

Is anything ever 'forgotten' online?

11 July 2016, by Keith W. Ross



How hard is it to find what people would prefer was forgotten? Credit: shutterstock.com

When someone types your name into Google, suppose the first link points to a newspaper article about you going bankrupt 15 years ago, or to a YouTube video of you smoking cigarettes 20 years ago, or simply a webpage that includes personal information such as your current home address, your birth date, or your Social Security number. What can you do – besides cry?

Unlike those living the United States, Europeans actually have some recourse. The European Union's "right to be forgotten" (RTBF) law allows EU residents to fill out an online form [requesting that a search engine \(such as Google\) remove links](#) that compromise their privacy or unjustly damage their reputation. A committee at the search company, primarily consisting of lawyers, will [review your request](#), and then, if deemed appropriate, the site will no longer display those unwanted links when people search for your name.

But privacy efforts can backfire. A landmark example of this happened in 2003, when actress and singer [Barbra Streisand sued a California couple](#) who took aerial photographs of the entire length of the state's coastline, which included Streisand's Malibu estate. [Streisand's suit argued](#) that her privacy had been violated, and tried to get the photos removed from the couple's website so

nobody could see them. But the lawsuit itself drew worldwide media attention; far more people saw the images of her home than would have through the couple's online archive.

In today's digital world, privacy is a regular topic of [concern and controversy](#). If someone discovered the list of all the things people had asked to be "forgotten," they could shine a spotlight on that sensitive information. Our research explored whether that was possible, and how it might happen. Our research has shown that hidden news articles can be unmasked with some hacking savvy and a moderate amount of financial resources.

Keeping the past in the past

The RTBF law does not require websites to take down the actual web pages containing the unwanted information. Rather, just the [search engine](#) links to those pages are removed, and only from results from searches for specific terms.

In most circumstances, this is perfectly fine. If you shoplifted 20 years ago, and people you have met recently do not suspect you shoplifted, it is very unlikely they would discover – without the aid of a search engine – that you ever shoplifted by simply browsing online content. By removing the link from Google's results for searches of your name, your brief foray into shoplifting would be, for all intensive purposes, "forgotten."

This seems like a practical solution to a real problem that many people are facing today. [Google has received requests](#) to remove more than 1.5 million links from specific search results and has removed 43 percent of them.

'Hiding' in plain sight

But our recent research has shown that a transparency activist or private investigator, with modest hacking skills and financial resources, [can find newspaper articles that have been removed from search results](#) and identify the people who

requested those removals. This data-driven attack has three steps.

First, the searcher targets a particular online newspaper, such as the Spanish newspaper El Mundo, and uses automated software tools to download articles that may be subject to delisting (such as articles about financial or sexual misconduct). Second, he again uses automated tools to get his computer to extract the names mentioned in the downloaded articles. Third, he runs a program to query google.es with each of those names, to see if the corresponding article is in the google.es search results or not. If not, then it is most certainly a RTBF delisted link, and the corresponding name is the person who requested the delisting.

As a proof of concept, we did exactly this for a subset of articles from El Mundo, a Madrid-based daily newspaper we chose in part because one of our team speaks Spanish. From the subset of downloaded articles, we discovered two that are being delisted by google.es, along with the names of the corresponding requesters.

Using a third-party botnet to send the queries to Google from many different locations, and with moderate financial resources (\$5,000 to \$10,000), we believe the effort could cover all candidate articles in all major European newspapers. We estimate that 30 to 40 percent of the RTBF delisted links in the media, along with their corresponding requesters, could be discovered in this manner.

Lifting the veil

Armed with this information, the person could publish the requesters' names and the corresponding links on a new website, naming those who have things they want forgotten and what it is they hope people won't remember. Anyone seeking to find information on a new friend or business associate could visit this site – in addition to Google – and find out what, if anything, that person is trying to bury in the past. One such site already exists.

At present, European law only requires the links to be removed from country- or language-specific

sites, such as google.fr and google.es. Visitors to google.com can still see everything. This is the source of a [major European debate](#) about whether the right to be forgotten should also require Google to remove links from searches on google.com. But because our approach does not involve using google.com, it would still work even if the laws were extended to cover google.com.

Should the right to be forgotten exist?

Even if delisted links to news stories can be discovered, and the identities of their requesters revealed, the RTBF law still serves a useful and important purpose for protecting personal privacy.

By some estimates, [95 percent of RTBF requests](#) are not seeking to delist information that was in the news. Rather, people want to protect personal details such as their home address or sexual orientation, and even photos and videos that might compromise their privacy. These personal details [typically appear in social media](#) like Facebook or YouTube, or in profiling sites, such as profileengine.com. But finding these delisted [links](#) for social media is much more difficult because of the huge number of potentially relevant [web pages](#) to be investigated.

People should have the right to retain their privacy – particularly when it comes to things like home addresses or sexual orientation. But you may just have to accept that the world might not actually forget about the time when as a teenager when your friend challenged you to shoplift.

This article was originally published on [The Conversation](#). Read the [original article](#).

Source: The Conversation

APA citation: Is anything ever 'forgotten' online? (2016, July 11) retrieved 21 September 2020 from <https://techxplore.com/news/2016-07-forgotten-online.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.