

February 21, 2013

Six Months Later – A Report Card on Google's Demotion of Pirate Sites

"Starting next week, we will begin taking into account a new signal in our rankings: the number of valid copyright removal notices we receive for any given site. Sites with high numbers of removal notices may appear lower in our results. This ranking change should help users find legitimate, quality sources of content more easily—whether it's a song previewed on NPR's music website, a TV show on <u>Hulu</u> or new music streamed from <u>Spotify</u>." – Amit Singhal, Senior Vice President of Engineering, Google, August 10, 2012

Executive Summary:

On August 10, 2012, Google announced that it would take into account in its search result rankings the number of valid copyright removal notices it has received for a given site.¹ Per its announcement, "sites with high numbers of removal notices may appear lower" in its search results. The result of the change should be to "help users find legitimate, quality sources of content more easily." Six months later, we have found no evidence that Google's policy has had a demonstrable impact on demoting sites with large amounts of piracy. These sites consistently appear at the top of Google's search results for popular songs or artists. Specifically:

- Over the six-month period, Google received notices for tens of millions of copyright removal requests concerning various sites, including multiple repeat notices of infringement of the same content on the same site;
- The sites we analyzed, all of which were serial infringers per Google's Copyright Transparency Report, were not demoted in any significant way in the search results and still managed to appear on page 1 of the search results over 98% of the time in the searches conducted;
- In fact, these sites consistently showed up in 3 to 5 of the top 10 search results;
- This is of particular concern as studies have shown that approximately 94% of users do not go beyond page 1 results;
- For 88% of our searches for mp3s and downloads of popular tracks, Google's "auto-complete" function suggested appending to the searches certain terms which are associated with sites for which it has received multiple notices of infringement, thus leading to illegal content;
- Well-known, authorized download sites, such as iTunes, Amazon and eMusic, only appeared in the top ten results for a little more than half of the searches. This means that a site for which Google has received thousands of copyright removal requests was almost 8 times more likely to show up in a search result than an authorized music download site. In other words, whatever Google has done to its search algorithms to change the ranking of infringing sites, it doesn't appear to be working.

¹ "An update to our Search Algorithms," posted by Amit Singh on August 10, 2012, available at <u>http://insidesearch.blogspot.com/2012/08/an-update-to-our-search-algorithms.html</u>.

1. Introduction

Six months after Google announced the implementation of a demotion signal related to copyright removal requests, it appears that this signal does not result in a meaningful change in search rankings of noticed sites on the first page of search results when a user searches for an mp3 or download of popular songs. This is of particular concern, as studies show that the vast majority of users only click through a result on the first page of search results for any given search. For example, a study of click-through rates by Chitika Insights using a sample of over 8 million clicks showed 94% of users clicked on a first-page result and less than 6% actually clicked to the second page to select a result displayed there.² Studies also show that a significant percentage of users rely on Google search to find music.³

Below is our analysis of the data collected to measure the impact of Google's demotion policy. While the analysis covers just a snapshot of the searches for songs on Google, it suggests that Google has knowledge of allegations of rampant infringement on these sites, but nonetheless continues to direct users to these sites via top search result rankings and auto-complete suggestions.

2. Methodology

For this analysis we performed searches for [artist] [track] mp3 and [artist] [track] download over a period of several weeks starting December 3, 2012.⁴ The tracks selected were based on the top 50 tracks on the Billboard Hot 100 list as of December 3, 2012. The specific queries are listed in Appendix A. We have labeled these queries as the "Top 50 Track Search Queries." We further analyzed the results for just the top 10 track queries. These are listed on Appendix A and labeled as the "Top 10 Track Search Queries."

We also observed the number of copyright removal requests Google had received for various sites that showed up in those search results, and classified these sites into sites for which Google had received notices of (a) more than 1,000 copyright removal requests, (b) more than 10,000 copyright removal requests, and (c) for a sample of the sites, more than 100,000 copyright removal requests.⁵

Separately, we collected data on the top 5 search results for *free* [artist] mp3 and *free* [artist] download for a period of several months starting in March, 2012. The specific queries are listed on Appendix A and labeled as "Artist Searches."

Finally, we observed what suggestions Google made via its auto-complete feature when we typed in searches for the tracks noted on Appendix A.

² See <u>http://insights.chitika.com/2010/the-value-of-google-result-positioning/</u>.

