Information Brokers, Fairness, and Privacy in Publicly Accessible Information

Andrea Slane*

The European Union, Canada, and the United States have each grappled with what counts as fair business practices in relation to information services that collect and package personal information that has ended up in one way or another online. On the open internet, this personal information often originates from two types of online sources: public records like arrests, mugshots, court decisions, and bankruptcy records; and user-generated content hosted on social media platforms and sites. This article argues that personal information that has been exposed to public view — be it by a government institution, another individual or organization, or by the data subject him or herself — should not be considered fair game to any and all subsequent commercial exploitation. The blunt concept of "public" information should be refined to a more nuanced understanding of "publicly accessible" information, where public access can be limited to particular purposes. By focusing on fairness in business dealings in publicly accessible personal information, it should be possible to move beyond a fixation on locating the elusive divide between private and public online information, and instead frame privacy as situated in a three-way balance of interests between the business, the public, and the data subject.

* Andrea Slane, PhD, Associate Professor in Legal Studies, University of Ontario Institute of Technology, Oshawa, Ontario: Andrea.slane@uoit.ca.

I. Introduction

In the last decade, online information brokers have come under increasing scrutiny from regulators in the European Union, Canada, and the United States. Each jurisdiction has grappled with where to draw the line regarding what kind of business practices are fair in each regime, especially where online businesses provide an information service that includes the collection and packaging of the personal information of individuals whose information has ended up in one way or another online. A comparison of these efforts reveals important variations and policy options, but also some common ground. This article explores these options and the decisions jurisdictions make to restrain the otherwise unimpeded flow of online personal information through information brokers.

Finding appropriate ways to regulate the way personal information flows through commercial business models is necessary, because the choices we make have implications for general commercial fairness in data processing. In particular, it is important to focus on privacy in publicly accessible personal information, since so much personal data is now generated from "public" online activity. This article will focus on recent legal and regulatory developments in the EU, Canada, and the US that deal with information products and services that collect, process, and package publicly accessible personal information. On the open internet, this personal information often originates from two types

of online sources: public records (like arrests, mugshots, court decisions, and bankruptcy records) and user-generated content hosted on social media platforms and sites.

Personal information that has been exposed to public view — be it by a government institution, another individual or organization, or by the data subject him or herself — should not be thought of as fair game to any subsequent commercial exploitation. The blunt concept of "public" information should be refined by shifting to a more nuanced understanding of "publicly accessible" information, where public access to that information can be limited to particular purposes. Each of the three jurisdictions has been engaged in determining what are fair purposes for accessing and subsequently exploiting personal information for commercial gain, albeit in their own distinct ways.

The concept of fairness permeates attempts to restrain commercial exploitation of publicly accessible personal information online. Fairness in business practices as they apply to individuals — whether they be customers or members of the broader public — governs the balance between the value we place in entrepreneurialism and the free market, the right of the public to the benefits provided by those business practices, and the rights of data subjects to be sheltered from certain types and magnitudes of informational harm. By focusing on fairness in business dealings in publicly accessible personal information, it should be possible to move beyond a fixation on locating the elusive divide between private and public online information, and instead frame privacy as situated in a three-way balance of interests among the business, the public, and the data subject.

In the US, efforts to articulate and manage the legitimate flow of personal information online have been spearheaded by the Federal Trade Commission ("FTC"), in particular its enforcement of fair credit reporting obligations and its intervention in unfair and deceptive business practices. In the EU and Canada, these efforts are rooted in data protection regimes that are intended to enforce fair information practices. This article compares how each of the three jurisdictions are working to determine to what extent, and how, existing consumer or data protection regimes should limit the commercial exploitation of

publicly accessible personal information about non-public figures.[1] Part II applies the EU's approach to the "right to be forgotten" as a starting point for exploring fairness in information location service provision, especially with regard to the Court of Justice of the European Union's ("CJEU") characterization of search engines as information brokers (or, in EU Data Protection Directive terms, "data controllers" that process personal information for commercial purposes).[2] Part III discusses how the US and Canada have each dealt with limits on the commercial exploitation of access to public records. Part IV explores how these jurisdictions have dealt with commercial exploitation of user-generated content containing personal information. Part V considers the problem of digital public culture — that is, how to deal with material containing personal information that is popular online, whether as "news" or as viral content like a meme. In an important sense, viral content can become part of the fabric of digital public culture in the same way that an event that is "newsworthy" merits public exposure and discussion even if it contains personal information and invades an individual's privacy. This section proposes newsworthiness as an arbiter of fairness for capitalizing

1. The distinction between public and private figures arises in the context of defamation and privacy litigation, especially First Amendment jurisprudence in the US. For the purposes of this article, non-public figures are persons whose actions and activities are subject to little or no *specific* public interest. See Susan M Gilles, "Public Plaintiffs and Private Facts: Should the 'Public Figure' Doctrine Be Transplanted into Privacy Law?" (2005) 83:4 Nebraska Law Review 1204. The usefulness of this distinction has also been considered as a way to align the right to be forgotten with the US First Amendment. See Michael L Rustad & Sanna Kulevska, "Reconceptualizing The Right to Be Forgotten to Enable Transatlantic Data Flow" (2015) 28:2 Harvard Journal of Law & Technology 349 at 354.
2. *Google Spain SL and Google Inc v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González* (13 May, 2014), Doc C-131/12, ECLI:EU:C:2014:317 (CJEU) [*Google Spain*]; EC, *Directive 95/46/EC of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data*, [1995] OJ, L 281/31 [Directive 95/46/EC].

on popular user-generated content (though the term itself requires significant refinement), with the aim of allowing for digital public culture to flourish while still protecting privacy of data subjects. Part VI explores the attitude that publicly accessible information is "free for the taking", and how the US and Canada have placed restrictions on businesses that try to unfairly capitalize on this perception.

Overall, the following analysis will demonstrate that broader principles of information fairness should guide choices about how to protect data subjects from the far more powerful forces of commercial enterprises that deal in personal information products and services.

II. The EU's "Right to Be Forgotten" as a Restraint on Commercial Exploitation of Personal Information Online

The EU's implementation of the right to be forgotten is a good starting point for discussing information brokers, fairness, and privacy in publicly accessible personal information, because this right is centrally concerned with whether ongoing public access to personal information that has already been made available online should be permitted. There are two major versions of the right to be forgotten, neither of which is very well captured by the concept of "forgetting". The first is the right to obscurity, which is a narrow procedural remedy for data subjects operating within existing data protection obligations in the EU. The right to obscurity arises from the 2014 CJEU decision in *Google Spain SL, Google Inc v Agencia Española de Protección de Datos (AEPD), Mario Costeja Gonzalez*,[3] which determined that data subjects have a right to require general search engines like Google to de-list certain links that appear in search results of their name, based on the characterization of search engines as information brokers.

The second is the right to erasure, which is a broader substantive right to require data controllers to erase certain online personal information;

3. *Google Spain*, *ibid.*

this will be implemented in the EU *General Data Protection Regulation*[4] ("*GDPR*") that comes into force in May 2018. This right to erasure applies to all data controllers including those that generate their own content (like news agencies), but when applied to secondary online information brokers (like search engines and hosts), it would mean ensuring that content does not appear in search results or otherwise on the hosting service, further reducing public accessibility of that information.[5] This article focuses on secondary information brokers that compile and present information garnered from other sources that do not originate with the business itself.

Two aspects of the *Google Spain* decision are particularly important to the following discussion: (1) the characterization of what Google does as information brokering — that is, the creation of a packaged profile of an individual, and (2) the determination that Google's activities are predominantly commercial rather than, for example, exercised in the public interest. A preliminary determination in *Google Spain* was based on whether Google and other general search engines are subject to the Data Protection Directive. The CJEU considered whether Google engaged in "processing" personal information as a "data controller" as set out in the Directive. The CJEU determined that it did, in that Google controls the algorithm that collects personal information from diverse online sources,

---

4. "[R]ight to be Forgotten, also known as Data Erasure, entitles the data subject to have the data controller erase his/her personal data, cease further dissemination of the data, and potentially have third parties halt processing of the data. The conditions for erasure, as outlined in article 17, include the data no longer being relevant to original purposes for processing, or a data subjects withdrawing consent. It should also be noted that this right requires controllers to compare the subjects' rights to "the public interest in the availability of the data" when considering such requests". See "GDPR Key Changes" *EU General Data Protection Regulation*, online: EUGDPR <eugdpr.org/key-changes.html>.
5. For a fuller discussion of the contours of the right to be forgotten and how it might be implemented in Canada, see Andrea Slane, "Search Engines and the Right to Be Forgotten: Squaring the Remedy with Canadian Values on Personal Information Flow"(forthcoming 2018) 55:2 Osgoode Hall Law Journal [Slane, "Squaring the Remedy"].

then collates and presents it to users in a ranked form.[6] When a person is searched by name, Google gathers available mentions across online sources and produces a profile that potentially has a greater impact on the privacy interests of the data subject than any one of those sources alone.[7]

