
Theo Bertram*, Elie Bursztein, Stephanie Caro, Hubert Chao, Rutledge Chin Feman, Peter Fleischer, Albin Gustafsson, Jess Hemerly, Chris Hibbert, Luca Invernizzi, Lanah Kammourieh Donnelly, Jason Ketover, Jay Laefer, Paul Nicholas, Yuan Niu, Harjinder Obhi, David Price, Andrew Strait, Kurt Thomas, and Al Verney

Three years of the Right to be Forgotten

Abstract: The “Right to be Forgotten” is a privacy ruling that enables Europeans to delist certain URLs appearing in search results related to their name. In order to illuminate the effect this ruling has on information access, we conduct a retrospective measurement study of 2.4 million URLs that were requested for delisting from Google Search over the last three and a half years. We analyze the countries and anonymized parties generating the largest volume of requests; the news, government, social media, and directory sites most frequently targeted for delisting; and the prevalence of extraterritorial requests. Our results dramatically increase transparency around the Right to be Forgotten and reveal the complexity of weighing personal privacy against public interest when resolving multi-party privacy conflicts that occur across the Internet.

1 Introduction

The “Right to be Forgotten” (RTBF) is a landmark European ruling governing the delisting of information from search results. It establishes a right to privacy, whereby individuals can request that search engines such as Google, Bing, and Yahoo delist URLs from across the Internet that contain “inaccurate, inadequate, irrelevant or excessive” information surfaced by queries containing the name of the requester [7]. Critically, the ruling requires that search engine operators make the determination for whether an individual’s right to privacy outweighs the public’s right to access lawful information when delisting URLs.

After the RTBF came into effect in May 2014, the broad applicability of the ruling raised significant questions about how search engine operators weighed and ultimately resolved delisting requests [23]. Additionally, while Google and Bing both publicly report the volume of RTBF requests [13, 21], there has been a demand for more public data on the types of information that search engines delist as well as on the entities seeking

RTBF delistings, as exemplified by a letter published from 80 academics [19]. Any such transparency requires striking a balance that also respects the privacy of the individuals involved.

In this work, we address the need for greater transparency by shedding light on how Europeans use the RTBF in practice. Our measurements study covers over three years of RTBF delisting requests to Google Search, totaling nearly 2.4 million URLs. From this dataset, we provide a detailed analysis of the countries and anonymized individuals generating the largest volume of requests; the news, government, social media, and directory sites most frequently targeted for delisting; and the prevalence of extraterritorial requests that cross regional and international boundaries.

We frame the key findings of our analysis as follows:

There are two dominant intents for RTBF delisting requests: 33% of requested URLs related to social media and directory services that contained *personal information*, while 20% of URLs related to news outlets and government websites that in a majority of cases covered a requester’s *legal history*. The remaining 47% of requested URLs covered a broad diversity of content on the Internet.

Variations in regional attitudes to privacy, local laws, and media norms influence the URLs requested for delisting: individuals from France and Germany frequently requested to delist social media and directory pages, while requesters from Italy and the United Kingdom were 3x more likely to target news sites.

Requests carry a local intent: over 77% of requests to delist URLs rooted in a country code top-level domain (e.g., peoplecheck.de) came from requesters in the same country; all of the top 25 news outlets targeted for delisting received a minimum of 86% of requested URLs from requesters in the same country.

Only 43% of URLs meet the criteria for delisting: delisting rates ranged from 23–52% by country, and 3–100% based on the category of personal infor-

*All authors are affiliated with Google.

mation referenced by a URL (e.g., political platforms, personal information).

Requests skew toward a small number of countries and individuals: France, Germany, and the United Kingdom generated 51% of URL delisting requests. Similarly, just 1,000 requesters (0.25% of individuals filing RTBF requests) requested 15% of all URLs. Many of these frequent requesters were law firms and reputation management services.

Requests predominantly come from private individuals: 85% of requested URLs came from private individuals, while minors made up 5% of requesters. In the last two years, non-government public figures such as celebrities requested the delisting of 41,213 URLs and politicians and government officials another 33,937 URLs.

2 The Right to be Forgotten

Before diving into our analysis, we provide a short background on the RTBF and how it compares to other privacy controls. We also discuss the process for how Google arrives at delisting verdicts, how Google enforces delistings, and recent rulings that recognize a RTBF in other countries.

2.1 Origin

In May 2014, the Court of Justice of the European Union established a RTBF [1]. It allows Europeans to request that search engines delist links present in search results containing an individual’s name, if the individual’s right to privacy outweighs public interest in those results. The delisted information must be “inaccurate, inadequate, irrelevant or no longer relevant, or excessive in relation to those purposes and in the light of the time that has elapsed.” The ruling requires that search engine operators conduct this balancing test and arrive at a verdict.

In the wake of the ruling, Google formed an advisory council drawn from academic scholars, media producers, data protection authorities, civil society, and technologists to establish decision criteria for particularly challenging delisting requests [8]. Google also added a transparency report to reveal the volume of requests and domains targeted for delisting [13]. Transparency around this process is challenging, as the requester’s identity

cannot be disclosed. This has led to the development of a set of proposed best practices from Data Protection Authorities for handling delisting requests [4]. In parallel, researchers have examined de-anonymization risks surrounding the RTBF [29].

Compared to other privacy techniques such as access control policies in social networks or anonymization services, the RTBF is a unique in how it addresses multi-party privacy conflicts. Under the RTBF, access to lawfully published, truthful information about a person may be restricted via search engines (e.g., a criminal conviction, or a photo). That older information is more likely to be access restricted than newer information is perhaps an important issue for both civil society and future generations.

2.2 Decision process

Individuals located in European Union countries, as well as Iceland, Liechtenstein, Norway, and Switzerland are eligible to submit a RTBF request, and can do so via Google’s online delisting form [12]. As part of this form, requesters must both verify their identity by submitting a document (a government-issued ID is not required) and provide a list of URLs they would like to delist, along with the search queries leading to these URLs and a short comment about how the URLs relate to the requester. Requesters also choose a country associated with the request, which is often their country of residence.

Google assigns every request to at least one or more reviewers for manual review. There is no automation in the decision making process. In broad terms, the reviewers consider four criteria that weigh public interest versus the requester’s personal privacy:

1. The validity of the request, both in terms of actionability (e.g., the request specifies exact URLs for delisting) and the requester’s connection to an EU/EEA country.
2. The identity of the requester, both to prevent spoofing or other abusive requests, and to assess whether the requester is a minor, politician, professional, or public figure. For example, if the requester is a public figure, there may be heightened public interest surrounding the content compared to a private individual.
3. The content referenced by the URL. For example, information related to a requester’s business may be of public interest for potential customers. Similarly, content related to a violent crime may be of inter-

est to the general public. Other dimensions of this consideration include the sensitive, private nature of the content and the degree to which the requester consented to the information being made public.

4. The source of the information, be it a government site or news site, or a blog or forum. In the case of government pages, access to a URL may reflect a decision by the government to inform society on a particular matter of public interest.

In assessing each of these criteria, reviewers may seek additional details from the requester. When applicable, reviewers will also take into account recommendation from Data Protection Authorities. Ultimately, a decision is made to delist the URL from Google Search or reject the request.

2.3 Delisting visibility

Delistings occur on result pages for queries containing a requester’s name on (1) Google’s European country search services¹; and (2) on all country search services, including *google.com*, for queries performed from geolocations that match the requestor’s country [10]. However, in 2015 the French data privacy regulator CNIL notified Google to extend the scope of delisted URLs globally, not just within Europe [9]. Google appealed this decision, and the matter is now under consideration by the Court of Justice of the European Union [2, 24].

2.4 Similar rulings in other countries

While this study focuses solely on the European RTBF, for completeness we note that other countries have adopted similar rulings. In July, 2015 Russia passed a law that allows citizens to delist links from Russian search engines that “violates Russian laws or if the information is false or has become obsolete” [26]. Turkey established its own version of the RTBF in October, 2016 [22]. As our subsequent findings identify large variations in the privacy attitudes of different countries, it is unlikely our results generalize to these other RTBF laws.

¹ At the time of the study, European country search services were determined by ccTLD, e.g., *google.fr*, *google.no*, etc.

Dataset	Summary
Requested URLs	2,367,380
Unique requesters	399,779
Unique hostnames present in requested URLs	395,907
Time frame	May 30, 2014–Dec 31, 2017

Table 1. Summary RTBF requests targeting Google Search included in our analysis.