³ See <u>IFPI Digital Music Report 2012</u>, p. 24 ("In fact, according to research done in the UK, 23 percent of consumers regularly download music illegally using Google as their means to find the content (Harris Interactive, September 2010). Further research in New Zealand highlights that 54 percent of users of unauthorized downloads said they found the music through a search engine (Ipsos MediaCT, October 2011).")

⁴ Google's auto-complete feature recommended appending "mp3" or "download" to an artist/track search 94% of the time for these artist/track combinations before the artist/track was fully typed into the search query.

⁵ As checked on <u>http://www.google.com/transparencyreport/removals/copyright/</u> on January 22, 2013 and January 23, 2013. We excluded from this count 2 sites for which Google had received over 10,000 notices but which have recently changed their practices to deter infringing activity on their sites.

3. Analysis

<u>3.1. Google has received a large number of copyright removal requests, including claimed serial infringement of sound recordings on various sites, as well as repeat notices of infringement of the same song on the same site.</u>

<u>3.1.1. Sites for which Google has received more than 1,000 or more than 10,000 copyright removal notices.</u> As noted above, we used the Google Copyright Transparency Report to identify those sites that consistently appeared in our search results for which Google had received notices of more than 1,000 or more than 10,000 instances of infringement⁶ as of January 22, 2013. There were more than 40 unique sites that showed up in the top 10 search results for the Top 50 Track Search Queries that satisfied these criteria. These sites are identified in Appendix B.

<u>3.1.2 Sample sites for which Google has received more than 100,000 copyright removal notices.</u> We further analyzed the number and class of infringements noticed on 8 sample sites for which Google had received notices of more than 100,000 instances of infringement. As shown on Table 1 below, Google has received notices of anywhere from 224,000 to more than 800,000 instances of infringement on these sites from a variety of copyright owners from the time Google started receiving notices for the site (typically sometime in 2011 or early 2012) through January 18, 2013.⁷ As further noted on Appendix B, Google had actually received more than 100,000 copyright removal requests for these sites by November 30, 2012.

Moreover, just for the period from August through December 2012, Google had received more than 100,000 copyright removal requests for all but one of these sites from RIAA alone.⁸ This period coincides with the period immediately after Google announced its demotion policy. This means that Google had received at least 100,000 copyright removal requests about sound recording infringements on these sites after the time it announced its demotion policy.

In addition, Table 1 shows the average number of repeat instances of infringement noticed to Google by RIAA for this period per site per track for the top 10 most popular tracks as reported by Billboard on December 3, 2012.⁹ These numbers suggest that, on average, RIAA provided notice of repeat infringement of the same track on the same site anywhere from once per week to once per day for the specified time frame.

⁶ We use "copyright removal requests" and "instances of infringement" interchangeably in this document.

⁷ Checked on <u>http://www.google.com/transparencyreport/removals/copyright/</u> on January 22, 2013. Per the transparency report, Google generally first received notices about these sites in 2011, and had continued to receive notices about these sites through January 2013, with a majority of the notices received in the second half of 2012.

⁸ RIAA's sister organization in the UK, BPI, has sent Google notices of more than 350,000 instances of infringement regarding Zippyshare.com since BPI started sending such notices. BPI has also sent Google notices concerning several of these and other sites as well.

⁹ The aggregate number is based on notices sent by RIAA during the specified period for the top 10 songs on the Billboard Hot 100 list as of December 3, 2012. The specific tracks used for this count are: "Diamonds" by Rihanna, "Die Young" by Ke\$ha, "One More Night" by Maroon 5, "Locked out of Heaven" by Bruno Mars, "Gangnam Style" by Psy, "Some Nights" by fun., "Ho Hey" by The Lumineers, "Home" by Phillip Phillips, "I Cry" by Flo Rida, and "Let Me Love You" by Ne-Yo. Some of these songs were not released until after August 2012. RIAA did not attempt to fully scan these sites for these tracks during this period.

Table 1 - # of removal requests received by Google for 8 sample sites

Notorious Site	# of Removal Requests All Time	# of Removal Requests Sent by RIAA Aug-Dec. 2012	Average # of Repeat Requests sent about the Same Track on the Same Site for the Top 10 Tracks AugDec. 2012
www.4shared.com	664,667	182,391	1.1 requests / week
www.audiko.com	224,123	135,686	3.2 requests / week
www.beemp3.com	578,641	137,587	1.0 requests / week
www.downloads.nl	822,670	433,342	7.2 requests / week
www.mp3chief.com	236,117	136,235	1.6 requests / week
www.mp3juices.com	311,320	139,199	2.2 requests / week
www.mp3skull.com	391,133	153,326	2.7 requests / week
www.zippyshare.com	614,117	69,816	4.4 requests / week

Thus, not only has Google received notices of hundreds of thousands of instances of claimed infringements on these sites generally, it has also received notices of more than 100,000 infringements of sound recordings on these sites after Google announced its demotion signal, and received several repeat notices of infringement of the same copyrighted works on the site.