As for the commercial nature of Google's activities, the CJEU focused on the most straightforward ways that Google makes money from searches, namely through its AdWords advertising program. AdWords uses a "pay per click" advertising model whereby advertisers bid for association with particular search terms, so that links to their sites come up at the top of search results, as tailored to the searcher's geographic area.[8] The CJEU wrote:

> [t]he very display of personal data on a search results page constitutes processing of such data. Since that display of results is accompanied, on the same page, by the display of advertising linked to the search terms, it is clear that the processing of personal data in question is carried out in the context of the commercial and advertising activity of the controller's establishment on the territory of a Member State, in this instance Spanish territory.[9]

The CJEU did not consider whether advertising appears on the same page of every type of search result, and whether this makes any difference to the overall analysis. In general, an individual's name, even a well-known public figure, does not serve as an AdWords linked keyword. However, if the individual's name is searched in conjunction with another term that is an AdWords keyword, then advertising links will appear. For example, if the complainant in the *Google Spain* case is searched in conjunction with the term "bankruptcy" (his complaint aimed to have Google de-list

---

6. *Google Spain*, *supra* note 2 at paras 32–33.
7. *Ibid* at para 37.
8. Rory Cellan-Jones, "How does Google make money?" *BBC News*, online: BBC iWonder <bbc.co.uk/guides/z9x6bk7>; Greg McFarlane, "How Does Google Make Its Money?" *Investopedia* (22 November 2012), online: Investopedia <investopedia.com/stock-analysis/2012/what-does-google-actually-make-money-from-goog1121.aspx>; Julia Love & Rishika Sadam, "Google parent Alphabet's profit up 29 percent on strong ad sales" *Reuters* (27 April 2017), online: Reuters <in.reuters.com/article/alphabet-results/google-parent-alphabets-profit-up-29-percent-on-strong-ad-sales-idINKBN17T2ZQ>.
9. *Google Spain*, *supra* note 2 at para 57.

links to public notices about past debt), then ads for debt relief services will appear at the top of the page.[10]

Nonetheless, having determined that Google is a "data controller" that "processes" personal information within an overall commercial business model that monetizes search results, the search engine is required, upon request, to remove links from the search results of a person's name where those links lead to information that is "inadequate, irrelevant, no longer relevant or excessive" to the purpose for which it was collected, unless there is a public interest in retaining the link to that information upon such a name search.[11]

For the most part, implementation of the *Google Spain* decision appears to be predominantly focused on results containing outdated personal information of non-public figures, where the privacy interests of the data subject outweigh the interests of the public in having access to that specific information through a search of that individual's name (such as a link revealing a long ago conviction for a minor crime).[12] It remains unclear whether the idea of "excessive to the purpose" could be meaningfully applied to a general search engine; if we characterize search engines' purpose for collection as providing a ranked compilation of *most* relevant publicly accessible online information related to that person, then "excessive" is a bit more refined than relevance alone. A search result could also be "excessive" if it returned highly sensitive information. Relevance and excessiveness must in any case be considered normative

---

10. "How Does Google Make Its Money: The 20 Most Expensive Keywords in Google AdWords" *Wordstream*, online: Wordstream <wordstream.com/articles/most-expensive-keywords>. This article used data from 2010–2011 and concluded that the most expensive pay per click word is "insurance" followed by "loans" and "mortgage".
11. The terms "inadequate, irrelevant, no longer relevant or excessive" come from the EU Data Protection Directive, Directive 95/46/EC, *supra* note 2, which requires at art 6(1)(c) that personal data must be "adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed".
12. "Transparency Report: Search removals under European privacy law" *Google*, online: Google <transparencyreport.google.com/eu-privacy/overview> [Google Transparency Report].

terms, akin to "newsworthiness", which is similarly not dependent on the judgement of a particular individual reader, but rather defines the contours of legitimate public interest in having the information.[13]

Relevance and excessiveness relate to newsworthiness, in that relevance implies a public interest in access to this information that outweighs the data subject's privacy interests, an interest that is calculated via the sensitivity of the information at issue. Along these lines, data protection regimes typically exclude the practice of "journalism" from data protection obligations.[14] Therefore, the collection of personal information about the subject of a news item legitimately in the public's interest, even when carried out by a for-profit news organization, is not constrained by obligations that would restrict public access to that news item.[15] In passing, the CJEU rejected the possibility that what search engines do is journalism.[16] The Advocate General's opinion on the case offered some credence to the idea that search engines serve as archives, but reiterated European jurisprudence that has held that news archives

13. Newsworthiness is most often used in the US context in relation to defamation, right of publicity and publication of private facts cases. It has often been criticized by US scholars who consider it to permit too much encroachment on freedom of expression. See *e.g.* Amy Gajda, *The First Amendment Bubble: How Privacy and Paparazzi Threaten a Free Press* (Cambridge: Harvard University Press, 2015); Amy Gajda, "The Present of Newsworthiness" (2016) 50:2 New England Law Review 145. Others consider newsworthiness to provide too easy a justification for violating privacy. See Dianna M Worley, "*Shulman v Group W Productions*: Invasion of Privacy by Publication of Private Facts — Where Does California Draw the Line Between Newsworthy Information and Morbid Curiosity?" (2000) 27 Western State University Law Review 535 at 535.
14. In Canada, see *Personal Information Protection and Electronic Documents Act*, SC 2000, c 5, s 7(1)(c) [*PIPEDA*]. See also Teresa Scassa, "Journalistic Purposes and Private Sector Data Protection Legislation: Blogs, Tweets and Information Maps" (2010) 35:2 Queen's Law Journal 733.
15. For analysis of problems with determining where to draw the line regarding journalism versus commercial speech that can be more heavily regulated, see Diane Leenheer Zimmerman, "Who Put the Right in the Right of Publicity?" (1998) 9:1 DePaul-LCA Journal of Art and Entertainment Law and Policy 35 at 55.
16. *Google Spain*, *supra* note 2 at para 85.

have a greater duty to ensure accuracy of historical information, since the urgency of publishing current affairs is absent.[17] Alternatively, Google tried to claim that it cannot be a data controller because it does not distinguish between different types of data and does not alter that data in presenting results.[18] The CJEU rejected this argument, stating that it makes no difference that Google does not distinguish between personal data and other information, nor does it matter that "[t]hose data have already been published on the internet and are not altered by the search engine".[19]

Several scholars have strongly critiqued Google's assertion that its service merely delivers up informational history, and so serves as a form of cultural memory.[20] For example, Julia Powles noted that many online service providers have been capitalizing on the concept of the internet as a public sphere when really it is "[j]ust an algebraic representation of privately owned services".[21] She warned against equating this privately owned and manipulated network with our commitment to maintaining public records and archives offline (or even digitally stored, but subject to some access controls). In effect, Google is trying to have it both ways: to be legally recognized as the guardian of transparency in the online info-world, and yet, to conceal the algorithm by which such information

17. See *Google Spain SL and Google Inc v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González* (25 June, 2013), Doc C-131/12, ECLI:EU:C:2013:424 at para 123.
18. *Google Spain*, *supra* note 2 at para 22.
19. *Ibid* at paras 28–29.
20. For Google's position, see Richard S Whitt, "'Through a Glass, Darkly': Technical, Policy, and Financial Actions to Avert the Coming Digital Dark Ages" (2017) 33:2 Santa Clara High Technology Law Journal 117.
21. Julia Powles, "The Case That Won't Be Forgotten" (2015) 47:2 Loy University of Chicago Law Journal 583 at 591.

is retrieved and monetized.[22] Google claims to use more than 200 factors when compiling its ranking of search results, with popularity being a dominant factor. But even this one factor, as Powles notes, tends to exacerbate the "man bites dog" problem long recognized in journalism — that what is most popular and sells the most "papers" is not necessarily what is most current, accurate, or most central to overall historical records regarding an individual.[23]

The dominance of the popularity factor is further skewed by the demographics of the audience that most actively uses Google — which has historically been Western, white, middle-class men, although this is slowly changing.[24] The legacy of this bias is evident in studies that have revealed that Google searches are often skewed to favour privileged perspectives — delivering search results that positively reflect whites and negatively reflect African-Americans for instance (*e.g.* "beautiful dreadlocks" turns up images of white people while "unprofessional

22. Richard Curtis, "Google Wants It — and Has It — Both Ways" *Publishing in the 21st Century* (blog) (30 May 2012), online: Publishing in the 21st Century <curtisagency.com/blog/2012/05/google-wants-it-and-has-it-both-ways.html >; Uta Kohl, "Google: The Rise and Rise of Online Intermediaries in the Governance of the Internet and Beyond (Part 2)" (2013) 21:2 International Journal of Law and Information Technology 187 at 191–98.
23. Powles, *supra* note 21 at 610.
24. Bias in machine learning is common, because machines learn from humans and unfortunately humans are biased, especially online. See *e.g.* Aylin Caliskan, Joanna J Bryson & Arvind Narayanan, "Semantics derived automatically from language corpora contain human-like biases" (2017) 356:6334 Science, online: Science <science.sciencemag.org/content/356/6334/183.full>; Tolga Bolukbasi et al, "Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings" (2016) arXiv 1607.06520v11, online: Cornell University Library <arxiv.org/pdf/1607.06520.pdf>.

hairstyles" brings up images of black people).[25] As Safiya Umoja Noble wrote:

> [i]t is dominant narratives about the objectivity and popularity of web search results that make misogynist or racist search results appear to be natural. Not only do they seem ["normal"] due to the technological blind spots of users who are unable to see the commercial interests operating in the background of search (deliberately obfuscated from their view), they also seem completely unavoidable because of the perceived ["popularity"] of sites as the factor that lifts websites to the top of the [results] pile.[26]

Further, Google has been called to task regarding how its AdWord algorithms work. One study found that searches of names associated with African-Americans were more likely to include ads for criminal record checks than neutral names or names associated with white people.[27] In other words, Google's business model delivers results and advertising skewed by existing social bias.