3 Dataset

Our dataset includes all European RTBF requests filed with Google Search from May 30, 2014 (when Google first implemented a delisting mechanism) to December 31, 2017. We provide a high level summary of this data in Table 1. Each record consists of two parts: (1) basic information related to URLs requested for delisting; and (2) additional annotations added by reviewers in the course of arriving at a verdict, described here.

3.1 Basic request data

Every request consists of the requester’s self-reported name, email address, country of residence, the date of the request, and the URLs requested for delisting, which we refer to from here on out as the *requested URLs*. Over the three year and a seven month period covered by our dataset, Google processed a total of 2,367,380 requested URLs submitted by 399,779 requesters². Additionally, each request includes a verdict per URL (i.e., delisted, rejected) and the date reviewers reached that verdict.

3.2 Annotations

Starting in January 22, 2016, Google reviewers began manually annotating each requested URL with additional categorical data for improving transparency around RTBF requests.

Category of site: The first annotation consists of a general purpose category that captures common strains of delisting requests related to social media, directory services, news, and government pages as shown in Table 2. These categories are not meant to be exhaustive.

² We omit any requested URLs that are still pending review at the end of our collection window. Over time, this may introduce discrepancies between metrics from Google’s Transparency Report and our report.

Label	Description
Directory	The URL relates to a directory or aggregator of information such as postal addresses or phone numbers for businesses or individuals. Examples include <i>118712.fr</i> which acts as a French business yellow page, or <i>192.com</i> which aggregates information about UK persons and businesses.
News	The URL belongs to a non-government media outlet or tabloid. Examples include <i>repubblica.it</i> an Italian newspaper, or <i>tv3.lt</i> a television station in Lithuania.
Social media	The URL references an account profile, photo, comment, or other content hosted on an online social network or digital forum. Examples include <i>facebook.com</i> and <i>vk.com</i> .
Government	The URL references a regional government’s legal or business records, or an official media outlet. Examples include <i>boe.es</i> , a bulletin for the Spanish government; and <i>thegazette.co.uk</i> , an official news outlet for the British government.

Table 2. Manually applied disjoint labels for different categories of sites. We denote URLs outside of these categories as Miscellaneous.

Rather, they enable tracking trends within categories and divergent interests between countries.

In total, 55.1% of requested URLs from 2016 onward fell into one of these categories. Reviewers labeled the other 44.9% as Miscellaneous. The large volume of miscellaneous URLs stems directly from the inherent diversity of content on the Internet. Examples of sites in this category include archival sites like *archive.is*, forums and blogs like *indymedia.org*, and articles or discussions on *wikipedia.org*.

Content on page: In the process of reviewing a URL to determine whether it meets the criteria for delisting, reviewers will categorize the information found on a page into a high-level description that captures the intent of the delisting as shown in Table 3. Most of these categories relate to varying degrees of potentially private information. Two categories—self-authored and name not found—relate to how information was authored or whether content is anonymized or does not reference the the requester by name. In total, reviewers determined a content category for 78.6% of requested URLs. The remaining 21.4% did not fall into one of these categories and were denoted Miscellaneous.

Requesting entity: The final annotation identifies special categories of individuals. These categories capture the degree to which information may be judged to have more of or less of a public interest, or the eligibility

of the requester. We provide a category breakdown in Table 4. In total, reviewers labeled 100% of requesters as belonging to one of these categories.

3.3 Ethics & Reproducibility

Our analysis broadly expands on the information currently available in Google’s Transparency Report [14]. Following the transparency report’s same standard, in the course of our analysis we never discuss details that might de-anonymize requesters or draw attention to specific URL content that was delisted. This emphasis on privacy creates a fundamental tension with standards for reproducible science. We cannot directly reveal a sample mapping between URLs and our annotations, nor can we reveal the exact URLs requested and the justification for delisting verdicts. We recognize these limitations upfront—but nevertheless argue it is critical to increase transparency surrounding how Europeans apply the RTBF and how it influences access to information on the Internet.

4 Delisting Requests

To start, we examine how the public’s use of the RTBF as a privacy tool has evolved over time. We explore the top countries generating requests, the impact of high-volume requesters on overall requests and delistings, and the number of corporations and government officials filing requests.

4.1 Timeline of requests

In order to understand whether RTBF requests are increasing in frequency, we calculate the number of URLs requested every month in Figure 1. During the first two months after RTBF enforcement began in May, 2014, requesters sought to delist roughly 282,000 URLs. From 2015 onward, RTBF requests fell to an average of approximately 45,000 URLs requested each month. Broken down by year, 39.7% of URLs were requested in the first year of enforcement, 24.9% in the second year, 22.0% in the third year, and the remaining 13.4% in the last seven months. We note this metric includes all requests, even those that reviewers ultimately declined to delist. In aggregate, only 43.3% of requested URLs met the criteria for delisting.

As an alternate metric of usage, we examine the arrival rate of new requesters as shown in Figure 2. After

Label	Description
Personal information	The requester’s personal address, residence, and contact information or images and videos.
Sensitive personal information	The requester’s medical status, sexual orientation, creed, ethnicity, or political affiliation.
Professional information	A requester’s work address, contact information, or neutral stories about their business activities.
Professional wrongdoing	References to the requester’s convictions of a crime, acquittals, or exonerations in a professional role.
Crime	References to the requester’s convictions of a crime, acquittals, or exonerations.
Political	Criticism of a requester’s political or government activities, or information about their platform.
Self authored	Requester authored the content.
Name not found	No reference to the requester’s name found in the content of the URL, though their name may appear in the URL parameters.

Table 3. Manually applied disjoint labels for the potentially private information appearing on a page requested for delisting. We denote content not falling into one of these categories as Miscellaneous.

Label	Description
Private individual	Default label for requesters not falling into a special category.
Minor	Requester is under the age of 18.
Government official, politician	Requester is a current or former government official or politician.
Corporate entity	Requester is filing on the behalf of a business or corporation.
Deceased person	Requester is filing on the behalf of a deceased person.
Non-governmental public figure	Requester known at an international level (e.g., a famous actor or actress), or has a significant role in public life within a specific region or area (e.g., a famous academic well known in their field).

Table 4. Manually applied disjoint labels for identifying categories of requesting entities.

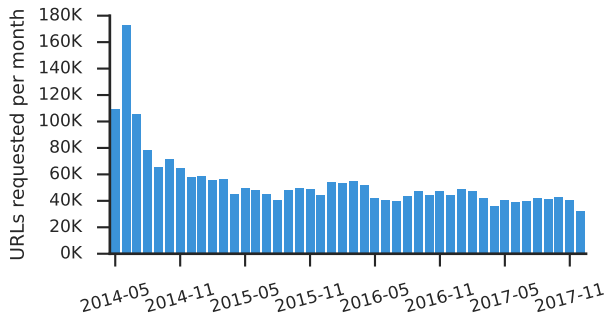


Fig. 1. URLs requested for delisting under the RTBF, including rejected requests. After an initial burst of requests as enforcement went into effect, there were an average of approximately 45,000 requested URLs per month since January 1, 2015.

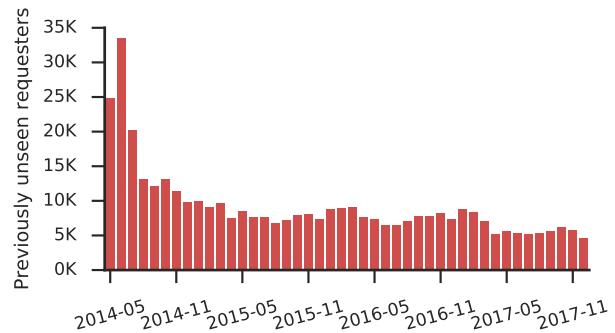


Fig. 2. Timeline of previously unseen requesters based on the date of their first request. Year after year, the number of new requesters declined.

the initial spike of interest, the number of new requesters since January 1, 2015 fell to an average of approximately 7,200 each month. Broken down by year, 43.5% of previously unseen requesters applied during the first year of enforcement, 23.8% in the second, 21.9% in the third year, and the remaining 10.8% in the last seven months. This indicates that while new Europeans continue to seek delistings under the RTBF, the total volume of previously unseen requesters has declined year after year.