3.2 These Sites Appear Frequently in the Top 10 Search Results for Searches for Popular Songs or Artists

Despite having received the notices of vast claimed infringements specified above, the sites in question continue to appear in the top 10 search ranking results for searches for an mp3 or download of popular tracks or popular artists. The following describes the incidence of sites appearing in the top 10 search results for the Top 50 Track Search Queries for which Google had previously received (a) more than 1,000 copyright removal requests, (b) more than 10,000 copyright removal requests, and (c) for a sample of sites, more than 100,000 copyright removal requests.

<u>3.2.1 Sites appearing in Top 10 search results for which Google had received notices of more than 1,000 instances of infringement.</u>

Figure 1 shows the breakdown by class of site that appeared in the top 10 search results for the Top 50 Track Search Queries conducted on a given day: January 25, 2013. As shown in Figure 1, on average:

- Nearly 50% of the top 10 search results (4.6 out of 10) were to sites for which Google had received more than 1,000 copyright removal requests by January 23, 2013;
- Conversely, well-known, authorized music download sites¹⁰ only appeared 0.6 times out of 10 in the top 10 search results. This means that, on average, an authorized download site appeared anywhere in the top 10 search results only in a little over half the searches.

¹⁰ We categorized Amazon, Apple/iTunes, and eMusic as well-known, authorized music download sites.

Figure 1 - Average number of times class of site appeared in top 10 search results for the Top 50 Track Search Queries on January 25, 2013



* The YouTube results often were to videos that included download links to unauthorized sites in the description of the video. We have not measured how many of these results contained such download links.

In other words, on average, for the Top 50 Track Search Queries, a site for which Google had received more than 1,000 copyright removal requests was almost 8 times more likely to appear in the top 10 search results than a well-known, authorized music download site.

<u>3.2.2 Sites appearing in Top 10 search results for which Google had received notices of more than 10,000</u> instances of infringement.

Figure 1 also shows the average number of times a site for which Google had received notices of more than 10,000 instances of infringement appeared in the top 10 search results. This class of site, on average, appeared in 3 to 4 of the top 10 search results. This class of site was 6 times more likely to appear in the top 10 search results than a well-known, authorized music download site.

In terms of frequency, 49 of the 50 searches conducted (98%) had at least 1 result in the top 10 search results to a site for which Google had received more than 10,000 copyright removal requests.

For the Top 50 Track Search Queries conducted on January 25, 2013, 98% of the time, Google returned search results that included in the top 10 rankings on average 3 to 4 sites for which Google had previously received more than 10,000 copyright removal requests.

3.2.3 Sample sites for which Google has received notices of over 100,000 instances of infringement.

Similar results appear when we analyze how often the sample sites noted above - *i.e.*, those for which Google had received more than 100,000 copyright removal requests by November 30, 2012 - appear in the top 10 results for the Top 50 Track Search Queries and the Top 10 Track Search Queries over an extended period of time.¹¹ Table 2 shows, by date of the search, the number of times these sites appeared in the top 10 search rankings on average for these queries.

¹¹ These sites are listed on Appendix B. As described on Appendix B, Google had received notices of >100,000 instances of infringement on these sites by at least November 30, 2012. Further, as noted above, in the last five months of 2012

Table 2 – Average # of times that a site for which Google has received more than 100,000 copyright removal requests appeared in the top 10 search results for the Top 50 Track Search Queries

	12/3/2012	12/15/2012 ¹²	12/20/2012	1/4/2013	1/11/2013	1/17/2013	1/25/2012
Top 10 tracks – Avg. # of times	1.9	2.1	1.8	2	2.1	2.2	1.8
Top 50 tracks – Avg. # of times	2.02	2.08	2.08	2.16	2.44	2.24	2.1

This means that, on average, 2 of the top 10 search results for each of the searches for an mp3 or download of the popular tracks were to sites for which Google had received hundreds of thousands of instances of infringement prior to the date of the search. This trend continued over the entire period measured.

Compare this to the number of times a well-known, authorized digital music download site appeared in the top 10 search results for the same searches, as shown in Table 3. Even the sample sites associated with more than 100,000 copyright removal requests were almost 4 times more likely to show up in a top 10 search result than a well-known, authorized digital music download site for these queries.