Google is constantly adjusting its algorithms and regularly attempts to address some of these concerns, but doing so merely reinforces the CJEU conclusion that Google indeed controls data collection, packaging, and presentation; Google search results are not neutral reflections of the material that is publicly available on the internet. Therefore, in terms of data protection and consumer protection, skewed results containing personal information should be addressed by requirements related to

25. Fiona Rutherford & Alan White, "This Is Why Some People Think Google's Results Are 'Racist'" *BuzzFeed* (12 April 2016), online: BuzzFeed <www.buzzfeed.com/fionarutherford/heres-why-some-people-think-googles-results-are-racist?utm_term=.kqpDg0ERB7#.dpKoZBwvqA>; Leigh Alexander, "Do Google's 'unprofessional hair' results show it is racist?" *The Guardian* (8 April 2016), online: The Guardian <theguardian.com/technology/2016/apr/08/does-google-unprofessional-hair-results-prove-algorithms-racist->.
26. Safiya Umoja Noble, "Google Search: Hyper-visibility as a Means of Rendering Black Women and Girls Invisible", (2013) 19 InVisible Culture, online: University of Rochester <ivc.lib.rochester.edu/google-search-hyper-visibility-as-a-means-of-rendering-black-women-and-girls-invisible/>.
27. Latanya Sweeney, "Discrimination in Online Ad Delivery" (2013) arXiv: 1301.6822 1, online: Cornell University Library <arxiv.org/pdf/1301.6822.pdf>.

relevance and excessiveness, more fairly balancing the interests of data subjects with the interests of searchers to easily find that information.

III. Using Fairness to Restrict Businesses that Facilitate Access to Public Documents Through Information Compilation Products

Since the advent of the internet, the easy accessibility of personal information has raised concerns about its use by the various gatekeepers of financial and professional opportunities — especially insurers, lenders, admissions officers, and potential employers.[28] Scholars and commentators have debated the best ways to address unfairness that can result from misuse of information found online — from legislation addressing the provision of the information, to legal restrictions on use, to ethical guidelines for these industries.[29] Parallel debates have focused on digitizing and facilitating public access to public documents, such as

28. "Number of Employers Using Social Media to Screen Candidates Has Increased 500 Percent over the Last Decade" *CareerBuilder* (28 April 2016), online: CareerBuilder <careerbuilder.ca/share/aboutus/pressreleasesdetail.aspx?sd=4%2F28%2F2016&id=pr945&ed=12%2F31%2F2016>; Jonathan A Segal & Joyce LeMay, "POINT/COUNTERPOINT: Should Employers Use Social Media to Screen Job Applicants?" *HR Magazine* (1 November 2014), online: SHRM <www.shrm.org/hr-today/news/hr-magazine/pages/1114-social-media-screening.aspx>; Kaitlin Mulhere, "Lots More College Admissions Officers Are Checking Your Instagram and Facebook" *Money* (13 January 2016), online: Time <time.com/money/collection-post/4179392/college-applications-social-media/>; Stephanie Armour, "Borrowers Hit Social-Media Hurdles: Regulators Have Concerns About Lenders' Use of Facebook, Other Sites" *The Wall Street Journal* (8 January 2014), online: Wall Street Journal <www.wsj.com/articles/borrowers-hit-socialmedia-hurdles-1389224469>.

29. Avner Levin, "Losing the Battle but Winning the War: Why Online Information Should Be a Prohibited Ground" (2015) 18:2 Canadian Labour and Employment Law Journal 379; Nathan J Ebnet, "It Can Do More Than Protect Your Credit Score: Regulating Social Media Pre-Employment Screening with the *Fair Credit Reporting Act*" (2012–2013) 97:1 Minnesota Law Review 306.

court decisions and documents, or arrest and detention records.[30]

If Google qualifies as a "data controller" for the purposes of the EU Data Protection Directive, then surely other online businesses that specifically provide a compilation of material about an individual found in public records would also qualify. In the US, restrictions on such businesses are relatively limited, but the FTC has initiated investigations and issued rulings against some of these businesses, including under the *Fair Credit Reporting Act*[31] ("*FCRA*"). The text of the Act is promising in that it defines a "consumer report" as communication of any information by a consumer reporting agency:

> [b]earing on a consumer's credit worthiness … character, general reputation, personal characteristics, or mode of living which is used or expected to be used or collected in whole or in part for the purpose of serving as a factor in establishing the consumer's eligibility for (A) credit or [personal, family or household] insurance … (B) employment.[32]

The *FCRA* also sets out restrictions on specific information that should not be provided as part of a consumer credit report, including outdated financial information (generally after 7 years), bankruptcies after 10 years, arrest records (generally after 7 years), and "[a]ny other adverse item of information, other than records of conviction of crimes" (generally after 7 years).[33] These time limits are related in spirit to the EU's restriction on data controllers dealing in outdated and no longer relevant information,

---

30. Amanda Conley et al, "Sustaining Privacy and Open Justice in the Transition to Online Court Records: A Multidisciplinary Inquiry" (2011–2012) 71:3 Maryland Law Review 772; Karen Eltis, "The Judicial System in the Digital Age: Revisiting the Relationship between Privacy and Accessibility in the Cyber Context" (2011) 56:2 McGill Law Journal 289.
31. In Canada, consumer reporting agencies are regulated by provincial legislation and require registration with a provincial authority. For instance, in Ontario, such agencies are governed by the *Consumer Reporting Act,* RSO 1990, c C-33. However, all businesses are subject to some form of data protection obligations, either the federal *PIPEDA* or substantially similar provincial legislation; for US, see *Fair Credit Reporting Act*, 15 USC § 1681a (1970) [*FCRA*].
32. *FCRA*, *ibid*, § 1681a(d)(1).
33. *Ibid*, § 1681c(a).

but the EU's definition of "data controller" is vastly broader than the *FCRA*'s definition of "consumer reporting agency".

The definition of a "consumer reporting agency" under the *FCRA* encompasses any person or organization that:

> [f]or monetary fees, dues, or on a cooperative nonprofit basis, regularly engages in whole or in part in the practice of assembling or evaluating consumer credit information or other information on consumers for the purpose of furnishing consumer reports to third parties.[34]

This definition is narrowed by the fact that "for the purpose of furnishing consumer reports" incorporates the definition of "consumer reports" as restricted to situations where the information collected is being provided "for the purpose of serving as a factor" in establishing creditworthiness or for employment purposes. The *FCRA* does not capture more general or unspecified purposes for collecting consumer information.

The consequences of this limitation are evident in the FTC's complaint brought forward by the US Attorney General against Spokeo in 2012.[35] Spokeo is an online service that assembles consumer information from online and offline sources to create "consumer profiles" to which it sells access to individuals or businesses. At the time of the complaint, Spokeo marketed its service specifically to the human resources industry as a background screening tool, offering high-volume access via subscription. After the court ruling against Spokeo for violating the *FCRA*, the end result has been that Spokeo no longer specifically markets its service to the human resources industry, but otherwise continues to operate its business in the same fashion, including by offering high-volume subscriptions.[36]

Since the Spokeo ruling, other personal information compilation services have also merely posted disclaimers that their services should not be used for *FCRA*-covered purposes. Truthfinder.com, for instance,

---

34. *Ibid*, § 1681a(f).
35. US, Federal Trade Commission, *United States v Spokeo Inc* (CV12-05001) (2012).
36. Spokeo instead claims that high volume subscriptions "generally appeal to professionals whose work routine includes constant people research". See "FAQs: what are quota upgrades?" *Spokeo*, online: Spokeo <www.spokeo.com/faqs-consumer>.

requires users to click an "I understand" button to enter the site, affirming consent to the statement that:

> TruthFinder does not provide consumer reports and is not a consumer reporting agency. We provide a lot of sensitive information that can be used to satisfy your curiosity, protect your family, and find the truth about the people in your life. You may not use our service or the information it provides to make decisions about consumer credit, employers, insurance, tenant screening, or any other purposes that would require [*FCRA*] compliance.[37]

Instead, Truthfinder's marketing is primarily aimed at individuals who want to learn "the truth about the history of your family and friends", although the service offers "Power Users" a discount for purchasing three months of unlimited searching.[38] Another *FCRA* disclaimer appears in tiny print at the bottom of the welcome page, stating:

> [t]he information available on our website may not be 100% accurate, complete, or up to date, so do not use it as a substitute for your own due diligence, especially if you have concerns about a person's criminal history. TruthFinder does not make any representation or warranty about the accuracy of the information available through our website or about the character or integrity of the person about whom you inquire.[39]

Truthfinder thus does not take any responsibility for the accuracy of its contents, despite what is implied in its name and marketing.

