The Right to Obscure:
A Mechanism and Initial Evaluation

Eric H. Huang
Stanford University
ehhuang@stanford.edu

Jaron Lanier
Microsoft Research
jalani@microsoft.com

Yoav Shoham
Stanford University
shoham@stanford.edu

Abstract

The recent landmark "right to be forgotten" ruling by the EU Court gives EU citizens the right to remove certain links that are "inaccurate, inadequate, irrelevant or excessive" from search results under their names. While we agree with the spirit of the ruling—to empower individuals to manage their personal data while keeping a balance between such right and the freedom of expression, we believe that the ruling is impractical as it provides neither precise criteria for evaluating removal requests nor concrete guidelines for implementation. Consequently, Google's current implementation has several problems concerning scalability, objectivity, and responsiveness.

Instead of the right to be forgotten, we propose the right to obscure certain facts about oneself on search engines, and a simple mechanism which respects the spirit of the ruling by giving people more power to influence search results for queries on their names. Specifically, under our proposed mechanism, data subjects will be able to register *minus terms*, and search results for their name queries that contain such terms would be filtered out. We implement a proof-of-concept search engine following the proposed mechanism, and conduct experiments to explore the influences it might have on users' impressions on different data subjects.

1 Introduction

Search engines have revolutionized the way people access information. Not only have they become a source of information about any concept or thing, they are also often the first stop when people search for information about other people. Search results for a person's name have thus effectively become his/her *online profile*, which can have profound influences on one's life. In fact, employers often conduct online searches about job applicants, and many search for potential dates' names before going on the first date. Yet, despite the importance of search results for their names, people cannot easily influence their online representations, if at all.

Gaining control over one's online presence is indeed something people desire, so much so that a vibrant industry of online reputation management have sprung up to address this need. However, these services come at a high cost. For instance, some commercial solutions for managing personal reputation are priced between $3,000 and $15,000 USD [Reputation.com, 2015], which is unaffordable for most people.

The lack of a cost-effective way to manage one's online reputation was the core issue in the 2010 case that led to the recent "right to be forgotten" ruling by the European Court. The case was about Mario Costeja González, a Spaniard who did not want the auction notice of his repossessed home in a newspaper in 1998 to appear on Google's search results on his name, since this notice is no longer relevant as he claimed. In May, 2014, the EU Court ruled that individuals have the right to ask search engines to remove certain links with personal information about them if the information is "inaccurate, inadequate, irrelevant or excessive" for the purposes of the data processing [EC, 2014b]. Search engine operators should assess the requests on a case-by-case basis, taking into account the type of information, sensitivity to the requester's private life, interest of the public, as well as the requester's role in public life.

In order to comply with the ruling, Google has implemented a web submission form [Google, 2015] through which citizens in the EU can request links to be removed from searches on their names on EU versions of Google (e.g., google.co.uk), but not versions of Google outside of EU (e.g., google.com). The requesters are asked to provide a formal identification and reasons for the requests. Each request will then be assessed case by case by a team of people hired by Google, as called for by the Court. The same webform approach was also adopted by Microsoft's Bing search engine in Europe.

Problems

There are three main problems with the ruling and Google's implementation. First, the inherently vague ruling offers no clear guidelines for implementation and criteria for assessing requests: how exactly should Google decide what information is "inaccurate, inadequate, irrelevant or excessive"? As a result, the assessments are bound to be subjective. This problem is further compounded by the lack of information Google has about the request, which is provided by the re-

quester who might be biased to leave out unfavorable information to his/her case. Moreover, this lack of concrete guidelines makes verifying compliance difficult and impractical.

Second, the case-by-case assessment called for by the EU Court demands vast resources and is therefore not scalable. While large companies such as Google have the resources to hire teams of paralegals to process the requests, it is much less viable for smaller companies who are not nearly as resourceful. Thus, this ruling could lessen competition and stifle innovation.

Finally, Google's implementation responds slowly to new information on the Internet, which could surface at any time. If it takes days, weeks, or even months to process a new request, significant damage may have already been done to the data subject (the person whose information is indexed by the search engine). This problem is again exacerbated for less resourceful companies.