4.2 Requests and delistings by country

We find a handful of countries heavily skew the volume of requests, as shown in Table 5. Combined, requesters from France, Germany, and the United Kingdom generated 50.6% of requested URLs. While these are some of the most populous nations covered under the RTBF, they also are some of the highest volume in terms of requests per capita as shown in Figure 3. For every thousand requesters in France, drawn from United

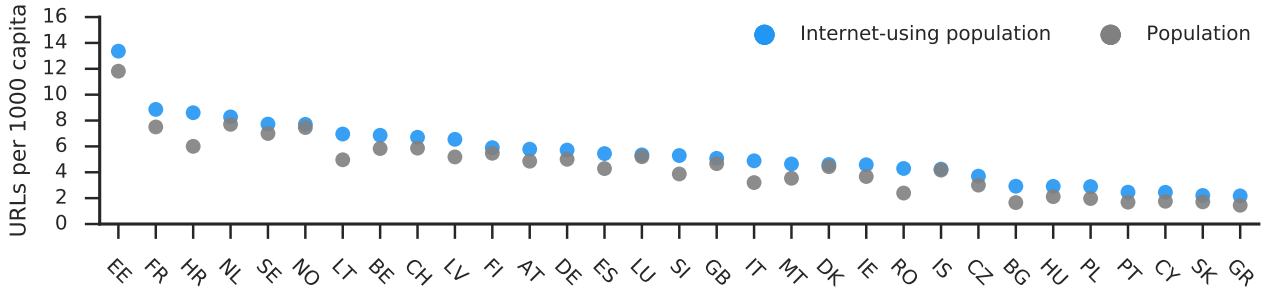


Fig. 3. URLs requested for delisting per capita, sorted by the Internet-using population of each country. For every thousand residents connected to the Internet, there were between 2.2–13.4 URLs requested for delisting per country.

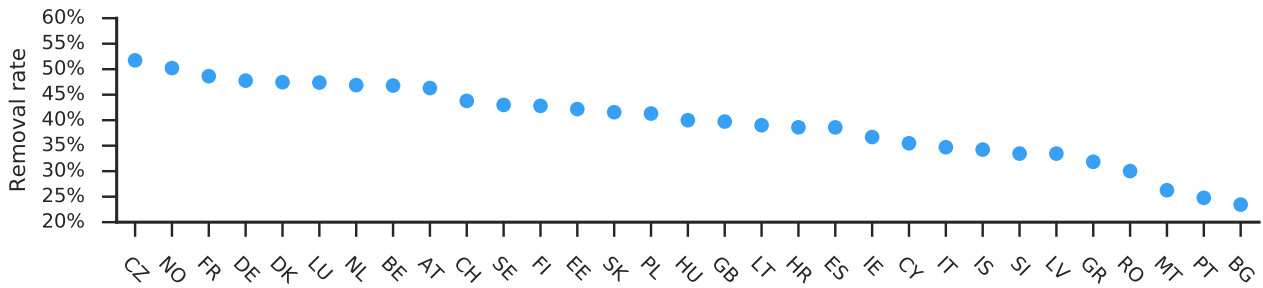


Fig. 4. Delisting rate of requested URLs, broken down by country. Delisting rates range between 23.4–51.7% with the Czech Republic having the highest delisting rate and Bulgaria the lowest.

Country	Code	Requested URLs	Breakdown
France	FR	483,709	20.4%
Germany	DE	409,038	17.3%
United Kingdom	GB	305,267	12.9%
Spain	ES	198,782	8.4%
Italy	IT	190,643	8.1%
Netherlands	NL	130,551	5.5%
Poland	PL	75,340	3.2%
Sweden	SE	68,328	2.9%
Belgium	BE	65,856	2.8%
Switzerland	CH	48,818	2.1%

Table 5. Top 10 countries by volume of URLs requested under the RTBF. See the Appendix for a breakdown for all countries.

Nations population estimates for 2015 [28], Google received approximately 7.5 requested URLs. If we restrict population to estimated Internet users, drawn from the International Telecommunication Union Internet penetration estimates for 2015 [18], then every thousand Internet users in France requested 8.9 URLs. In comparison, requesters from Germany and the United Kingdom generated roughly half that volume, with 5.7 and 5.1 requested URLs per thousand Internet-using residents respectively. These findings suggest that usage of the RTBF differs across countries, with requesters from Estonia filing the most requests per capita, and Greece

the least (excluding countries with fewer than 1,000 requests). As we explore later in Section 5, multiple factors may help explain this including regional news and government practices on reporting private information as well as social media penetration.

We provide a detailed breakdown of delisting decision outcomes per country in Figure 4. Delisting rates ranged from 51.7% for URLs requested by individuals from the Czech Republic, down to 23.4% for URLs requested by individuals from Bulgaria. Among the largest countries by volume, France and Germany had similar delisting rates—48.6% and 47.8% respectively—while the United Kingdom had a lower delisting rate at 39.7%. As we show more in Section 5, this spectrum of delisting rates stems from requesters of some countries more frequently targeting news or government websites where there is a public interest value that ultimately leads to rejecting the delisting.

4.3 High-volume requesters

As individuals have varying privacy expectations and digital footprints, we examine the volume of requested URLs per requester. We find a small number of requesters made heavy use of the RTBF to delist URLs as shown in Figure 5. The top thousand requesters (0.25%

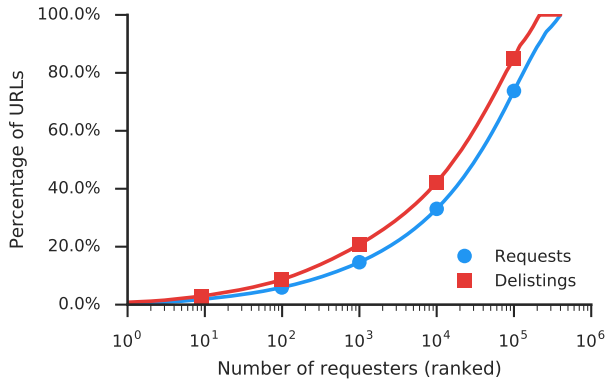


Fig. 5. CDF of all requested and delisted URLs, ranked by the highest volume requester. The top thousand requesters generated 14.6% of URL requests and 20.8% of eventual delistings.

of all requesters) generated 14.6% of requests and 20.8% of delistings. These mostly included law firms and reputation management agencies, as well as some requesters with a sizable online presence. Of the entities in this group, 17.1% resided in Germany, 16.0% in France, and 15.1% in the United Kingdom. The most prolific requester sought to delist 5,768 URLs. In the remaining long tail, 35.1% of requesters sought to delist only a single URL, and 75.2% five or fewer URLs. These results illustrate that while hundreds of thousands of Europeans rely on the RTBF to delist a handful of URLs, there are thousands of entities using the RTBF to alter hundreds of URLs about them or their clients that appears in search results.

4.4 Requesting entities

As the RTBF explicitly outlines public interest as one of the balancing criteria when judging delistings, it’s also important to understand the categories of entities requesting to delist URLs. Starting in 2016, Google began categorizing requesters into coarse groups (discussed in Section 3). We provide a breakdown of all requested URLs since January, 2016 based on these categories in Table 6. A majority of requested URLs—84.5%—were from private individuals. Minors constituted 5.4% of all requested URLs, a special group with a delisting rate nearly twice as high as private individuals. Government officials and politicians generated 3.3% of requested URLs and had a lower delisting rate than private individuals. The low delisting rates for government officials highlights one aspect of the public interest balance that Google strikes when applying the RTBF. We

Requesting entity	Requested URLs	Breakdown	Delisting rate
Private individual	858,852	84.5%	44.7%
Minor	55,140	5.4%	78.0%
Non-governmental public figure	41,213	4.1%	35.5%
Government official or politician	33,937	3.3%	11.7%
Corporate entity	22,739	2.2%	0.0%
Deceased person	4,402	0.4%	27.2%

Table 6. Breakdown of all requested URLs after January 2016 by the categories of requesting entities. Private individuals make up the bulk of requests.

note that corporate entities never have content delisted under the RTBF.

4.5 Processing time

With tens of thousands of RTBF requests filed each month, we examine the time it takes Google to review each request and reach a verdict, where a request may include multiple URLs. In the first two months after Google began delisting URLs under the RTBF, it took a median of 85 days to reach a decision from the time a requester first filed a request. This delay resulted from reviewers establishing cases for what constituted the inaccurate, inadequate, irrelevant, or excessive information. After January 1, 2017, a URL took a median of 4 days to process.

Reviewing requests is critical for catching fraudulent and erroneous submissions. In one case, an individual, who was convicted for two separate instances of domestic violence within the previous five years, sent Google a delisting request focusing on the fact that their first conviction was “spent” under local law. The requester did not disclose that their second conviction was not similarly spent, and falsely attributed all the pages sought for delisting to the first conviction. Reviewers discovered this as part of the review process and the request was ultimately rejected. In another case, reviewers first delisted 150 URLs submitted by a businessman who was convicted for benefit fraud, after they provided documentation confirming their acquittal. When the same person later requested the delisting of URLs related to a conviction for manufacturing court documents about their acquittal, reviewers evaluated the acquittal documentation sent to Google, found it to be a forgery, and reinstated all previously delisted URLs.