A very similar *FCRA* disclaimer appears on commercial mugshot and arrest record websites, which offer a way to acquire a compilation of this subset of public records, generally scraped from law enforcement and detention centre websites that make such information available online to the public.[40] Debates about the value and purpose of making these sorts of pre-conviction and non-conviction documents a matter of public record have included the public interest argument that publicly inspectable records help ensure the transparency and fairness of the criminal justice system.[41] However, making such records easy to acquire feeds more into

37. *Truthfinder*, online: Truthfinder <www.truthfinder.com>.
38. *Ibid*.
39. *Ibid*.
40. See *e.g. Mugshots*, online: Mugshots <www.mugshots.com> [Mugshots].
41. Danielle Bruno, "Note: Mugshots Or Public Interest? Why FOIA Exemption 7(C) Does Not Categorically Exempt Booking Photographs from Disclosure" (2016) 78 University of Pittsburgh Law Review 95.

the socially punitive approach to persons in conflict with the law. From this perspective, easy-to-access mugshots and arrest records not only allow people to protect themselves from these individuals, but also heighten the effects of conflict with the law through public shaming, even when an individual has not been convicted of a crime. Some US states and counties have made arrest and detention records publicly available online, while others are more restrictive in their release of this information.[42] Publicly available law enforcement and jail websites generally include disclaimers warning that errors and inaccuracies in the information provided are common, and reiterating the basic criminal justice tenet of innocence until proven guilty.[43] EU and Canadian law enforcement organizations generally do not make such information freely available online.[44]

Commercial mugshot and arrest record websites feature similar disclaimers to law enforcement and jail sites.[45] However, commercial sites tend to retain mugshot, arrest, and detention records indefinitely, still

---

42. Martin A Holland, "Note: Identity, Privacy and Crime: Privacy and Public Records in Florida" (2012) 23 University of Florida Journal of Law & Public Policy 235.
43. For instance, see "Johnson County Iowa Jail Roster Disclaimer" *Johnson County Iowa*, online: Johnson County Iowa <www.johnson-county.com/Sheriff/JailRoster/Index> [Johnson County Iowa].
44. In Canada, public disclosure of personal information by the government without the individual's consent is generally prohibited by *Privacy Act*, RSC 1985, c P-21, s 8.
45. Mugshots.com prominently displays such a disclaimer, including (in ALL CAPS) that "[T]HE MUGSHOTS AND/OR ARREST RECORDS PUBLISHED ON MUGSHOTS.COM ARE IN NO WAY AN INDICATION OF GUILT AND THEY ARE NOT EVIDENCE THAT AN ACTUAL CRIME HAS BEEN COMMITTED. ARREST DOES NOT IMPLY GUILT, AND CRIMINAL CHARGES ARE MERELY ACCUSATIONS. A DEFENDANT IS PRESUMED INNOCENT UNLESS PROVEN GUILTY AND CONVICTED. FOR LATEST CASE STATUS, CONTACT THE OFFICIAL LAW ENFORCEMENT AGENCY WHICH ORIGINALLY RELEASED THE INFORMATION". See Mugshots, *supra* note 40.

without updating or correcting incorrect information.[46] For example, Mugshots.com, the most prominent of these sites, calls itself a "Google for Mugshots", states the following:

> [t]he website is a search engine for Official Law Enforcement records, specifically booking photographs, mugshots. Originally collected and distributed by Law Enforcement agencies, booking records are considered and legally recognized as public records, in the public domain. Mugshots.com republishes these Official Records in their original form ("as is") under the First Amendment to the United States Constitution, the freedom to publish true and factual information. Our intent is to provide a legitimate and useful service for both the private and public sectors.[47]

The site recognizes no irony in the disconnect between characterizing the First Amendment as guaranteeing "the freedom to publish true and factual information" and a disclaimer denying responsibility for accuracy. In response to the questions "[m]y record was expunged"; "[I] was pardoned"; "[m]y case was dismissed"; and "[w]ill you remove my mugshot?", the Mugshots.com FAQ page states, "[a]s you may be aware [e]xpungement and pardon only apply to certain government agencies' databases, and not all of them. Certainly not to the private sector".[48] In other words, according to Mugshots.com, whatever balancing the public sector engages in to justify granting a pardon or expungement does not apply to public records that are archived by private entities.

Sites like Mugshots.com capitalize on US First Amendment jurisprudence, which permits further dissemination of truthful

---

46. The duration which arrest records are kept by public offices in US states varies. The Hillsborough County Florida Sheriff's Office posts the following notice: "Arrest information is a Public Record under Florida State Law unless it has been ordered sealed or expunged. Online arrest inquiries are available for adult arrests occurring since January 1, 1995 for which the Hillsborough County Sheriff's Office has an electronic record". See "Arrest Inquiry" *Hillsborough County Sheriff's Office*, online: HCSO <webapps.hcso.tampa.fl.us/ArrestInquiry#>. The Johnson County Iowa Jail Roster only contains names of individuals who are or have been held by the Johnson County Sheriff within the last 48 hours. See Johnson County Iowa, *supra* note 43.
47. "About" *Mugshots*, online: Mugshots <mugshots.com/about.html>.
48. "FAQ" *Mugshots*, online: Mugshots <mugshots.com/faq.html>.

information if it was released by its original custodian, even if the release itself was against the law or public policy.[49] Thus, even if the original source cannot vouch for the accuracy of the information, mugshot websites in the US are currently under no obligation to update inaccurate or outdated information, even in the face of a direct complaint. However, even in the US, commercial mugshot and arrest record websites have come under fire for using a business model whereby individuals can pay a fee to have their profile removed, altered, or updated, prompting some US states to enact legislation that prohibits the use of public records in this sort of business model, especially where the person has not been convicted.[50] Most states do not prohibit it, so Mugshots.com, until at least September 2017, continued to offer "content removal services" through UnpublishArrest.com, which it bills as its exclusive "licensee" to specifically handle removal and editing requests to Mugshots.com. In May 2018, the state of California charged four proprietors of Mugshots.com with extortion, money laundering, and identity theft in relation to this fee-for-removal scheme, and as of this writing the site now simply refuses to remove content at all, standing on the claim to be entitled to

49. *Florida Star v BJF*, 491 US 524 (1989) [*Florida Star*].

50. "Mug Shots and Booking Photo Websites" *National Conference of State Legislatures* (23 October 2017), online: NCSL <ncsl.org/research/telecommunications-and-information-technology/mug-shots-and-booking-photo-websites.aspx>; Bruno, *supra* note 41; Sean P Sullivan, "Mugshot 'extortion' website ban signed by Christie" *NJ.com* (23 July 2017), online: NJ.com <nj.com/politics/index.ssf/2017/07/christie_signs_bill_banning_mugshot_extortion.html>; David Harris, "New law forces websites to pull mug shots of the acquitted" *Orlando Sentinel* (19 June 2017), online: Orlando Sentinel <orlandosentinel.com/news/breaking-news/os-public-records-mugshots-florida-20170619-story.html>.

republish information issued by law enforcement agencies "as is".[51]

In addition to the fee structure, the overarching business model for a site like Mugshots.com is advertising driven. The dynamics of the ads it runs capitalize on both sides of the online personal information market. On one hand, there are prominent ads for Cleansearch.net, which offers to remove results from general search engines and so targets data subjects. On the other hand, there are ads to fee-charging profile compilation services — mostly via search boxes that look like they are merely additional internal search engines to Mugshots.com, but actually bring the searcher to an external site — and so target data seekers. Links lead searchers to BeenVerified.com (which often uses the slogan "This Site's Deep Search Can Reveal More Than Google"), Peoplelooker.com, Instantcheckmate.com, and Truthfinder.com — all of which offer personal information profile compilation for a fee, either per report or as a monthly subscription.[52] Mugshots.com also employs Google AdSense, which delivers sidebar ads tailored to the search history of individual users, regardless of the content of the website.

In 2013, in response to criticism of the business practices of commercial mugshot websites like Mugshots.com, Google implemented a voluntary change to its algorithm to demote name search results linking to such sites; they are not de-listed entirely, but appear lower on the

---

51. Until the scheme was dismantled in late 2017, mugshots.com charged USD$399 to remove, permanently publish, or edit one arrest record. See e.g. *Internet Archive: Wayback Machine* (27 September 2017), online: Internet Archive: Wayback Machine <https://web.archive.org/web/20170927005616/http://unpublisharrest.com/>; *Internet Archive: Wayback Machine* (3 November 2017), online: Internet Archive: Wayback Machine <https://web.archive.org/web/20171103230426/https://chase44.wufoo.com/forms/zr7v2lm1svib3r/>; Cyrus Farivar, "All of Mugshots.com's alleged co-owners arrested on extortion charges" *Ars Technica* (17 May 2018), online: Ars Technica <https://arstechnica.com/tech-policy/2018/05/all-of-mugshots-coms-alleged-co-owners-arrested-on-extortion-charges>; "FAQ" *Mugshots*, *supra* note 48.