Our Contributions
Our goal in this work is not to implement the right to be forgotten ruling, due to the above-mentioned problems. Rather, we propose the right to obscure certain facts about oneself on search engines, which respects the spirit and intent of the ruling: "empowering individuals to manage their personal data while explicitly protecting the freedom of expression and of the media" [EC, 2014a], and "making sure that the people themselves—not algorithms—decide what information is available about them online when their name is entered in a search engine" [EC, 2014c]. We propose a simple mechanism that empowers individuals to have more control over their online presence on search engines via a query management approach. Specifically, a person would be able to remove certain links from the search results for his/her name by simply registering a set of *minus terms*, or keywords that they do not want to appear in the search results. For example, in the case of the ruling, Mario Costeja González could enter the term "auction" as a minus term that would always be appended to general queries on his name under the proposed scheme. When people search for "Mario Costeja González", the search results returned will be for the search string "Mario Costeja González -auction", effectively filtering out the undesirable personal information.

In order to evaluate the effectiveness of the proposed mechanism, we carry out initial experiments and demonstrate how the mechanism can help ordinary individuals protect their online reputation, and point out where it could fall short with regard to public interest.

2 Proposed Mechanism

We first introduce the design criteria that should be met by a practical mechanism that adheres to the spirit of the ruling. We then describe our mechanism, go through common use cases and address potential attacks. Finally, we comment on how the mechanism follows the ruling's spirit and the design criteria.

2.1 Design Criteria

Transparency. The mechanism should be easy to understand and not resort to blackbox algorithms.

Scalability. The mechanism should be able to handle a large number of requests with few resources so that it is implementable by small businesses.

Objectivity. The need for human oversight should be minimized. If human oversight is needed, decisions should be made as objectively as possible. For instance, verifying identity is objective, while deciding what information is excessive is relatively more subjective and disputable.

Responsiveness. The mechanism should be responsive to new links in the search results without requiring the subject to constantly monitor and submit new requests, especially if the new link is a duplicate of a link in previously submitted requests.

Public interest. The mechanism should keep the public interest at heart while empowering people to manage their personal data. It should not over-empower public figures and criminals; they should not be able to effectively hide information that may be of public interest.

2.2 Our Mechanism

At the high level, the mechanism allows users to register *minus terms* which would be used to filter out search results for queries on the requester's name. Specifically, when people search for the requester's name, any link containing the registered minus terms would be removed from the search results. We now describe in detail the protocol for the requesters and the search providers.

Requesters
Protocol The requester first provides his name and some form of legal identification. Upon successful ID verification, the requester can register a set of keywords as *minus terms*, which specify the keywords that should not appear in the search results. The requester can optionally register a set of *plus terms* which cannot be used as minus terms by other people with the same name. The purpose of plus terms is to mitigate conflicts of interest and to counter the demotion attack which will be described later in this section. In addition, we forbid registering frequent words in the language as minus terms, in order to counter the promotion attack discussed later in the section. The list of registered terms will be viewable only to those people with the same name whose IDs have been verified.

Implementation Minimally, there would be a simple webform that lets people register and manage their minus and plus terms. More convenient features could be added to aid requesters. For example, to help determining the right minus terms, there could be a panel showing the search results returned for a search on the requester's name. The requester can then select results which he/she wishes to remove. Then, the minus terms that would remove the selected results would be automatically generated and suggested to the requester.

Search Providers
Search providers will maintain a database whose entries consist of the requester's name and the associated registered terms.
Protocol
1. Given a search query, the search provider determines whether the query is a name query by matching it against names in its database.

2. A search string matches a name entry in the database if it matches at least the given name and the surname of an entry[1] (middle name is optional), and does not include additional non-matching terms.

3. The list of minus terms registered by the requester with a name match will be returned, which would only include minus terms that are not also registered as plus terms by any other person with the same name.