More typically though, requests were rejected due to errors on the part of the requester. Common errors included requests for URLs not in Google’s search index and invalid or incomplete URLs (e.g., requesting facebook.com rather than a specific profile page). For requests filed after 2016 when reviewers started annotating requests, 24.5% were rejected due to errors on the part of the requester.

5 Content Targeted

In total, requesters have sought to delist URLs related to 395,907 different hostnames on the Internet. These requests capture a spectrum of potentially personal information: some requesters seek to control their digital footprint exposed through social networks and directory services, while others delist URLs related to news sources and government reports. We explore these divergent applications in terms of the hostnames targeted, the categories of content present on requested pages, and ultimately whether there are better mechanisms for the public to resort to for controlling personal information on popularly requested sites.

5.1 Frequently targeted sites

We present a breakdown of the top 20 hostnames requested for delisting from May 2014–Dec 2017 in Table 7.³ We rely on hostnames rather than second-level domains to differentiate multiple services hosted on the same domain (e.g., wordpress.com). Roughly half of the sites listed are social networks and forums, the most popular including Facebook, YouTube, Google+, Google Groups, and Twitter. In practice requests to these sites may be greater as many operate ccTLDs that are not reflected in the top 20 (e.g., facebook.fr, facebook.de). The other half includes directory sites that aggregate contact details and personal content from other sites, like 118712 and Profile Engine. Combined, the top 20 sites represent 11.2% of requested URLs. Delisting rates varied drastically based on site, ranging 12.1%–91.4%.

³ We exclude *www.* prefixes from hostnames. We also omit 56,804 requested URLs during May 2014–Dec 2017 from this and all subsequent content analysis as they were not well-formed URLs.

Hostname	Description	Requested URLs	Delisting rate
facebook.com	Social	43,423	60.0%
plus.google.com	Social	31,696	43.1%
youtube.com	Social	23,590	43.1%
twitter.com	Social	23,103	47.2%
groups.google.com	Social	17,452	54.5%
annuaire.118712.fr	Directory	15,928	80.5%
profileengine.com	Directory	13,535	91.4%
flashback.org	Other	12,674	12.1%
scontent.cdninstagram.com	Social	11,823	90.2%
badoo.com	Social	10,823	62.4%
myspace.com	Social	7,507	48.5%
copainsdavant.linternaute.com	Directory	6,716	48.9%
societe.com	Directory	6,659	20.4%
wherevent.com	Directory	6,447	91.2%
192.com	Directory	6,043	80.8%
infobel.com	Directory	6,036	78.2%
pbs.twimg.com	Social	5,839	61.9%
verif.com	Directory	5,740	20.0%
linksunten.indymedia.org	Other	5,654	57.4%
linkedin.com	Social	5,563	44.4%

Table 7. Top hostnames requested for delisting during May 2014–Dec 2017 along with a description for content typically hosted on the domain. Social media (and its CDN equivalents), forums, and directory aggregation services dominate the list.

To provide a perspective of the long tail of remaining hostnames, we measure the cumulative fraction of requests covered by successively adding the next most popular hostname in Figure 6. We find 20.9% of requested URLs target only 100 hostnames and 40.8% of requests target the top 1,000 hostnames. In contrast, 60.3% of hostnames received only a single request and 89.6% of hostnames at most 5 requests. Our results indicate that the privacy concerns of Europeans concentrate on a small fraction of the hundreds of millions of hostnames on the Internet.

5.2 Categories of sites

In order to gain a broad perspective of the categories of sites targeted by requests, we rely on two years of labels from January, 2016 onward assigned to each requested URL (discussed in Section 3). These labels cover four mutually exclusive categories: social media content, news sources, directory and information aggregators, and government pages. We provide a breakdown of requested URLs by category in Table 8. Directories that display names, addresses, and other personal information made up 18.8% of requested URLs. Social media made up 13.9% of requested URLs. Delisting rates for

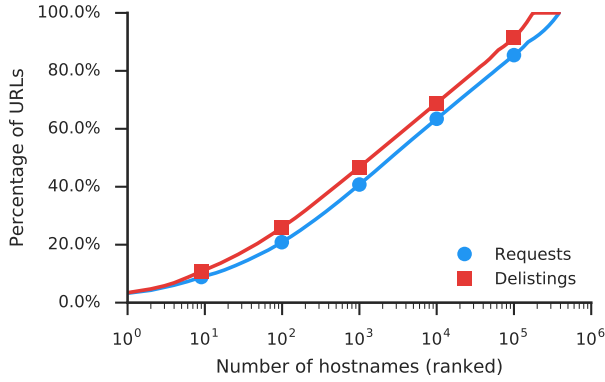


Fig. 6. CDF of all requested and delisted URLs, ranked by the most frequently targeted hostnames. The top thousand hostnames received 40.8% of requests and 46.8% of delistings.

Category of site	Requested URLs	Breakdown	Delisting rate
Directory	187,476	18.8%	51.0%
News	173,860	17.5%	32.2%
Social Media	138,007	13.9%	53.6%
Government	24,718	2.5%	19.5%
Miscellaneous	471,067	47.3%	44.9%

Table 8. Breakdown of requested URLs during the two years from Jan 2016–Dec 2017 based on the category of site requested.

these two categories in aggregate surpassed 51%. News media, entirely absent from the top 20 requested hostnames, made up 17.5% of requested URLs; a reflection of the volume of independent media outlets on the Internet. Far less frequent, government pages made up only 2.5% of requested URLs. Given the public interest nature of news and government records, delisting rates were lower than average—32.2% and 19.5% respectively. The remaining 47.3% of URLs did not fall within a well-defined category.

We provide a monthly breakdown of requested URLs by the category of site in Figure 7. We find requests targeting news-related content steadily increased from 14.7% to 19.1%, while social media requests generally declined from 15.5% to 12.5%, other than a renewed burst in the last month focused largely on Instagram. These trends may reflect requesters adapting their privacy concerns over time. For social media, entities may be finding other mechanisms to proactively reign in their social media footprint (or are less concerned). In contrast, news or directory services remain outside a requester’s control.

Cut another way, we find broad variations in the categories of sites requested by individuals of different

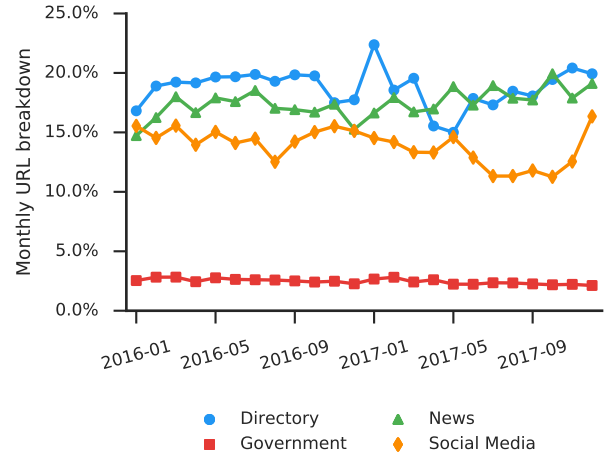


Fig. 7. Monthly breakdown of the categories of sites requested for delisting. Requests to delist social media URLs have trended downward (other than a burst in the last month), while news media has increased.

countries as shown in Table 9. Focusing only on the top five countries by volume of requests, we find requesters from Italy and the United Kingdom were far more likely to target news media in their requests (32.5% and 25.2% of requested URLs respectively). This correlates with diverging journalistic practices: news sources in Italy and the United Kingdom are more prone to reveal the identity of individuals in relation to articles covering crimes. In contrast, news sources in Germany and France tend to anonymize the parties in articles covering crimes. Requesters in France and Germany were most concerned with information exposed on social media and via directory services compared to other countries. Finally, in Spain, 10.6% of requested URLs targeted government records. This may stem from Spanish law which requires the government to publish ‘edictos’ and ‘indultos’. The former are public notifications to inform missing individuals about a government decision that directly affects them; the latter are government decisions to absolve an individual from a criminal sentence or to commute to a lesser one. These variations expose how RTBF usage varies by country, illustrating the challenge of one-size-fits-all privacy policies.