52. One-month subscriptions tend to hover just under USD $30. See for instance *Truthfinder*, online: Truthfinder <https://www.truthfinder.help/cost/>.

results list.[53] Searchers are free to choose to go directly to the site and partake in the service and its economy directly, but Google has chosen to make it more difficult for a searcher who is not specifically looking for this sort of information to inadvertently find it. The fix is not foolproof however. For example, using Google to search the uniquely spelled name of a woman whose image is posted on the non-consensual pornography website MyEx.com, along with her state of residence, produces a results list prominently containing links to multiple sites detailing her arrest record, including both law enforcement institutions and Mugshots.com. Further, as noted above, if a person's name is entered followed by the search term "arrest", not only are these sites likely to rise to the top, but the results will include paid AdWords links to commercial public records compilation services like Truthfinder.com.

The public policy commitments related to public access to pre-conviction and non-conviction information, as well as criminal conviction records, vary significantly by jurisdiction, in ways that profoundly shape this market for sensitive personal information.[54] Many scholars have noted the influence of a longer tradition of personality rights protection in continental Europe, which is widely considered to be the backdrop for the current embrace of "the right to be forgotten". Apart from not providing public access to past criminal conviction records, some European countries even forbid public discussion of past criminal convictions by media organizations, including documentary filmmakers attempting to explore historical crimes.[55] In Canada, criminal convictions, pre-conviction status, and non-conviction records are not

53. Barry Schwartz, "Google Launches Fix to Stop Mugshot Sites from Ranking: Google's MugShot Algorithm" *Search Engine Land* (7 October 2013), online: Search Engine Land <searchengineland.com/google-launches-fix-to-stop-mugshot-sites-from-ranking-googles-mugshot-algorithm-173672>.

54. James B Jacobs, *The Eternal Criminal Record* (Cambridge: Harvard University Press, 2015).

55. See discussion of European approach in Franz Werro, "The Right to Inform v the Right to Be Forgotten: A Transatlantic Clash" in Aurelia Colombi Ciacchi et al, eds, *Liability in the Third Millennium* (Baden-Baden: Nomos, 2009) 285 at 290.

freely open to the public; they are housed in a law enforcement database — Canadian Police Information Centre ("CPIC") — and are only made available upon legitimate request, usually with the consent of the data subject (for instance, when a person wants to volunteer in a school). The rationale for these restrictions is based on the principle that such personal information is always sensitive, that ongoing public disclosure is highly likely to negatively affect the individual, and that the inability to shield this information from ongoing public disclosure damages the individual's chances of rehabilitation and reintegration.[56]

Further, Canada makes "record suspensions" available to eligible individuals who apply for them, similar to European jurisdictions, although unlike some European countries, Canada does not prevent the reporting or republishing of information about past crimes. A record suspension (formerly referred to as a pardon) removes a criminal conviction record from the parts of the CPIC database that are available to the public upon legitimate request.[57] Access to the full record is

56. Jeannie Stiglic, "Hard to check criminal records of others: Only legal way is through court documents" *CBC News* (13 January 2012), online: CBC <cbc.ca/news/canada/hard-to-check-criminal-records-of-others-1.1145038>. This is not to say that injustices do not continue to be perpetuated against people who have been in conflict with the law, since many potential employers require police record checks without much justification other than prejudice. See Canadian Civil Liberties Association, "False Promises, Hidden Costs: The Case for Reframing Employment and Volunteer Police Record Check Practices in Canada", by Abby Deshman (Toronto: CCLA, May 2014), online: CCLA <ccla.org/recordchecks/falsepromises>. See also Canadian Civil Liberties Association, "Presumption of Guilt? The Disclosure of Non-Conviction Records in Police Background Checks", by Graeme Norton (Toronto: CCLA, May 2012), online: CCLA <ccla.org/cclanewsite/wp-content/uploads/2015/02/Presumption-of-Guilt.pdf>.

57. Convictions for which a record suspension has been granted may still be released pursuant to a Police Vulnerable Sector Check, which is sought by people seeking employment or volunteering in a position of authority or trust relative to vulnerable persons. For instance, see Ontario Provincial Police, "Criminal Record Checks and Police Checks" (OPP, 26 October 2017), online: OPP <opp.ca/index.php?id=115&entryid=56a1276d8f94acdb5824a3d7>.

retained by police, and public accessibility according to the above noted restrictions can be reinstated if the individual commits another offence.[58] Overall, the US is far less generous in its protection of people with criminal records, and pardons are much more rare — there are some other administrative means of providing limited relief from the burden of having a criminal record, but none of them affect previous, existing, or future publication of the fact of conviction.[59]

The key issue here is ease of access versus obscurity, or put more materially, public accessibility to conviction records upon legitimate request versus accessibility by mere payment of a fee. Further, websites that provide public records can choose whether to allow their contents to be crawled and indexed by general search engines like Google. Most court and tribunal websites, as well as legal information repositories like the various Legal Information Institute sites, offer internal search tools but opt not to permit external search engines to index their content. In Canada, the Office of the Privacy Commissioner ("OPC") ruled complaints against Globe24h, a website based in Romania, to be well-founded.[60] The website had scraped content from Canadian legal information sites, including CanLII, and allowed the reposted court and tribunal documents to be searched by external search engines.[61] One aspect of the Globe24h business model was to charge a fee to individuals

58. For Canada, see Royal Canadian Mounted Police, "Dissemination of Criminal Record Information policy", (RCMP, 24 June 2014), online: RCMP <rcmp-grc.gc.ca/en/dissemination-criminal-record-information-policy>.

59. *Collateral Consequences Resource Center: Collateral Consequence of Criminal Conviction and Restoration of Rights: News, Commentary, and Tools*, online: CCRC <ccresourcecenter.org>; Peter Leasure & Tia Stevens Andersen, "The Effectiveness of Certificates of Relief as Collateral Consequence Relief Mechanisms: An Experimental Study" (2016) 35 Yale Law & Policy Review Inter Alia 11, online: Yale University <www.ylpr.yale.edu/sites/default/files/IA/leasure.certificates_of_relief.produced.pdf>.

60. Office of the Privacy Commissioner of Canada, *Website that generates revenue by republishing Canadian court decisions and allowing them to be indexed by search engines contravened PIPEDA*, *PIPEDA* Report of Findings #2015-002 (Ottawa: OPC, 5 June 2015).

61. *Ibid*.

wishing to have their personal information removed from the site, as well as employing advertising. Globe24h claimed the documents were in the public domain and that it was free to repost the material and to make it more easily accessible to searchers, and refused to comply with the OPC's finding.

One complainant, joined by the OPC, brought the case to the Federal Court, which affirmed the findings of the OPC and held that Globe24h was disrupting the balancing done by courts, tribunals, and publicly accessible legal databases like CanLII, between the open court principle and the privacy interests of people whose personal information appears in these documents.[62] The Court ruled that making court documents searchable by general search engines does not further the interests of the open-court principle that justifies courts and tribunals making information public. Consequently, Globe24h is required to obtain the consent of data subjects in order to republish decisions and documents and make them externally searchable. The Court endorsed the OPC's support of a corrective court order requiring Globe24h to remove Canadian cases containing personal information, to take steps to remove these decisions from search engine caches, and to take steps to ensure that any documents reposted were not indexed by search engines. The Court also granted a declaratory order that the complainant can then take to Google per its voluntary removal policy for court orders. Unlike the CJEU, the Canadian Court was not asked to determine whether general search engines like Google would be required to de-index links to this material coming up in a name search for a data subject.

Globe24h argued that it should qualify for either the journalistic purpose or the publicly available information exemptions to application of Canada's private sector data protection legislation, the *Personal Information Protection and Electronic Documents Act*[63] ("*PIPEDA*"). The Court ruled that Globe24h was not engaging in a journalistic purpose when it republished court and tribunal documents and allowed them to be indexed by search engines, relying on the Canadian Association

62. *AT v Globe24h.com*, 2017 FC 114.
63. *Ibid* at para 29, referring to *PIPEDA*, *supra* note 14.

of Journalists' definition as suggested by the OPC. According to that definition, an activity qualifies as journalism only when: (1) its purpose is to inform the community on issues the community is interested in; (2) the presentation of the information involves an element of original production; and (3) it incorporates a "[s]elf-conscious discipline calculated to provide an accurate and fair description of facts, opinion and debate at play within a situation".[64] Thus, fairness once again provides a core measure for whether personal information is being made more easily publicly accessible in the public interest. The Court also rejected the defendant's efforts to use the "publicly available" exemption, stating the exemption only applies if the defendant's collection, use, or disclosure relates directly to the purpose for which the information appears in the public record or the original source. Again, court and tribunal records or documents are only exempt from further obligations if their republication furthers the open-court principle.[65]

This ruling suggests that general search engines like Google would also potentially be subject to *PIPEDA* in Canada, in that its search results that contain personal information would similarly not meet the criteria for either of these exemptions, though the OPC has not yet taken this stance.[66]

---

64. *Ibid* at para 68.
65. *Ibid* at para 78.
66. In *Google Inc v Equustek Solutions Inc*, the Supreme Court of Canada upheld an interlocutory injunction of worldwide reach against Google, requiring it to de-list the defendant's websites that sold wares in violation of plaintiff's intellectual property rights. The court rejected Google's claim that such an injunction interferes with freedom of expression and international comity, stating that there was no evidence on the record that any jurisdiction across the world would view the particular speech at issue as protected speech (that is, speech aiming to pass off the wares of the plaintiff as the defendant's). Google could apply for a variance if it was able to prove that protected speech was at issue. The case suggests that where there is variance between jurisdictions, that de-listing should be limited geographically to those jurisdictions where de-listing is considered a justified restriction on freedom of expression. See *Google Inc v Equustek Solutions Inc*, 2017 SCC 34 at paras 46–48.