4. If there are multiple requesters whose name matches the search query, the union of all minus terms (that are also not plus terms) corresponding to all such requesters is returned.

5. A search is carried out with the search query. The links that include any of the minus terms should be excluded from the results.

6. If any link is removed from the results, the search provider should notify the searcher that there are one or more links removed.

Implementation Most search engines already implement an advanced search feature to exclude certain keywords in search results. Typically, this is done by adding a minus sign in front of the keyword one wishes to exclude, then appending it to the search query. Thus, our mechanism can be easily implemented by search providers by prepending each of the minus terms with a minus sign then appending them to the search query, then returning the search results of the modified query. To maintain low latency, the modified queries can be cached and only updated when new terms are registered.

2.3 Use Cases

Basic

John Alpha Smith would like a link about an expired auction of his house removed from the search results. In order to remove the link, he determines a keyword in the link, "auction", and requests that term be a minus term. We now describes what happens when given the following search strings:

"John Alpha Smith": As the search query consists solely of his name, it is determined as a name query. The returned results should be the normal search results for the query but with links containing the term "auction" removed.

"John Smith": This query would still match "John Alpha Smith" as matching the middle name is optional per the proposed protocol. Also, since this query does not uniquely identify John Alpha Smith, minus terms requested by other "John Smith"s should also be used. Thus, the returned results should not contain "auction" as well as any other minus terms requested by other "John Smith"s.

"John Alpha Smith auction": The query would match the requester's name. However, since the query also contains other keywords, the registered minus term should not be used as this is a specific query. Thus, unfiltered search results would be returned.

"John Alpha Smith job": Same as the previous case.

---

[1]Note that this step is modular and can be modified according to the naming conventions in different regions.

Competing Interests

Case 1. Jane Alpha Smith (JAS) would like Link A about her divorce removed, say a link written by his vengeful ex-husband. Jane Beta Smith (JBS) blogs about her divorce in Link B and would like search engines to display it.

JAS initially requests "divorce" to be a minus term. However, JBS notices that his blog has been removed from searches on her name, so she requests "divorce" as a plus term. In order to remove Link A, JAS finds a minus term that JBS would not object to, so it has to be one of the more specific terms that do not appear in links about JBS, such as her spouse's name, or other details about the divorce. As a result, those that search for "Jane Smith" or "Jane Alpha Smith" would not see results about JAS's divorce. This is sensible as potential employers or new friends would not learn about JAS's divorce if she doesn't want them to. She now has some influence over her private data online.

Case 2. Two people with exactly the same name get into a fight at Bar X which is reported in an article online. "John Smith"-1 would like the link removed so he requested "Bar X" to be a minus term, while "John Smith"-2 would like the link to be shown so he requested "Bar X" to be a plus term.

When people search for "John Smith", the link would not be filtered because "Bar X" has been requested as a plus term, even though it has also been requested as a minus term. According to our protocol, a term would not be returned as a minus term if someone has requested it to be a plus term.

In this example, it is as though the mechanism is not in effect. Note that we are implicitly favoring the public's right to information in this conflicting case. The reason for this decision is partly to prevent the potential attacks described below.

2.4 Potential Attacks

Demotion Attack

Sometimes an attacker may want to demote another person's presence in search results by removing links about that person. In fact, this attack is often used by competing businesses on Google's existing process for removing copyright infringing contents, and makes up over 50% of the removal requests [Urban, 2006].

Consider the following example. "John Smith"-1 is a physician who wants to promote his clinic, while "John Smith"-2 is a lawyer who wants to promote his firm. In order to promote their businesses, both of them are incentivized to request links about the other person to be removed, so that a general search on "John Smith" would return each one's own business, but not the other person's.

Our mechanism lets people register plus terms to counteract this attack. Specifically in this example, JS-1 may register "law" as a minus term, which would filter out JS-2's web page by itself. JS-2 notices this and sees that "law" is registered as a minus term (recall that minus terms are only viewable to people with the same name). JS-2 can then register "law" as a plus term, essentially canceling "law" as a minus term. Thus, JS-2's web page will not be filtered.