<u>Table 3 – # of times, on average, that amazon.com, apple.com or emusic.com appeared in top 10 search</u> <u>results</u>

	12/3/2012	12/15/2012	12/20/2012	1/4/2013	1/11/2013	1/17/2013	1/25/2013
Top 10 tracks – Avg. # of times	0.8	0.5	0.4	0.6	0.6	0.6	0.6
Top 50 tracks – Avg. # of times	0.8	0.6	0.3	0.6	0.6	0.6	0.6

In terms of frequency, for the searches conducted during the time period noted in Figure 2, the top 10 search results for 40 to 43 of 50 (80-86%) searches included a site for which Google had received more than 100,000 copyright removal requests. Figure 2 compares the frequency of this class of sites in the top 10 search results to the frequency of a well-known, authorized digital music download site in the top 10 search results. It also compares the frequency that one site for which Google has received more than 100,000 copyright removal requests, mp3skull.com, appears in the 1st ranked position, versus the frequency that any one of the well-known, authorized digital music download sites appears in the 1st ranked position.

alone, Google had received notices of >100,000 instances of infringement of sound recordings on each of these sites from RIAA and/or BPI.

¹² Three of the 50 searches were not conducted on this date.



Figure 2 - Frequency of certain sites appearing in top 10 search results for the Top 50 Track Search Queries

As shown above, one site associated with more than 100,000 instances of infringement, mp3skull.com, appeared more often in the 1st ranked position than all of the well-known, authorized digital music download sites in the top 10 search results.

Figure 3 breaks down, by site, the average percentage of times the identified sample sites appeared in the top ten search results over the Top 50 Track Search Queries.



Figure 3 – Average % of time a site for which Google had received more than 100,000 copyright removal requests appeared in the top 10 search results for [artist] [track] download or [artist] [track] mp3 for 50 popular songs¹³

Thus, despite Google having received notices of hundreds of thousands of claimed infringements on these sites, including many repeat infringements about the same copyrighted work on those sites, and including continued notices of infringement over the period in question, these sites continue to be ranked on the first page of search results when a user searches for [artist] [track] mp3 or [artist] [track] download.

¹³ Percentage reflects the sum of the number of times site showed up in the top 10 rankings for each of the 50 searches, divided by 50.

<u>3.3 There does not appear to be any meaningful decrease in ranking of these sites since the demotion signal</u> was implemented

As a means of testing whether or not there had been a significant decrease in the top search results of sites that engage in infringing activity after August 10, versus before, we looked at several more measurements. We ran a similar study for the six searches *free* [artist] mp3 or *free* [artist] download (3 of each) from March 23, 2012 through January 25, 2013, observing the top 5 search results during this period. This covers a significant period of time before the demotion signal went into effect around August 10-12, 2012, and a significant period of time thereafter. As noted below, this data suggests that sites for which Google received notices of high numbers of infringement were not demoted in any meaningful respect in the search rankings for this class of search.

Figure 4 shows the number of times any of the sample sites noted above appeared in the top 5 search results, averaged over the six searches described above. This data suggests that the sites in question – the sample of sites for which Google had received more than 100,000 instances of infringement – were not demoted in any meaningful respect for this class of search queries, but rather increased in frequency in the top 5 rankings after the demotion signal went into effect around August 10-12, 2012.



3.4 Google's Auto-Complete feature suggests terms associated with sites for which it has received multiple notices of infringement

This problem may be exacerbated by Google's auto-complete function, which suggests that users append terms associated with sites for which Google has received notices of multiple instances of infringement. We observed what Google recommended in auto-complete when typing in [artist] [track] mp3 or [artist] [track] download, using the top 50 tracks noted in Appendix A.¹⁴ In this case, **88% of the time (42/50), Google recommended appending to those searches terms associated with sites for which Google has received**

¹⁴ As observed on January 23, 2013.

notices of more than 1,000 instances of infringement. These terms included skull (associated with mp3skull.com), zippy (associated with zippyshare.com), sharebeast (associated with sharebeast.com), 4shared.com (associated with 4shared.com), mp3ye (associated with mp3ye.com), .nl (associated with downloads.nl), dopehood (associated with dopehood.com) and juices (associated with mp3juices.com).

Conclusion

Based on the data described herein, the search rankings for sites for which Google has received large numbers of instances of infringement do not appear to have been demoted by Google's demotion signal in any meaningful way, at least with respect to searches for downloads or mp3s of specific tracks or artists. More data should be collected and analyzed to determine whether this is a broader trend among searches aimed at creative content acquisition.