IV. Businesses that Facilitate and Package User-Generated Content

The second major category of material that is to varying degrees public, or more accurately publicly accessible, is user-generated content. This may appear on social networking platforms or through websites serving as a forum for user-posted material. Indeed, Google reports that most of the top 10 sites for which it receives de-listing requests after the CJEU ruling are sites that host user-provided content.[67] The regulation of sites and services that host user-posted content has been controversial, given the widely recognized policy of immunizing hosts from liability for third-party-provided content. The degree of immunity varies by jurisdiction; the EU provides hosts immunity from liability for user-posted content but revokes that immunity if the host does not respond promptly to notice of illegal content, whereas the US provides broad immunity through the *Communications Decency Act*[68] ("*CDA*"), section 230, which imposes no obligation on hosts to respond to complaints about user-posted material. Whether general information location services like Google could (or should) be considered mere hosts or intermediaries of third-party content that turn up in search results, and hence be wholly or partially sheltered from liability, is an open question and would likely be answered differently by the US and EU, with Canada undecided.[69]

The US has struggled with host immunity in its efforts to curtail businesses that specifically profit from user-generated content in the category of non-consensual pornography ("revenge porn") — that is, sites that encourage users to post intimate images for public consumption

67. Google Transparency Report, *supra* note 12. The top 10 sites include social media juggernauts Facebook, YouTube, Twitter, and Instagram (last visited on 5 June 2017).

68. EC, *Directive 2000/31/EC of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce')*, [2000] OJ L 178/1, art 14, online: EUR-Lex <eur-lex.europa.eu/eli/dir/2000/31/oj> [Directive 2000/31/EC]; *Communications Decency Act*, 47 USC § 230 [*CDA*], is the common name for Title V of the *Telecommunications Act* of *1996*.

69. Slane, "Squaring the Remedy", *supra* note 5.

without the consent of the person pictured. While many jurisdictions have now created criminal offences that prohibit posting such images without consent, for the most part these offences do not apply to the host or platform that houses them or else are not enforced against such hosts.[70] In the US, with the exception of host sites whose operators have been found guilty of other offences (*e.g.* hacking, identity theft, or extortion), the operators of online businesses that exploit the criminal acts of users have so far generally been assumed to be sheltered by *CDA*, section 230.[71]

To date, only one site has been investigated and ruled against by the FTC, which found that defendant Craig Brittain, who operated the site IsAnybodyDown, had:

> [u]nfairly disseminated photographs of individuals with their intimate parts exposed, along with personal information about them, for commercial gain and without the knowledge or consent of those depicted, despite the fact that he knew or should have known that the individuals had a reasonable expectation that their image would not be disseminated in that manner.[72]

What the FTC means by "in that manner" is dissemination on commercial or for-profit pornography websites, ordering that Brittain must remove all photos for which he did not have proof of consent and going forward, he must secure proof of consent of the person pictured before allowing a user to post that person's intimate image.[73]

Brittain, like the website operators convicted of criminal offences, also engaged in further unfair and deceptive business practices, such as tricking women into sending him intimate photos by posing as another woman on Craigslist, operating a "bounty system" to facilitate posting of specific people's images, and a fee-for-removal model. Nonetheless, the

---

70. Andrea Slane & Ganaele Langlois, "Debunking the Myth of Not My Bad: Sexual Images, Consent, and Online Host Responsibilities in Canada" (2018) 30:1 Canadian Journal of Women and the Law 42 [Slane & Langlois, "Debunking the Myth"].
71. *CDA*, *supra* note 68.
72. US, Federal Trade Commission, *Analysis of Proposed Consent Order to Aid Public Comment: In the Matter of Craig Brittain*, File No 132 3120, (FTC, 29 January 2015), online: FTC <https://www.ftc.gov/system/files/documents/cases/150129craigbrittainanalysis.pdf>.
73. Slane & Langlois, "Debunking the Myth", *supra* note 70.

FTC includes the more common and not as obviously unfair practice of soliciting users to post intimate images of other people without ensuring consent in its list of unfair business practices.[74] However, this kind of practice continued to be used by other non-consensual pornography sites, despite the FTC ruling against Brittain. For example, the still-operational website MyEx.com, operational until January 2018 when it went offline as part of a settlement with the FTC, invited users to post images of their former lovers, along with identifying information, and disavowed any obligation to ensure users had the photo subject's consent. MyEx.com monetized traffic to and from the site in various ways: in addition to the general Google Analytics tracking tool, MyEx.com employed Advertising.com (a tracker that matches ads with the content and types of users of a website), EroAdvertising (a more specialized targeted advertising tracker for porn-related advertising), and Adult Webmaster Empire (an affiliate program, whereby websites like MyEx.com are compensated for driving traffic onto a range of other commercial porn websites).[75]

In the US, websites like MyEx.com, like other websites that host third-party content, have assumed they are immune from any responsibility regarding material posted by users, under section 230 of the *CDA*. However, this immunity is based on the assumption that such websites are not serving as data controllers that process the personal information of consumers (albeit non-users of the service) when they provide a specific hosting service like this one. Following *Google Spain*, it is clear that the EU takes a different approach, considering business models to be processing personal information even when they are simply compiling and packaging information posted by others.[76] The EU data protection requirements are supplemented by the conditional immunity

74. US, Federal Trade Commission, *In the Matter of Craig Brittain: Complaint* (C-4564) (2016) at para 5.
75. Ganaele Langlois & Andrea Slane, "Economies of Reputation: The Case of Revenge Porn" (2017) 14:2 Communication & Critical/Cultural Studies 120; *US Federal Trade Commission and State of Nevada v EMP Media Inc, et al*, Stipulated Order for Permanent Injuction on Monetary (2018) 2:18-cv-00035 at 5-6.
76. *Google Spain*, *supra* note 2 at para 29.

provided to hosts by the EU E-Commerce Directive, which requires businesses to take down illegal material posted by third parties upon notification.[77] It is unlikely that a non-consensual pornography business would be able to comply with data protection requirements in the EU at all, but at the very least this kind of business would be required to take non-consensually posted intimate material down without charging a fee.

Canada has not formally applied its private-sector data protection regime to non-consensual pornography-hosting websites, although the OPC does claim to have successfully advocated on behalf of complainants to have images taken down.[78] In other online contexts, the OPC has several times imposed data protection obligations on a service provider that allows users to post or otherwise offer up a non-user's personal information; for example, in 2009, the OPC found Facebook to have violated *PIPEDA* with regard to a feature that prompted users to provide the email addresses of people they know who were not yet users of Facebook.[79] The OPC found that there should be "[a] clear distinction between activities conducted by Facebook users for strictly personal reasons and activities in which Facebook itself is involved".[80] To illustrate, the OPC continued:

> [w]hen users post information about non-users to their profiles, Walls, or News Feeds, such postings are made for personal purposes and as such fall outside the purview of the Act. The Act would apply only where Facebook uses non-users'

77. Directive 2000/31/EC, *supra* note 68.
78. Office of the Privacy Commissioner of Canada, "Online Reputation: What are they saying about me?" (Ottawa: OPC, 21 January 2016), online: OPC <priv.gc.ca/en/opc-actions-and-decisions/research/explore-privacy-research/2016/or_201601/> [OPC, "Online Reputation"].
79. Office of the Privacy Commissioner of Canada, *Report of Findings into the Complaint Filed by the Canadian Internet Policy and Public Interest Clinic (CIPIC) Against Facebook Inc Under the Personal Information Protection and Electronic Documents Act*, by Elizabeth Denham, *PIPEDA* Report of Findings #2009-008 (Ottawa: OPC, 16 July 2009) [Denham]. See also Office of the Privacy Commissioner of Canada, "Facebook investigation follow-up complete" (Ottawa: OPC, 22 September 2010), online: OPC <priv.gc.ca/en/opc-news/news-and-announcements/2010/bg_100922/>.
80. Denham, *ibid* at para 306.