5.2.1 News Media

Examining each category in more depth, we present a breakdown of the top 10 news outlets and tabloids targeted by delisting requests during January 2016–December 2017 in Table 10. We also include a count of all historical requests to the same domains for the entire

Country	Directory	Social	News	Gov't	Misc
France	28.5%	17.5%	11.7%	1.7%	40.5%
Germany	20.8%	16.3%	10.3%	0.9%	51.7%
Italy	10.5%	10.3%	32.5%	2.1%	44.6%
Spain	20.1%	12.3%	18.6%	10.6%	38.4%
United Kingdom	14.9%	10.8%	25.2%	2.4%	46.7%

Table 9. Categories of sites requested by the top five requesting countries. Requesters from Italy and the United Kingdom are the most likely to target news media, compared to France and Germany where the major concern is personal information exposed via social networks and directory aggregators.

period of our dataset. The institutions represented provide news coverage to some of the most populous nations in the European Union, including the United Kingdom, Italy, and France. Delisting rates for the January 2016–December 2017 period ranged from 18.1–40.5%.

In order to illuminate aspects of the balancing tests that Google conducts for articles on news media outlets, we provide three examples of requests and their outcomes. In one case, a requester who held a significant position at a major company sought to delist an article about receiving a long prison sentence for attempted fraud. Google rejected this request due to the seriousness of the crime and the professional relevance of the content. In another request, an individual sought to delist an interview they conducted after surviving a terrorist attack. Despite the article’s self-authored nature given the requester was interviewed, Google delisted the URL as the requester was a minor and because of the sensitive nature of the content. Lastly, a requester sought to delist a news article about their acquittal for domestic violence on the grounds that no medical report was presented to the judge confirming the victim’s injuries. Given that the requester was acquitted, Google delisted the article.

We note that the RTBF is just one technical mechanism currently available to delist news stories. Websites can also add a `Disallow` directive in their `robots.txt` to instruct crawlers to ignore any listed URLs. For one popular Italian news site, roughly 20% of URLs requested under the RTBF for delisting appeared in their `disallow` directive, including URLs that Google ultimately rejected to delist on the grounds of public interest. We cannot definitively state what legal mechanism or internal decision lead the news site to add a `disallow` directive for these URLs, but the end result is that the news articles are now removed from *all* search results regardless of the search query or origin of the search request. Whether such site-level controls should

Hostname	Requested URLs	Delisting rate	Requested URLs (all time)
dailymail.co.uk	1,793	27.4%	3,586
ricerca.gelocal.it	1,022	28.6%	2,253
telegraph.co.uk	949	29.9%	2,025
leparisien.fr	880	31.9%	1,822
ricerca.repubblica.it	864	23.0%	1,793
247.libero.it	792	40.5%	1,615
mirror.co.uk	781	32.8%	1,367
ouest-france.fr	738	18.1%	1,142
bbc.co.uk	712	22.4%	1,413
thesun.co.uk	676	32.8%	1,043

Table 10. Top 10 news media sites requested from Jan 2016–Dec 2017.

Hostname	Requested URLs	Delisting rate	Requested URLs (all time)
facebook.com	15,416	52.1%	30,363
twitter.com	11,408	50.1%	20,219
plus.google.com	11,287	33.2%	18,049
scontent.cdninstagram.com	10,388	90.8%	7,397
youtube.com	8,110	40.6%	20,804
pbs.twimg.com	4,382	60.4%	3,635
m.facebook.com	3,088	59.3%	3,740
myspace.com	2,618	46.6%	6,944
linkedin.com	2,442	37.6%	3,101
groups.google.com	2,393	45.4%	16,410

Table 11. Top 10 social media sites requested from Jan 2016–Dec 2017.

play a role in enforcing the RTBF remains an area of discussion [11].

5.2.2 Social Media & Communities

We provide a breakdown of the top ten most popular social media sites targeted for delisting in the last two years in Table 11. Facebook, Twitter, Google+, Instagram, and YouTube were the most frequent target of delisting requests. Delisting rates for the top 10 sites from January 2016–December 2017 ranged from 33.2–90.8%.

All of these social networks provide privacy controls to restrict access to self-authored content, as well as mechanisms to take down abusive content (e.g., revenge porn) that violates the site’s terms of service. These tools may be better suited to individuals seeking solely to control their social media footprint, with the exception of information leaked due to asymmetric privacy controls (e.g., photo tagging, mentions). Furthermore, these privacy controls also extend to search queries per-

Hostname	Requested URLs	Delisting rate	Requested URLs (all time)
annuaire.118712.fr	10,162	77.5%	14,777
societe.com	3,631	16.2%	6,231
profileengine.com	3,191	88.0%	12,394
infobel.com	2,859	75.8%	5,559
verif.com	2,395	14.2%	5,348
copainsdavant.linternaute.com	2,107	44.7%	6,177
192.com	1,930	73.0%	5,668
gepatroj.com	1,904	52.6%	2,077
wherevent.com	1,633	89.5%	6,102
nwumber.fr	1,540	96.1%	1,642

Table 12. Top 10 directory sites requested from Jan 2016–Dec 2017.

Hostname	Requested URLs	Delisting rate	Requested URLs (all time)
boe.es	1,708	25.9%	3,926
beta.companieshouse.gov.uk	1,599	5.4%	1,602
infogreffre.fr	932	8.3%	1,763
bocm.es	543	58.1%	1,027
juntadeandalucia.es	271	40.4%	518
thegazette.co.uk	251	1.3%	525
sede.asturias.es	215	46.0%	354
legifrance.gouv.fr	208	7.7%	384
boc.cantabria.es	194	65.2%	436
caib.es	191	59.5%	397

Table 13. Top 10 government sites requested from Jan 2016–Dec 2017.

formed on the respective sites for an individual, whereas the RTBF only covers search engine providers.

Whenever such privacy controls are relevant, Google steers requesters towards their use. For example, in one request, an individual sought to delist their Facebook profile. Google noted that Facebook has procedures to limit the visibility of the content in question for all search engines, not just Google, and recommended that the requester utilize these controls. In another case, an individual requested to delist a URL to a page that had taken a self-published image and reposted it. With no social media privacy controls to intervene, Google delisted this URL.

5.2.3 Directories

Directory services index and aggregate information available publicly through social media channels and business listings (e.g., Yellow Pages) such as names, addresses, phone numbers, and more. We provide a breakdown of the most popularly requested sites in Table 12. The most popular site, 118712.fr, is a discovery service in France for looking up professionals and individuals. Delisting rates for all these sites during the January 2016–December 2017 period ranged from 14.2–96.1%. Delisting rates were higher for directories aggregation personal information compared to professional information.

Examples of requests include a requester that sought to delist several URLs leading to a directory page displaying their address and phone number. Google delisted all of the URLs. In another case, a requester sought to delist a URL that listed the requester’s directorships in various companies. Google denied the request as the URL contained information of professional

relevance which appeared neither inaccurate nor outdated.

We note that most directory sites provide their own search functionality as well. While successful RTBF requests delist directory pages for individuals on Google Search, the public can still perform a direct search via any of the popular directory services if no additional RTBF action is taken on those sites directly. Discrepancies between search indexes can lead to possible privacy risks, such as identifying requesters [29].

5.2.4 Government Records

We provide a breakdown of the top 10 most popular government and government-affiliated websites requested for delisting over the last two years in Table 13. Popular hosts include *boe.es* which posts government bulletins and *companieshouse.gov.uk* which serves as a government-operated directory of businesses in the UK. Among the most popular sites, delisting rates for January 2016–December 2017 ranged from 1.3–65.2%.

As mentioned at the start of Section 5.2, many of these requests relate to ‘edictos’ and ‘indultos’ posted by the Spanish government. In one such case, a requester sought to delist 3 URLs for a Spanish government page containing a notice summoning the requester to city hall for a matter related to tax debt. Google delisted all of the URLs. In another request, an individual sought to delist a government page from 2015 that described how the requester had received a prison sentence for managing several companies whilst being subject to bankruptcy order and restrictions. Google denied the delisting request as the information was published by a government entity and there was significant professional relevance regarding the content.

Hostname	Requested URLs	Delisting rate	Requested URLs (all time)
flashback.org	5,649	7.3%	12,510
goo.gl	3,590	0.3%	2,673
linksunten.indymedia.org	2,646	59.5%	5,635
camgirl.gallery	1,771	9.2%	1,653
archive.is	1,730	17.2%	1,854
webcamrecordings.com	1,480	64.8%	1,262
lh3.googleusercontent.com	1,358	47.2%	2,198
pressreader.com	1,249	39.4%	1,819
sexescort1.com	1,174	16.6%	1,136
issuu.com	1,035	34.2%	2,407

Table 14. Top 10 miscellaneous sites requested from Jan 2016–Dec 2017. We note that *googleusercontent.com* is a CDN for multiple types of content and that *goo.gl* is a URL shortener.