Appendix A

Search Queries

Top 50 Tracks Search Queries per Billboard Hot 100 List as of December 3, 2012

rihanna diamonds download
kesha die young mp3
maroon 5 one more night download
bruno mars locked out of heaven mp3
psy gangnam style download
fun. some nights mp3
lumineers ho hey download
phillip phillips home mp3
flo rida i cry download
ne-yo let me love you mp3
taylor swift we are never ever getting back together download
alex clare too close mp3
justin bieber beauty and a beat download
chris brown don't wake me up mp3
justin bieber as long as you love me download
kanye west clique mp3
florida georgia line cruise download
p!nk try mp3

miguel adorn download
kendrick lamar swimming pools mp3
alicia keys girl on fire download
nicki minaj va va voom mp3
ed sheeran the a team download
imagine dragons it's time mp3
cassadee pope over you download
neon trees everybody talks mp3
p!nk blow me download
swedish house mafia don't you worry child mp3
owl city carly rae jepsen good time download
pitbull don't stop the party mp3
juicy j bandz a make her dance download
adele skyfall mp3
ellie goulding lights download
carly rae jepsen call me maybe mp3
macklemore ryan lewis thrift shop download
mumford & sons i will wait mp3
frank ocean thinkin bout you download
One republic feel again mp3
luke bryan kiss tomorrow goodbye download
hunter hayes wanted mp3
11

flo rida whistle download
carrie underwood blown away mp3
gotye somebody that i used to know download
train 50 ways to say goodbye mp3
french montana pop that download
script hall of fame mp3
taylor swift i knew you were trouble download
david guetta titanium mp3
of monsters and men little talks download
lil wayne no worries mp3

Top 10 Track Search Queries Per Billboard Hot 100 List as of December 3, 2012

rihanna diamonds download
kesha die young mp3
maroon 5 one more night download
bruno mars locked out of heaven mp3
psy gangnam style download
fun. some nights mp3
lumineers ho hey download
phillip phillips home mp3
flo rida i cry download
ne-yo let me love you mp3

Artist Searches

Free Madonna MP3 Free Katy Perry MP3 Free Bruno Mars MP3 Free Rihanna Download Free Kelly Clarkson Download Free Lady Antebellum Download

Appendix B

>1K Sites, >10K Sites, Sample >100K Sites

Sites for which Google had received notices of more than 1,000 or more than 10,000 instances of infringement (copyright removal requests), as observed on the Google Transparency Report on January 23, 2013

sites	>1,000 copyright removal requests	>10,000 copyright removal requests		
2shared.com	y	y		
4shared.com	y	y		
airmp3.me	y	y		
audiko.net	y	y		
beemp3.com	y	y		
codemymp3.com	у	у		
dilandau.eu	у	у		
downloads.nl	у	у		
emp3world.com	у	у		
filestube.com	у	у		
freemp3box.com	y	y		
freemp3x.com	у	у		
isohunt.com	у	у		
kat.ph	y	y		
loudtronix.me	у	у		
mp3bear.com	у	у		
mp3chief.com	у	у		
mp3juices.com	у	у		
mp3lemon.org	у	у		
mp3searchy.com	у	у		
mp3skull.com	у	у		
mp3ye.eu	у	у		
musicaddict.com	у	у		
myfreemp3.eu	у	у		
prostopleer.com	у	у		
rlslog.net	у	у		
searchmp3.mobi	у	у		
torrentreactor.net	у	у		
viperial.com	у	у		
zippyshare.com	У	у		
aimini.net	У	n		
banashare.com	У	n		
filecrop.com	У	n		
mp3.li	У	n		
mp3oak.com	У	n		

sites	>1,000 copyright removal requests	>10,000 copyright removal requests
mp3olimp.com	у	n
sharebeast.com	у	n
soundowl.com	У	n
tumblr.com	у	n
zippytune.com	У	n

Sample of sites for which Google had received notices of more than 100,000 instances of infringement (copyright removal requests), as observed on the Google Transparency Report on November 30, 2012 and again on January 22, 2013

www.4shared.com
www.audiko.com
www.beemp3.com
www.downloads.nl
www.mp3chief.com*
www.mp3juices.com
www.mp3skull.com
www.zippyshare.com

* # of infringements on this site was not observed on the Google Transparency Report on November 30, 2012, but RIAA had sent Google notices of more than 100,000 instances of infringement on this site by November 30, 2012.