> personal information for purposes of its own.[81]

Determination of when a business is using non-user personal information "for purposes of its own" within a business model based on advertising, traffic direction, and close ties to information removal services also varies by jurisdiction; Canada must choose between the approaches used in the EU and the US. The business model of non-consensual pornography websites — *i.e.* solicitation and monetization of sensitive personal information of non-users — should count as using personal information for the business's own purposes. By this logic, the OPC could also consider, as the CJEU did, that search engines like Google specifically profit from search results, although profiting from search of a person's name is less clear.

What is clear is that the OPC considers indexing by search engines to increase the effects of privacy concerns about information posted online, whether that information is public documents as in the Globe24h case described above, or is posted by users. In a 2012 finding against the Canadian youth-oriented social networking site, Nexopia, the OPC found that allowing user profiles and all their contents to be indexed by search engines as a default setting was not within the scope of what a reasonable person would expect from a social networking site, even if, as Nexopia argued, it markets itself as a more outward-facing, public exposure-oriented alternative to Facebook.[82] The OPC recommended that "visible to friends" should be the default privacy setting, and to make it obvious and explicit that choosing "visible to all" would include indexing via external search engines.[83]

While the EU, the US, and Canada clearly use different approaches, all three jurisdictions distinguish between businesses that merely host third-party content, and businesses that assist in creating content that uses personal information of users or non-users as part of its profit-

---

81. *Ibid*.
82. Office of the Privacy Commissioner of Canada, *Social networking site for youth, Nexopia, breached Canadian privacy law*, *PIPEDA* Report of Findings #2012-001 (Ottawa: OPC, 18 February 2013) at para 71.
83. *Ibid* at para 107.

making activity.[84] In response to public pressure to curtail the effects of non-consensual pornography businesses on victims, many mainstream US-based companies have voluntarily chosen to make it easier for these data subjects to successfully request removal of intimate images they did not consent to have publicly posted, including Reddit, Facebook, Twitter, Microsoft, and Google.[85] Facebook recently announced it would employ a photo identification system to block the reposting of such images it

84. *Fair Housing Council of San Fernando Valley v Roommates.com, LLC*, 521 F (3d) 1157 (9th Cir 2008) (US); Mary Anne Franks, "The Lawless Internet? Myths and Misconceptions About CDA Section 230", *Huff Post* (blog) (18 December 2013), online: Huffington Post <huffingtonpost.com/mary-anne-franks/section-230-the-lawless-internet_b_4455090.html>.

85. All of these services announced new policies in 2015 regarding take-down of intimate images housed on their services upon complaint by the victim. Andrea Peterson, "Reddit is finally cracking down on revenge porn" *The Washington Post* (24 February 2015), online: Washington Post <washingtonpost.com/news/the-switch/wp/2015/02/24/reddit-is-finally-cracking-down-on-revenge-porn/?utm_term=.df3926415b93>; Hayley Tsukayama, "Twitter updates its rules to specifically ban 'revenge porn'" *The Washington Post* (11 March 2015), online: Washington Post <www.washingtonpost.com/news/the-switch/wp/2015/03/11/twitter-updates-its-rules-to-specifically-ban-revenge-porn/?utm_term=.46ee8ea4384f>; Vindu Goel, "Facebook Clarifies Rules on What It Bans and Why" *New York Times: Bits* (blog) (16 March 2015), online: NY Times <bits.blogs.nytimes.com/2015/03/16/facebook-explains-what-it-bans-and-why/?mcubz=0>; Alyssa Newcomb, "How Microsoft Is Waging War Against Revenge Porn" *ABC News* (23 July 2015), online: ABC <abcnews.go.com/Technology/microsoft-waging-war-revenge-porn/story?id=32639751>; Jeff John Roberts, "Google to remove 'revenge porn' links at victims' request" *Fortune* (19 June 2015), online: Fortune <fortune.com/2015/06/19/google-revenge-porn-removal/>.

had already taken down.[86] In the US, these are voluntary policies and are limited to non-consensual pornography, nonetheless these voluntary policies are efforts to distinguish ethical platforms from unethical ones, where the former engage in striking a normatively fair balance between their incentives to make information easily accessible and the interests of people whose personal information is circulating.

Harnessing privacy invasion for profit via the attention economy drives many online business models. If this economy is to be fair, then an appropriate balance is needed between competing stakeholder interests, a balance that considers the sensitivity of the information (often correlated with harm or risk of harm to the data subject), and the public interest in easy access to that information.

V. Viral Content: When Online Personal Information Becomes Part of Public Culture

The public interest in access to content that includes the personal information of others is malleable, especially in an online context where viral distribution of some online material may render it a part of public culture. However, here too balancing of interests — by way of an analysis akin to newsworthiness — can help determine whether virality is sufficient to justify ongoing easy access to that content.

In the attention economy, the "subculture of humiliation" ensures

86. Matt Burgess, "Facebook is using photo-matching to tackle 'revenge porn'" *Wired* (6 April 2017), online: Wired <wired.co.uk/article/facebook-revenge-porn-tools>. In the Facebook Moderation Guidelines leaked to the press in May 2017, the internal Facebook document stated that Facebook had flagged more than 50,000 posts as related to non-consensual intimate imagery and sextortion in the month of January 2017 alone. The guidelines set out an escalation and removal protocol. See Nick Hopkins, "Revealed: Facebook's internal rulebook on sex, terrorism and violence" *The Guardian* (21 May 2017), online: The Guardian <www.theguardian.com/news/2017/may/21/revealed-facebook-internal-rulebook-sex-terrorism-violence>; see especially "What Facebook says on sextortion and revenge porn" *The Guardian* (22 May 2017), online: The Guardian <www.theguardian.com/news/gallery/2017/may/22/what-facebook-says-on-sextortion-and-revenge-porn>.

that public circulation of sensitive personal information can be especially profitable for online businesses. For example, consider materials that mock or otherwise harass a person with disabilities. One of the earliest and most controversial cases of this sort involved an Italian court's conviction of three Google executives for criminal privacy invasion in 2010, charges that arose from users posting a video of an autistic boy being physically bullied to Google Video (its video-sharing platform prior to its purchase of YouTube).[87] The decision was widely criticized as misconstruing service provider obligations both in the US and in Europe, and the decision was overturned by an Italian appellate court in 2012. The Court of Appeals found that Google served as a host and had no obligation to monitor user postings, and had responded promptly by removing the video once expressly notified.[88]

While not discussed in the case, had Google not removed the video upon being notified, it likely would have been liable for the criminal offences charged, and also subject to data protection obligations related to the sensitive personal information of the autistic boy pictured in the video. Following *Google Spain*, Google would have been found to be a "data controller" profiting from the exploitation of this video; in the two months in which it was publicly available, Google Video algorithms had ranked the video highly in the "funny video" category and the Google AdWords service had automatically associated specific search terms with the video.[89] In other words, Google collected profits from the public

87. Manuela D'Alessandro, "Google executives convicted for Italy autism video" *Reuters* (24 February 2010), online: Reuters <www.reuters.com/article/us-italy-google-conviction-idUSTRE61N2G520100224>; Ernesto Apa & Oreste Pollicino, *Modeling the Liability of Internet Service Providers: Google vs Vivi Down, A Constitutional Perspective* (Milan: Egea, 2013).

88. "Court of appeals overturns conviction of Google Italy executives, redefines liability of hosting providers under privacy legislation" *Lexology* (26 March 2013), online: Lexology <www.lexology.com/library/detail.aspx?g=b36ffdc4-ee2b-4dfb-ae83-01bcb15ff5f7>.

89. Bruno Carotti, "The *Google — Vivi Down* Case: Providers' Responsibility, Privacy and Internet Freedom" in Sabino Cassese et al, eds, *Global Administrative Law: The Casebook* (Institute for Research on Public Administration, 2012) 117.

availability and popularity of this video via an advertising model that capitalized on people searching for and viewing it.

The new *GDPR* in the EU would likely further require a hosting platform like Google Video to remove the video on request as part of the "right to be forgotten". The US consumer protection regime surely would not, because Google had no hand in creating or posting the video, nor did it specifically solicit this type of content, unlike the common practice on non-consensual pornography sites. The voluntary moderation guidelines leaked from Facebook in 2017 also reveal that photos mocking people with disabilities have until recently not been considered the kind of material that should be removed (the Facebook guidelines even included an image of a person with Down Syndrome as an example).[90] In other words, mocking people with disabilities is deemed a matter of freedom of speech, offensive but protected, although it is unclear in the guidelines and the discussion of them whether a request from the person pictured (or his or her guardian) would prompt a different action from Facebook than a general user's complaint about an objectionable image of an unknown person with disabilities.[91]

Identification is an aggravating factor in privacy invasion, and an online image of an identifiable person (*e.g.* showing a face) becomes something else entirely when that image is associated with a name. Images of an identifiable person may still contain sensitive personal information

---

90. Nick Hopkins, "Revealed: Facebook's internal rulebook on sex, terrorism and violence" *The Guardian* (21 May 2017), online: The Guardian <https://www.theguardian.com/news/2017/may/21/revealed-facebook-internal-rulebook-sex-terrorism-violence>; Julia Angwin & Hannes Grassegger, "Facebook's Secret Censorship Rules Protect White Men From Hate Speech But Not Black Children" *ProPublica* (28 June 2017), online: ProPublica <https://www.propublica.org/article/facebook-hate-speech-censorship-internal-documents-algorithms>.