5.2.5 Miscellaneous

We provide a breakdown of the top hostnames that fell into the miscellaneous category in Table 14. Adult websites, blogs, and forums make up 8 of the top 10 sites. Delisting rates for all of the most popular pages ranged from 0.3–65.0% for the period of January 2016–December 2017. We note URLs shortened via *goo.gl* are not in Google’s search index, leading to the low delisting rate shown. Instead, Google asks requesters to provide the full URL. Similarly, the low delisting rate for adult websites stems from a bias where a small number of requesters generated the vast majority of requests, where the content was still relevant to their ongoing career as an adult performer.

Examples of requests for this category include a requester that sought to delist multiple URLs to a website displaying nude images from the individual’s previous profession at an adult cam website. Google delisted all of the URLs given the sensitive nature of the content and the lack of professional relevance due to the individual’s change to a career outside the adult industry. In another case, a prominent former politician requested to delist a URL leading to a Wikipedia page that detailed a conviction the requester claimed was spent. Google denied the request due to the public role of the requester and the significant public interest regarding the conviction referenced at the URL.

5.3 Information present on page

As a final measure of how Europeans utilize the RTBF in practice, we examine the relation between each requester and the information present at URLs requested

Content on page	Requested URLs	Breakdown	Delisting rate
Professional information	182,012	23.8%	16.7%
Self authored	79,429	10.4%	31.9%
Crime	64,597	8.4%	47.4%
Professional wrongdoing	54,065	7.1%	14.5%
Personal information	53,067	6.9%	97.8%
Political	25,806	3.4%	2.8%
Sensitive personal information	18,203	2.4%	90.1%
Name not found	123,501	16.2%	100.0%
Miscellaneous	163,918	21.4%	67.9%

Table 15. Potentially private information appearing on URLs requested for delisting from Jan 2016–Dec 2017. Most delistings relate to professional information and self-authored content (e.g., a social media post).

between January 2016–December 2017. For the purposes of this analysis, we exclude all misfiled URLs (discussed in Section 4.5). We provide a general summary of the content referenced by URLs in Table 15 using the categories we previously outlined (see Section 3). The most commonly requested content related to professional information, which rarely met the criteria for delisting (16.7%). Many of these requests pertain to information which is directly relevant or connected to the requester’s current profession and is therefore in the public interest to have indexed in Google Search. Personal information (e.g., a home address, contact information) made up only 6.9% of requests, while more sensitive personal information (e.g., medical information, sexual orientation) made up only 2.4% of requests. The delisting rate for both of these categories exceeds 90%, indicating that any potential public interest is usually outweighed by the sensitive nature of the content⁴. Likewise, instances of when a URL appeared in the search results for a requester’s name, but nevertheless had no reference to the requester in the content, had a delisting rate of 100% due to no competing public interest. In contrast, only 2.8% of requests targeting a requester’s political platform met the criteria for delisting.

We find that where potentially private information appears is heavily dependent on the category of site in-

⁴ We note that sensitive personal information appearing to have a lower delisting rate than personal information is the result of a bias where that information appears: personal information most commonly appears on directory sites with limited public interest, whereas sensitive personal information appears more often on news articles where there are public interests.

Content on page	Directory	Social	News	Gov't	Misc
Crime	2.1%	2.5%	22.8%	8.9%	7.0%
Personal information	20.1%	3.7%	1.9%	2.5%	4.5%
Political	0.5%	1.1%	7.1%	5.0%	3.7%
Professional information	35.6%	6.7%	18.2%	48.6%	24.7%
Professional wrongdoing	1.4%	1.6%	20.3%	4.2%	5.9%
Self authored	3.3%	34.6%	4.9%	1.3%	9.0%
Sensitive personal information	0.9%	0.9%	1.6%	1.0%	3.9%
Name not found	13.8%	24.9%	9.0%	6.3%	18.1%
Miscellaneous	22.4%	24.1%	14.1%	22.2%	23.3%

Table 16. Breakdown of information found per category of site, summed vertically. The majority of social media sites requested for delisting were self-authored posts, whereas news articles predominantly dealt with criminal or professional wrongdoing.

involved in a RTBF delisting request, as shown in Table 16. In the case of news, requests most frequently targeted articles covering crimes or professional wrongdoing. Conversely, individuals seeking to delist social media content most often targeted self-authored posts. These results again highlight two classes of individuals using the RTBF: those seeking to delist personal information appearing on social media and directory sites, and those targeting news and government sites that report crime-related or professionally relevant information respectively.

6 Extraterritorial Requests

As a final measurement, we examine how information requested for delisting under the RTBF relates to the audience of that information. In particular, we explore whether requests span beyond the national boundaries or even the continental boundaries of the requester. However, mapping geographic boundaries to Internet sites or audiences in a generalized way is extremely difficult. We consider two proxy metrics: requests to news outlets and other pages headquartered outside the requester’s country, and requests to delist content from sites whose country code top-level domain (ccTLD) differs from the requester’s country.

6.1 Regional & International News

For the top 25 news hostnames that received delisting requests from January 2016–December 2017, we calculate the fraction of requests that originated from entities residing in the same country as the media outlet’s head-

quarters. As shown in Table 17, we find all 25 of the top news hostnames received over 86% of requests from local individuals. The Irish Independent, headquartered in Ireland, received 10.3% of URL delisting requests from UK individuals (possibly in Northern Ireland). This concentration indicates that most requesters have local intent for their delistings.

While requests from Europeans to non-European media outlets occurred, as shown in Table 18, they happened an order of magnitude less frequently compared to European media outlets (previously shown in Table 10). The most popular targets included Bloomberg, the Financial Times, and the New York Daily News. This mirrors our previous finding that the majority of requests to news outlets originate from local requests.

6.2 Directories, Social, and Government

We repeat the same local analysis, this time for the top 10 hostnames categorized as directory, social media, and government pages from January 2016–December 2017. We present our results in Table 19. For directory services and government pages, there is a clear banding of local requests. More than 92.5% of requests to 7 of the top 10 directory services originated from local requesters. The same was true of 82.5% of requests to all of the top 10 government pages. Only social media and a single directory page with a global user base received delisting requests from a variety of European countries. These results illustrate again the local nature of a majority of requests.

6.3 Country code domains

As an alternate measure of extraterritorial requests, we examine the relationship between the locale of a requester (e.g., France) and the ccTLDs of the URLs requested for delisting (e.g., *.fr* in the case of *number.fr*). We compare this to requests for non-EU ccTLDs and non-ccTLDs altogether such as *.com*. We note these ccTLDs relate only to the publisher’s site and not to any country specific search portal.

We present our results in Table 20 for the top five countries by volume of requests. We find 31.6–51.3% of requests targeted a host registered to a ccTLD, of which 77.2–88.2% were within the requesters locale. As with our analysis of news media, our results highlight the local nature of a majority of RTBF requests (excluding URLs with no clear locale).

Media outlet	BE	DE	ES	FR	GB	HR	IE	IT	Other
hln.be	91.4%	0%	0%	0.5%	1.6%	0%	0.2%	0%	6.2%
nieuwsblad.be	96.1%	0%	0%	0%	0%	0%	0.2%	0.5%	3.1%
bild.de	0%	97.2%	0%	0.9%	1.7%	0%	0%	0%	0.2%
elmundo.es	0%	0.9%	96.1%	1.3%	0.2%	0%	0%	0.6%	0.9%
elpais.com	0%	0.2%	97.7%	0.7%	0.2%	0%	0%	0.8%	0.5%
expansion.com	0.7%	0%	98.0%	0.5%	0%	0%	0%	0.2%	0.5%
ladepeche.fr	0.5%	0%	0%	99.3%	0%	0%	0%	0.2%	0%
lavoixdunord.fr	1.1%	0%	0%	98.6%	0%	0%	0%	0%	0.2%
leparisien.fr	0.3%	0.1%	0%	97.4%	0.7%	0%	0%	0.7%	0.8%
letelegramme.fr	0%	0.3%	0%	99.5%	0.2%	0%	0%	0%	0%
estrepublicain.fr	0.3%	0.3%	0%	99.5%	0%	0%	0%	0%	0%
ouest-france.fr	0.1%	0.3%	0%	99.3%	0%	0%	0%	0.1%	0.1%
bbc.co.uk	0%	0.1%	0.1%	0.3%	97.2%	0%	0.7%	0.3%	1.3%
dailymail.co.uk	0.5%	1.2%	0.2%	2.2%	91.2%	0.1%	0.6%	0.8%	3.1%
mirror.co.uk	0.8%	0.6%	0%	0.9%	94.1%	0%	0.6%	0.5%	2.4%
news.bbc.co.uk	0%	1.0%	1.0%	2.1%	91.6%	0%	1.3%	1.3%	1.7%
standard.co.uk	0.2%	1.2%	0%	0.6%	95.2%	0%	0.2%	0.2%	2.3%
telegraph.co.uk	0.3%	1.5%	0.2%	3.5%	91.4%	0%	0.4%	0.5%	2.2%
theguardian.com	0.2%	1.3%	0.4%	4.0%	86.8%	0.2%	1.7%	1.3%	4.2%
thesun.co.uk	0.7%	0.4%	0.1%	3.0%	92.0%	0%	1.0%	0%	2.7%
index.hr	0%	1.8%	0%	0.5%	1.8%	91.7%	0%	0%	4.3%
independent.ie	0%	0.2%	0%	0.4%	10.3%	0%	87.7%	0%	1.4%
247.libero.it	0%	0.1%	0.1%	0.1%	0.6%	0%	0%	98.9%	0.1%
ricerca.gelocal.it	0%	0.1%	0.2%	0%	0.1%	0%	0%	99.1%	0.5%
ricerca.repubblica.it	0.1%	0%	0.1%	0%	0%	0%	0%	98.8%	0.9%