Note that registering a keyword as a plus term does not affect what would normally appear in the search results without filtering, nor the ordering of search results, which are de-

termined by the search provider. The option to register plus terms serves only as a way to nullify minus terms in case of the abuse of process.

Promotion Attack

Without a set of restricted minus terms, the proposed mechanism would allow people to promote their own content by filtering out organic search results. Concretely, someone wanting to promote his/her own content can simply register some of the most common words in the dictionary such as "the", "a", "it", etc., to filter out most of the links. Then, he/she would create a link with the content he/she wants to promote, without those chosen minus terms. This effectively puts only the links the abuser wants to show in the search results while removing the organic search results.

When the most frequently appearing words in a language cannot be registered as minus terms, the abuser would need to find a significantly larger set of keywords to filter out unwanted content. Also, since the mechanism only filters content for name-only queries, this attack would not have any effect on search queries with other keywords in addition to a name.

Revelation Attack

Under some circumstances, one may be interested to prevent another person from removing a link. For example, if the attacker created a link to spread a rumor about someone, he would like to keep the link in the search results. Under our proposed mechanism, it is possible to disallow someone from removing a link by registering plus terms, if the attacker has the same name or colludes with someone with the same name. However, observe that the end result of this attack is simply the same as without the mechanism. That is, the worst case of this attack is reverting to the status quo.

Metasearch Engine

One can imagine a metasearch engine that takes a name as input, makes thousands of queries with common minus terms, and aggregates results to highlight contents that were filtered. We note that this metasearch engine is also a search engine and should follow the mechanism if it is enforced as a law. Even if it does not follow, it still requires people to put in extra effort (going to another search engine they do not normally use) if they want to know more about the data subject, which is one of our design goals.

2.5 Other Concerns

Favoring People with Unique Names. Since people with the same name share the same set of minus and plus terms under the proposed mechanism, one property of the mechanism is that it implicitly favors those with uncommon names. However, people with the more common names are also more obscure and harder to search for, and are thus favored by search engines currently.

Security of Registered Terms. A direct consequence of implementing the proposed mechanism is that search engine providers would maintain a database of the registered terms, whose security is extremely important. People should be made aware of the possibility of data breaches of such databases. Thus, in most cases, the question people face under the proposed mechanism is whether they accept having unwanted links in their name searches, or delist them but risk having the terms going public if hacked. If they choose to register terms and the database is hacked, presumably the registered terms would correspond to links on the first few pages in the search results, which would always be exposed anyways if no terms were registered.

2.6 Discussion

Now that we have described in detail how the mechanism works and how it is robust to some potential abuses, we discuss how it adheres to the spirit of the ruling and the design criteria.

As demonstrated in the example use cases, the proposed mechanism empowers individuals to manage their personal data by giving them some control over the search results of general queries on their names. It is a simple way to let individuals remove unwanted content when people search for their names without resorting to expensive online reputation-management services. It is also an effective way to curtail online harassment; one would not be able to slut-shame, spread rumors, and post revenge porn as easily.

However, our mechanism does not give individuals complete control over search results under their names as the mechanism only delists links from the search results under certain name searches, but does not delete contents from webmasters. Moreover, the mechanism only applies to search queries consist solely of one's name, and does not apply to queries with keywords other than a name. This strikes a good balance between giving individuals some control over their online presence, and the freedom of expression and of the media.

Under the proposed mechanism, the filtered information is still on the Internet, but no one would be able to start with a person's name and learn all about him/her without giving that person the ability to influence how they appear online. Searchers would have to put in more effort and come up with relevant keywords to find what they are looking for, more similar to the pre-Internet times.

The proposed mechanism also scores well on the design criteria. It is *transparent* as it is straightforward and simple. It is *scalable* as it requires minimal human oversight and can automatically process requests upon identity verification. It is *objective* as verifying identity is uncontroversial, and minus term registration is automatically approved. This design is *responsive* as minus terms would filter out old and new matching content, thus relieving data subjects of the duty to constantly monitor their online persona. Finally, *public interest* is not compromised as searchers are still able to reach all the information by putting in more effort.

