0001-00243150

```
>>> > Source code is useful, so as soon as you figure out how to share
>>> with other Googlers, I'd benefit from reading it.
> >> >
>>> > Also, how many search queries per second do you get at peak?
> >> How many
>>> videos are in your index? How fast is it growing?
>>>>
>>> > David
>>>>
>>> On 11/14/06, Matthew Rizzo <mrizzo@youtube.com> wrote:
>>> >> Hello all,
>>> >> How do you guys want to move forward on this? Let me know what
>>> >> information you need from me. I don't have intranet access yet I
>>> >> think that will come tomorrow or thursday.
> >> >>
>>> Feel free to use my google email if that is easier for you
>>> >> (mrizzo@google.com)
> >> >>
>>> >> Rizzo
>>> >> On Nov 14, 2006, at 3:31 PM, Cuong Do wrote:
> >> >>
> >> > David.
>>>>>>
>>> >> I'm CC'ing Matt Rizzo on this e-mail. He is the keeper of
>>> YouTube
>>> >> search and will likely be leading the effort to integrate into
>>>> Google Search from the engineer side.
> >> >> >
>>>>> Cuong
>>>>>
>>> >> On Nov 14, 2006, at 2:59 PM, David Chandler wrote:
>>>>>
> >> >> Hi Cuong,
> >> >> >>
>>> >> Jeremy tells me you wrangle youtube's backend. I'm the
>>> >> wrangler of
>>> >> Google Video's search index, and I'd appreciate it if you could
>>> >> put me
>>> >> in contact with the right engineers to talk about getting
>>> >> youtube's
>>> >> docs in our search index. I hope we can do a better job of
>>> >> ranking
>>> >> them than you currently do and hope to take some of the load
>>> >> off of
>>> >> your shoulders so you can focus on youtube's strengths.
> >> (Google
>>> >> Video
>>> >> and Google Search also both want to search over youtube's
> >> videos.)
>>>>>>
>>>>> Regards,
>>> >> David
> >> >>
>>>>> WO
>>> >> cel
>>>>>
> >> >>
```

Highly Confidential GO0001-00243151

>>>>>

> >> > >>

>

Highly Confidential G00001-00243152