Table 17. Top 25 hostnames categorized as news targeted for delisting, broken down by the country of the requester. We find requesters were overwhelmingly located in the same country as the outlet. We note some hostnames belong to the same outlet.

Hostname	Requested URLs	Delisting rate	Requested URLs (all time)
bloomberg.com	107	25.5%	274
nydailynews.com	94	32.8%	143
ft.com	82	12.7%	196
nytimes.com	69	27.8%	199
huffingtonpost.com	58	38.0%	148
nypost.com	58	27.8%	134
reuters.com	47	16.7%	95
timesofindia.indiatimes.com	37	23.1%	68
vice.com	37	17.6%	103
washingtonpost.com	33	22.7%	71

Table 18. Top 10 non-EU news hostnames targeted for delisting from Jan 2016–Dec 2017. These requests occurred an order of magnitude less frequently than EU news outlets.

7 Related Work

Previous studies on the RTBF have largely focused on examining the competing interests between the right to privacy and freedom of expression, as well as the technical challenges of enforcement [3, 5, 15, 25]. Apart from these legal and policy analyses, information on how the

RTBF has been used in practice is limited to the transparency reports from Google and Bing [13, 21].

In the closest study to our own, Xue et al. analyzed 283 URLs publicly disclosed by media outlets as being delisted through the RTBF [29]. They found 62 of the articles related to miscellaneous matters, while the remaining 221 articles referenced events surrounding sexual assault, murder, terrorism, and other criminal activity. Our own work expands on the analysis of the categories of content delisted under the RTBF. Our dataset is also unbiased to any particular segment of delistings.

A subset of RTBF requests relate to multi-party privacy conflicts. These arise due to diverging views of how content should be shared and how broadly [6, 16, 17, 20, 27]. The RTBF expands the surface area for such conflicts to occur, even beyond social networks, due to the inclusion of any websites appearing in search results. Equally challenging, the RTBF leaves the resolution of these privacy conflicts to both search engines and Data Protection Authorities rather than the parties in conflict.

Hostname	BE	DE	ES	FR	GB	IE	IT	Other
infobel.com	8.8%	9.1%	63.4%	4.1%	0.5%	0.0%	3.5%	10.5%
annuaire.118712.fr	0.1%	0.0%	0.1%	99.5%	0.0%	0%	0.0%	0.2%
copainsdavant.linternaute.com	1.2%	0.5%	0.0%	97.3%	0.2%	0%	0.1%	0.6%
gopatroj.com	0%	0%	0.1%	98.9%	0%	0%	0%	1.1%
nuwber.fr	0.1%	0.1%	0%	99.6%	0.1%	0%	0%	0.1%
societe.com	0.3%	0%	0.1%	99.0%	0.2%	0%	0.1%	0.4%
verif.com	0.3%	0%	0.2%	99.0%	0.2%	0%	0.0%	0.4%
192.com	0.3%	0.6%	0.5%	2.5%	92.5%	0.6%	0.2%	2.7%
profileengine.com	3.6%	11.6%	6.7%	30.2%	14.0%	0.7%	3.5%	29.7%
wherevent.com	5.7%	29.8%	3.1%	26.2%	2.9%	0.1%	2.9%	29.2%
facebook.com	3.8%	20.8%	7.3%	21.8%	10.0%	1.0%	5.3%	30.0%
groups.google.com	1.4%	26.5%	5.6%	16.7%	10.5%	0.1%	11.7%	27.5%
linkedin.com	3.7%	9.3%	9.9%	33.2%	11.9%	1.0%	5.0%	26.1%
m.facebook.com	3.0%	17.8%	9.5%	19.3%	10.5%	0.2%	5.7%	33.8%
myspace.com	3.2%	22.5%	5.7%	32.8%	12.9%	0.6%	5.5%	16.7%
pbs.twimg.com	1.6%	16.1%	6.9%	35.1%	11.7%	0.6%	6.0%	22.0%
plus.google.com	3.9%	20.8%	9.0%	28.3%	7.3%	0.4%	4.9%	25.4%
scontent.cdninstagram.com	3.6%	27.9%	5.5%	18.2%	2.5%	0.1%	18.0%	24.1%
twitter.com	2.9%	11.7%	6.2%	37.6%	13.8%	1.1%	4.0%	22.8%
youtube.com	3.6%	15.6%	6.4%	23.2%	12.0%	0.8%	6.6%	31.8%
boc.cantabria.es	0%	0%	100.0%	0%	0%	0%	0%	0%
bocm.es	0%	0%	98.0%	1.7%	0.2%	0%	0%	0.2%
boe.es	0%	0.2%	99.2%	0.1%	0.1%	0%	0.1%	0.4%
caib.es	0%	2.6%	97.4%	0%	0%	0%	0%	0%
juntadeandalucia.es	0.4%	0%	99.3%	0%	0%	0%	0%	0.4%
sede.asturias.es	0%	0%	99.5%	0%	0.5%	0%	0%	0%
infogrefe.fr	0%	0%	0%	99.6%	0.3%	0%	0%	0.1%
legifrance.gouv.fr	0.5%	0%	0.5%	98.6%	0%	0%	0%	0.5%
beta.companieshouse.gov.uk	0.4%	3.4%	1.3%	2.5%	82.5%	1.3%	1.1%	7.4%
thegazette.co.uk	0%	4.4%	0%	0.8%	93.6%	0.4%	0%	0.8%

Table 19. Breakdown of requests for the top directory, social media, and government pages. As with news media, we find requesters are overwhelmingly located in the same country as the site operator, with the exception of social media.

	GB	FR	DE	ES	IT
Non-ccTLD	64.8%	68.4%	51.9%	64.7%	48.7%
ccTLD	35.2%	31.6%	48.1%	35.3%	51.3%
Local ccTLD	77.2%	80.1%	82.3%	83.2%	88.2%
Other EU ccTLD	10.0%	9.9%	9.8%	5.8%	5.5%
Non-EU ccTLD	12.7%	10.0%	7.9%	11.0%	6.3%

Table 20. Breakdown of extraterritorial requests based on the ccTLD of requested URLs. We note the ccTLD refers only to a publisher’s site and is unrelated to country specific search portals.

8 Conclusion

In this paper, we presented an in-depth analysis of how the RTBF affects access to information on Google Search. Our analysis covered nearly 2.4 million URLs that Europeans requested to delist over the last three years and seven months. We identified two dominant categories of RTBF requests: delisting personal information found on social media and directory sites; and

delisting legal history and professional information reported by news outlets and government pages. Our results showed the intent behind these requests was nuanced and stemmed in part from variations in regional privacy concerns, local media norms, and government practices. While a majority of URLs were requested by private individuals—85%—we found that politicians and government officials requested to delist 33,937 URLs and non-governmental public figures another 41,213 URLs. Overall, the RTBF can lead to a reshaping of search results for certain individuals, where just 1,000 entities (0.25% of roughly 400,000 Europeans) requested to delist over 346,000 URLs.

Acknowledgements

We thank Nina Taft, Beckett Madden-Woods, and Thibault Guiroy from Google and Luciano Floridi from

the Oxford Internet Institute for their valuable feedback in drafting this report. We also thank all the other Googlers involved in the RTBF delisting process.