3 Experiments

In order to better understand how the mechanism might actually affect searchers' impressions on the data subjects, as well as how this effect is related to the searchers' agenda, we carry out experiments to explore the following research questions:

RQ1. Does the mechanism indeed give people more control over their online presence on search engines? How does

the mechanism's effect differ for well-known public figures versus ordinary individuals?

RQ2. How does the mechanism affect the accessibility of information?

RQ3. How does the searcher's search agenda play a role?

3.1 Setup

Overview

We implemented the proposed mechanism in a custom search engine using a search service API provided by Microsoft Bing. We recruit searchers from an online crowdsourcing platform, and ask searchers about their impressions about a given data subject before conducting any searches. We then ask them to research the person with a specified agenda using our custom-built search engine. Finally, we ask them again about their impressions about the person after they've conducted searches on our search engine.

Task

Upon accepting our task, workers are directed to a web page where they are provided with an overview of the task at the top. Before they conduct any searches, workers are asked how much they know about the data subject, on a scale of 1 to 5, as well as what their impressions are on the data subject({Strongly negative, Negative, Neutral, Positive, Strongly positive}). Then we present a brief description of the search engine with our mechanism, as well as an example. Workers do not know if there are any minus terms registered for the data subject in their task. They are instructed to research the data subject using our search engine, then answer the same questions about their knowledge and impression on the data subject after they finished researching. All instructions and questions are presented on one single web page in the order described.

Data Subjects and Search Agenda

We selected four data subjects, consisted of both public figures and ordinary individuals covering different real-world scenarios. To understand how the agenda of searchers play a role, we chose several agenda for each of the selected data subjects.

1. Lindsay Lohan. An American entertainer who has had troubles with alcohol and drug addictions, chosen to represent a fairly well-known celebrity who may be interested in keeping some personal information from the public. Agenda: find out if she was involved in any scandal.

2. Darleen A. Druyun. A former U.S. Air Force civilian official who was involved in a corruption scandal, chosen to represent a not so well-known public figure who may want to hide the information hurting her image. Agenda: find out more information about this candidate for a political position.

3. Jane Doe (generic name used). A marketing representative who was reported by another individual on shesa-homewrecker.com for breaking up a couple. She is chosen to represent an ordinary person whose online reputation is hurt by a damaging link prominently displayed on

the first page of name search results. Agenda: find out whether she could be an employee at a marketing firm.

4. John Doe (generic name used). A registered sex offender, chosen to represent an obscure person with a criminal history which is shown as his top name search results. Agenda A: find out more information about this new friend you just met. Agenda B: find out more information in order to decide whether to invite this new neighbor to your house-warming party. Agenda C: find out if this person has committed any crime in the past.

Minus Terms

We manually went through each data subject's name search results and registered minus terms in order to filter out results that are negative to his/her image, according to our best judgment.

Treatments

For each data subject-agenda pair, there are two treatments: with and without filtering. For each data subject we also include a treatment with no specified agenda. Thus, a total of 20 treatments were carried out. Each treatment is completed by approximately 30 participants recruited on Amazon Mechanical Turk, and pays $0.15 USD. To ensure work quality, we imposed standard qualification requirements of at least 95% approval rate on over 100 tasks. We also ensured that no participant researched the same data subject more than once.

3.2 Results

Considering the ordinal nature of the Likert scale items, the main metric we consider is the direction of change of the searcher's impression after conducting research on the search engine: positive, negative, or unchanged. This captures the influence of the new information from the search engine on the searchers' impressions of the data subject. We use the Mann-Whitney U tests for testing statistical significance and report the p-values. Figure 1 summarizes the results. We now explore our research questions in light of the experimental data.

RQ1. We found that the mechanism does allow less well-known individuals to effectively influence their online presence. However, it has relatively limited effects on well-known public figures like Lindsay Lohan. This is sensible as people are less likely to learn new information from the search engine if they already know a lot about the subject, thus they are less affected by the search results and the filtering mechanism. Indeed, the median of searcher's prior knowledge on Lindsay Lohan is 3 on a scale of 5, whereas for all other data subjects we selected, the median is 1 (no prior knowledge). Moreover, search engines are often not the only source of information about public figures as people are exposed to information about public figures via many other channels, such as TV news, newspapers/websites, etc. Thus, the mechanism's effect is less pronounced on public figures than ordinary individuals.

RQ2. People looking for the filtered information can still discover it with the proposed mechanism in place. As shown in John Doe and Lindsay Lohan's cases, when given the agenda to specifically look for the hidden information,
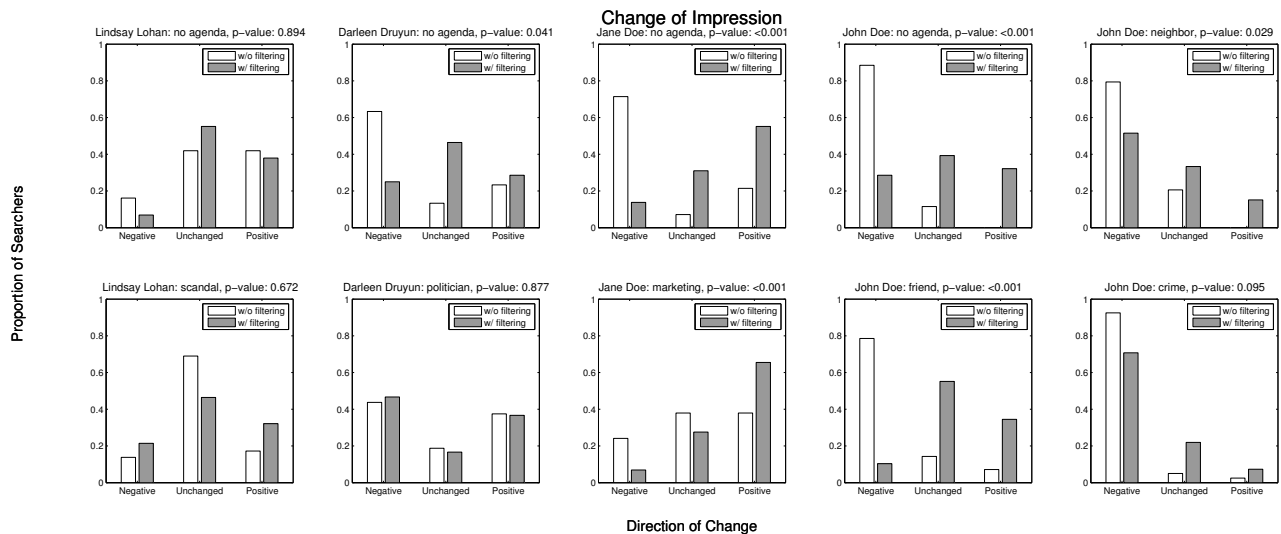
Figure 1: Distribution of the direction of change of searchers' impressions on data subjects, across all agenda. White bars are data with filtering, while darker bars are data without filtering. *p*-values are also reported.

searchers are still able to find it even with filtering. As designed, information is not deleted nor "forgotten" under the proposed mechanism. People are simply given more influence over their online persona when other people search for their names.

RQ3. As the data shows, the distribution of the directions of change of searchers' views on subjects can differ dramatically across different agenda. Therefore, the agenda of the searcher is an important factor when evaluating different removal mechanisms for search engines.

We observe that filtering's effect is most prominent when the filtered information is least relevant relative to the searcher's agenda. For example, in John Doe's case, we can say that a sex offender's record is least relevant when the searcher has no search agenda, more relevant when researching a new friend, even more relevant when deciding to invite some neighbor to a party, and most relevant when researching someone's criminal history. As the filtered content gets more relevant to the agenda, the more likely people are to find the filtered content, which is desirable. However, in the case of criminals, it may still be concerning and dangerous if some people missed that information. The proposed mechanism can easily handle this by not allowing criminals to register minus terms, which can be implemented at the identity verification step by looking up the requester's criminal record.