References

- [1] Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González. <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A62012CJ0131>.
- [2] Google Inc. v CNIL (C-507/17). <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:62017CN0507>, 2017.
- [3] Meg Leta Ambrose and Jef Ausloos. The right to be forgotten across the pond. *Journal of Information Policy*, 2013.
- [4] Article 29 Data Protection Working Party. Guidelines on the implementation of the Court of Justice of the European Union judgment on “Google Spain and Inc v. Agencia Española de Protección de Datos (AEPD) and Mario Costeja González” c-131/12. http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp225_en.pdf, 2014.
- [5] Jef Ausloos. The ‘right to be forgotten’—worth remembering? *Computer Law & Security Review*, 2012.
- [6] Andrew Besmer and Heather Richter Lipford. Moving beyond untagging: photo privacy in a tagged world. In *Proceedings of the Conference on Human Factors in Computing Systems*, 2010.
- [7] European Commission. Factsheet on the “right to be forgotten” ruling. http://ec.europa.eu/justice/data-protection/files/factsheets/factsheet_data_protection_en.pdf, 2016.
- [8] David Drummond. Searching for the right balance. <https://europe.googleblog.com/2014/07/searching-for-right-balance.html>.
- [9] Julia Fioretti. France fines google over ‘right to be forgotten’. <http://www.reuters.com/article/us-google-france-privacy-idUSKCN0WQ1WX>, 2016.
- [10] Peter Fleischer. Adapting our approach to the european right to be forgotten. <https://www.blog.google/topics/google-europe/adapting-our-approach-to-european-right/>, 2016.
- [11] Lucianoi Floridi. “the right to be forgotten”: a philosophical view. *Jahrbuch für Recht und Ethik-Annual Review of Law and Ethics*, 23(1):30–45, 2015.
- [12] Google. Google search index removal form. https://support.google.com/legal/contact/lr_eudpa?product=websearch.
- [13] Google. Search removals under european privacy law. <https://transparencyreport.google.com/eu-privacy/overview>.
- [14] Google. Transparency Report. <https://transparencyreport.google.com/eu-privacy/overview>, 2018.
- [15] David Hoffman, Paula Bruening, and Sophia Carter. The right to obscurity: How we can implement the google spain decision. *North Carolina Journal of Law & Technology*, 2015.
- [16] Hongxin Hu, Gail-Joon Ahn, and Jan Jorgensen. Multi-party access control for online social networks: model and mechanisms. *IEEE Transactions on Knowledge and Data Engineering*, 2013.
- [17] Panagiotis Ilia, Iasonas Polakis, Elias Athanasopoulos, Federico Maggi, and Sotiris Ioannidis. Face/off: Preventing privacy leakage from photos in social networks. In *Proceedings of the Conference on Computer and Communications Security*, 2015.
- [18] International Telecommunication Union. Percentage of individuals using the Internet. <https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx>, 2017.
- [19] Jemima Kiss. Dear google: open letter from 80 academics on ‘right to be forgotten’. <https://www.theguardian.com/technology/2015/may/14/dear-google-open-letter-from-80-academics-on-right-to-be-forgotten>, 2015.
- [20] Airi Lampinen, Vilma Lehtinen, Asko Lehmuskallio, and Sakari Tamminen. We’re in it together: interpersonal management of disclosure in social network services. In *Proceedings of the conference on human factors in computing systems*, 2011.
- [21] Microsoft. Content removal requests report. <https://www.microsoft.com/en-us/about/corporate-responsibility/crrr>, 2016.
- [22] Begüm Yavuzdoğan Okumuş and Bensu Aydın. Turkey’s court of constitution officially recognizes right to be forgotten. <http://gun.av.tr/turkeys-court-of-constitution-officially-recognizes-right-to-be-forgotten/>, 2016.
- [23] Julia Powles and Rebekah Larsen. Academic commentary: Google spain. <http://www.cambridge-code.org/googlespain.html>, 2016.
- [24] Mathieu Rosemain and Julia Fioretti. French court refers ‘right to be forgotten’ dispute to top eu court. <http://www.reuters.com/article/us-google-litigation-idUSKBN1A41AS>, 2017.
- [25] Jeffrey Rosen. The right to be forgotten. *Stanford Law Review Online*, 2011.
- [26] RT. Russia’s ‘right to be forgotten’ bill comes into effect. <https://www.rt.com/politics/327681-russia-internet-delete-personal/>, 2016.
- [27] Kurt Thomas, Chris Grier, and David M. Nicol. unFriendly: Multi-Party Privacy Risks in Social Networks. In *Proceedings of the Privacy Enhancing Technologies Symposium*, 2010.
- [28] United Nations, Department of Economic and Social Affairs, Population Division. World population prospects (2017). <https://esa.un.org/unpd/wpp/>, 2017.
- [29] Minhui Xue, Gabriel Magno, Evandro Cunha, Virgilio Almeida, and Keith W Ross. The right to be forgotten in the media: A data-driven study. *Proceedings on Privacy Enhancing Technologies*, 2016.

Appendix

Country	Code	Requested URLs	Breakdown
Austria	AT	42,138	1.8%
Belgium	BE	65,856	2.8%
Bulgaria	BG	11,887	0.5%
Croatia	HR	25,439	1.1%
Cyprus	CY	2,040	0.1%
Czech Republic	CZ	31,890	1.3%
Denmark	DK	25,208	1.1%
Estonia	EE	15,540	0.7%
Finland	FI	29,984	1.3%
France	FR	483,709	20.4%
Germany	DE	409,038	17.3%
Greece	GR	16,289	0.7%
Hungary	HU	20,727	0.9%
Iceland	IS	1,376	0.1%
Ireland	IE	17,253	0.7%
Italy	IT	190,643	8.1%
Latvia	LV	10,334	0.4%
Lithuania	LT	14,559	0.6%
Luxembourg	LU	2,948	0.1%
Malta	MT	1,512	0.1%
Netherlands	NL	130,551	5.5%
Norway	NO	38,777	1.6%
Poland	PL	75,340	3.2%
Portugal	PT	17,594	0.7%
Romania	RO	47,586	2.0%
Slovakia	SK	9,318	0.4%
Slovenia	SI	8,027	0.3%
Spain	ES	198,782	8.4%
Sweden	SE	68,328	2.9%
Switzerland	CH	48,818	2.1%
United Kingdom	GB	305,267	12.9%

Table 21. RTBF requests per country from May 2014–Dec 2017, excluding countries with fewer than 1,000 requested URLs.

Country	Directory	Social	News	Gov't	Misc
Austria	17.0%	16.7%	14.8%	2.1%	49.5%
Belgium	15.0%	17.6%	16.2%	0.8%	50.4%
Bulgaria	5.1%	6.4%	21.1%	2.0%	65.4%
Croatia	10.0%	13.0%	28.0%	3.5%	45.6%
Cyprus	7.2%	22.0%	16.4%	1.6%	52.8%
Czech Republic	13.8%	8.3%	15.3%	1.6%	61.0%
Denmark	8.9%	13.2%	40.9%	0.6%	36.4%
Estonia	14.1%	11.0%	23.3%	2.0%	49.6%
Finland	15.9%	13.7%	13.7%	1.6%	55.0%
France	28.5%	17.5%	11.7%	1.7%	40.5%
Germany	20.8%	16.3%	10.3%	0.9%	51.7%
Greece	5.7%	9.4%	24.1%	2.9%	57.9%
Hungary	10.4%	12.0%	16.7%	2.1%	58.8%
Iceland	9.5%	8.7%	21.4%	0.8%	59.6%
Ireland	9.8%	11.4%	35.8%	1.2%	41.8%
Italy	10.5%	10.3%	32.5%	2.1%	44.6%
Latvia	11.7%	13.9%	20.1%	3.5%	50.7%
Lithuania	6.8%	10.4%	23.9%	2.8%	56.1%
Luxembourg	18.2%	17.5%	16.1%	2.2%	46.1%
Malta	6.6%	11.3%	25.8%	9.8%	46.5%
Netherlands	13.4%	17.4%	12.8%	0.4%	56.0%
Norway	17.2%	14.1%	18.4%	0.8%	49.5%
Poland	19.5%	12.9%	13.7%	1.9%	52.1%
Portugal	13.1%	9.4%	13.5%	10.7%	53.3%
Romania	9.7%	4.7%	17.3%	3.8%	64.6%
Slovakia	23.0%	12.0%	18.1%	2.4%	44.5%
Slovenia	16.1%	12.0%	27.1%	2.3%	42.5%
Spain	20.1%	12.3%	18.6%	10.6%	38.4%
Sweden	26.6%	9.9%	7.4%	0.4%	55.7%
Switzerland	20.6%	13.3%	15.5%	2.4%	48.3%
United Kingdom	14.9%	10.8%	25.2%	2.4%	46.7%

Table 22. Types of sites requested per country for Jan 2016–Dec 2017, excluding countries with fewer than 1,000 requested URLs.