4 Related Work

Due to the recency of the right to be forgotten ruling, there has not been much academic work written on it. Zimmer [Zimmer, 2008] explores the consequences of the efforts of search engines to aggregate all the increasing amount of personal information online driven by the Web 2.0 infrastructures, and outlines possible spaces for intervention. Unlike our work, the author did not propose a concrete mechanism.

User intention has been considered by many in the in-

formation retrieval literature to be valuable for improving search relevancy and personalization [Jansen *et al.*, 2007; Rose and Levinson, 2004; Liu *et al.*, 2002]. To the best of our knowledge, our work is the first to study the relationship between user's search agenda and the resulting change in impression on the search subject. We note that user's search agenda is an important factor to consider when evaluating mechanisms that balance data subject's right to obscure and the freedom of expression and of the media.

5 Conclusion

We proposed a simple mechanism that empowers people to manage their personal data on search engines while still allowing for freedom of expression and of the media. The mechanism takes a query-management approach that is scalable and easy to use for data subjects. Moreover, unlike Google's implementation, it can easily be extended to domain-specific search engines, such as search on social websites, online forums, etc. We conducted experiments which suggest our mechanism does protect people's online persona, while not over-empowering public figures. While we believe that our approach finds a better balance and is cheaper, more responsive, and easier to implement than link removal, the mechanism must be tested extensively for situations where it may fall short and look for other potential abuses of the process, just like any such systems involving several parties with differing interests.

6 Acknowledgments

References

[EC, 2014a] EC. Factsheet on the "right to be forgotten" ruling. http://ec.europa.eu/justice/data-protection/files/factsheets/factsheet_data_protection_en.pdf, 2014. Accessed: 2014-11-02.

[EC, 2014b] EC. Judgment of the court in case c-131/12. http://curia.europa.eu/juris/document/\\document.jsf;?docid=152065\&doclang=EN, 2014. Accessed: 2014-11-10.

[EC, 2014c] EC. Mythbuster: The Court of Justice of the EU and the "right to be forgotten". http://ec.europa.eu/justice/data-protection/files/factsheets/factsheet_rtbf_mythbusting_en.pdf, 2014. Accessed: 2014-11-02.

[Google, 2015] Google. Legal help. https://support.google.com/legal/contact/lr_eudpa?product=websearch, 2015. Accessed: 2015-04-30.

[Jansen *et al.*, 2007] Bernard J. Jansen, Danielle L. Booth, and Amanda Spink. Determining the user intent of web search engine queries. In *Proceedings of the 16th International Conference on World Wide Web*, WWW '07, pages 1149–1150, New York, NY, USA, 2007. ACM.

[Liu *et al.*, 2002] Fang Liu, Clement Yu, and Weiyi Meng. Personalized web search by mapping user queries to categories. In *Proceedings of the Eleventh International Conference on Information and Knowledge Management*, CIKM '02, pages 558–565, New York, NY, USA, 2002. ACM.

[Reputation.com, 2015] Reputation.com. Fix negative search results — reputation.com. http://www.reputation.com/reputationdefender, 2015. Accessed: 2015-04-30.

[Rose and Levinson, 2004] Daniel E. Rose and Danny Levinson. Understanding user goals in web search. In *Proceedings of the 13th International Conference on World Wide Web*, WWW '04, pages 13–19, New York, NY, USA, 2004. ACM.

[Urban, 2006] Jennifer Urban. Efficient process or "chilling effects"? takedown notices under section 512 of the digital millennium copyright act. http://www.law.berkeley.edu/files/bclt_Urban_Efficient-Process_or_Chilling_Effects.pdf, 2006. Accessed: 2014-11-10.

[Zimmer, 2008] Michael Zimmer. The externalities of search 2.0: The emerging privacy threats when the drive for the perfect search engine meets web 2.0. http://www.firstmonday.dk/ojs/index.php/fm/article/view/2136/1944, 2008. Accessed: 2014-11-09.