91. Nick Hopkins, "How Facebook allows users to post footage of children being bullied: Leaked guidelines on cruel and abusive posts also show how company judges who 'deserves our protection' and who doesn't" *The Guardian* (22 May 2017), online: The Guardian <https://www.theguardian.com/news/2017/may/22/how-facebook-allows-users-to-post-footage-of-children-being-bullied>.

(as with the autistic boy), but if the image is publicly associated with the name of a specific person, the degree of invasiveness is magnified. This distinction was described by Ghyslain Raza, a then 14-year-old boy who gained unwanted internet fame as the "Star Wars Kid" beginning in 2003, when a video he privately recorded of himself wielding a pretend lightsaber was found and posted by mocking classmates. He noted that it was only when his name was released by a media organization that the harassment became much worse, opening him up not only to bullying by people he already knew offline (his schoolmates) but also to random unknown individuals online.[92] So while many people have argued that the "Star Wars Kid" video entered public culture, along with its many benign user-generated variations, it is much more difficult to argue that the video and its variations should continue to be associated with Raza's name.[93]

This brings us back to the issue of name search results in search engines, and the way that Google, after the *Google Spain* decision, now distinguishes between requests to delist news articles that are, or are not, associated with a person. Even the newsworthiness of an article published by a dedicated news site wanes as time goes on if the individual named therein is no longer in the public eye.[94] Google lists 23 examples of news articles that were requested to be delisted and the decision it made in relation to each; in the 11 examples where Google granted the delisting, most dealt with articles referring to minor crimes, quashed convictions,

---

92. Rebecca Hawkes, "Whatever happened to Star Wars Kid? The sad but inspiring story behind one of the first victims of cyberbullying" *The Telegraph* (4 May 2016), online: The Telegraph <telegraph.co.uk/films/2016/05/04/whatever-happened-to-star-wars-kid-the-true-story-behind-one-of/>; Jonathan Trudel, "Return of the 'Star Wars Kid'", *Maclean's* (27 May 2013) 28.
93. Meg Leta Ambrose, "You Are What Google Says You Are: The Right to be Forgotten and Information Stewardship" (2012) 17:07 International Review of Information Ethics 21; Limor Shifman, "An anatomy of a YouTube meme" (2012) 14:2 New Media & Society 187.
94. Meg Leta Ambrose, "It's About Time: Privacy, Information Life Cycles, and the Right to Be Forgotten" (2013) 16:2 Stanford Technology Law Review 369.

crime victims or relatives of crime victims who were not public figures, as well as an article naming a contestant in a contest who was a minor at the time.[95] Of the 12 examples where delisting was not granted, most involved crimes by public officials, serious crimes, serious accusations regarding people in positions of trust, professional misconduct/discipline, fraud, and one example of a "media professional" requesting removal of links to articles reporting on embarrassing content he had posted himself.[96] Together, these examples show that Google has attempted to establish guidelines regarding what sort of personal information contained in news stories remains in the public interest enough to warrant ongoing public association with a person's name, and what does not.

While internet service providers, including Google, have not yet had to deal with the stronger right of erasure in the new *GDPR*, the key will be proportionality in balancing the public interest in access to information that has entered into public circulation against the ongoing privacy interests of the individuals named or otherwise identified. In some cases, an image — like the "Star Wars Kid" video — might acquire the status of a shared cultural document, the factual content of which is not particularly sensitive. In most others, determining whether delisting or deindexing is the most appropriate way to address the privacy interest of the subject will depend on both the degree of sensitivity of the information revealed (a child's autism-related reaction to physical confrontation is clearly more sensitive than a child's goofy playacting) and the degree to which the document in which the information appears has acquired or maintained newsworthiness (as distinguished from prurient or morbid curiosity)[97] or the public culture equivalent thereof. Widespread creative adaptation of a popular culture meme weighs in favour of keeping the Star Wars Kid video available, although it should be disassociated from the young man's

95. California passed a bill providing a means for minors to remove material they have posted themselves. See US, SB 568, *An Act to Add Chapter 22.1 (Commencing with Section 22580) to Division 8 of the Business and Professions Code, Relating to the Internet*, 2013–14, Reg Sess, Cal, 2015 (enacted).
96. Google Transparency Report, *supra* note 12.
97. Worley, *supra* note 13.

name unless he chooses otherwise, while mean-spirited humour found in the humiliation of a person with disabilities does not.

VI. Publicly Available ≠ Free for the Taking

Policies regarding how and whether to constrain businesses that profit from access to publicly available documents on the public internet have implications for how to regulate "big data", another front on which the privacy interests of data subjects may clash with business interests in monetizing publicly accessible information. These same authorities are beginning to question the idea that, although people leave behind a trail of information wherever they go and whatever they do online, this information is free for the taking. However, as with publicly accessible information packaged for open public consumption by information location services, to date regulators have only targeted the most egregious business practices.

For example, the FTC's 2015 decision and order against the website, Jerk.com, found the site operators to be engaging in unfair and deceptive business practices related to harvesting profile content from Facebook via an application program interface ("API") that allowed third-party application developers to access even content that was set to be shared only with "friends".[98] The operators of Jerk.com claimed that their content was created by their users, when in fact it was largely created by the operators themselves, from personal information scraped from Facebook and other "publicly accessible" sources, many of which contained full names and images, buttons for users to vote whether or not the person was a "jerk", and fields for users to fill in further information about that person.[99] These profiles were then made available for indexing by general search engines.[100] Jerk.com's business model included selling USD $30 memberships, requiring a USD $25 "customer service fee" to

98. US, Federal Trade Commission, *In the Matter of Jerk, LLC and John Fanning: Complaint* (No. 9361) (2014) at paras 7, 10–11 [Jerk.com Complaint].
99. US, Federal Trade Commission, *In the Matter of Jerk, LLC and John Fanning: Opinion* (No. 9361) (2015).
100. Jerk.com Complaint, *supra* note 98 at para 9.

communicate with administrators, and third-party advertising.[101]

The Respondents claimed that their enterprise amounted to speech protected by the First Amendment because the Facebook photos and profile information were "publicly available" and that Facebook was to blame for making that material accessible. In other words, once Facebook failed to ensure that its users' private information was protected, any developer could use that information however they chose.[102] Indeed the Respondent tried to argue that the First Amendment was implicated because the FTC's order impinges on "[j]erk.com posting publicly available information derived from the internet".[103] The FTC (and the US Court of Appeal that upheld its decision) rejected that claim, in essence finding that Jerk.com misrepresented the source of its profile content, thereby misleading consumers as to how it had obtained it.[104] The ruling is narrow, in that it does not directly deal with the problem of whether a business that exploits a technological weakness that renders personal information publicly accessible gains the right to process or package it in whatever way it pleases, provided that they are honest with consumers about the source.[105] It is unclear, then, whether the First Amendment would protect the right to publish personal information scraped from the internet via a security weakness, given the seminal 1989 US Supreme Court freedom of speech decision in *Florida Star v BJF*,[106] where a newspaper was permitted to defy restrictions on publicizing a rape victim's identity because police had been negligent in including her name in a police report. In that case, the onus on protecting sensitive personal information was placed entirely on the public authority that improperly released it; the distinction with the Jerk.com case could

101. *Ibid* at para 5.
102. Trial Brief of Respondent John Fanning in *Fanning v Federal Trade Commission* (No 15-1520) (2016) at 3, stating "nothing prohibited the publication on jerk.com information made accessible to the public by Facebook through the internet".
103. *Ibid*.
104. Jerk.com Complaint, *supra* note 98 at para 10.
105. US, Federal Trade Commission, *Fanning v Federal Trade Commission* (No 15-1520) (2016).
106. *Florida Star*, *supra* note 49.

come down to the difference between newsworthy material held by a public entity that media organizations utilized precisely as news upon its improper public release, versus private, non-newsworthy material that is improperly accessible and that has been utilized for a commercial purpose devoid of public interest.

In Canada, the OPC has made stronger statements about inappropriate exploitation of public access to personal information when it conducted an investigation into Google's data collection practices for its location-based services (Google Maps), where Google was discovered to have collected a significant amount of "payload data" from unencrypted WiFi networks in the course of the data-gathering operations of its Street View cars.[107] These data included the full names, telephone numbers, and addresses of many Canadians, as well as complete email messages, email headers, IP addresses, machine hostnames, and the contents of cookies, instant messages, and chat sessions.[108] While Google claimed that the data collection was inadvertent, the OPC nonetheless took the opportunity to stress that even if a WiFi network is unencrypted and therefore publicly accessible, that does not mean any private data travelling across that network are free for the taking:

> [n]otwithstanding the fact the personal information collected was sourced from unprotected networks (and was in some cases fragmented), it is impossible to conceive that a reasonable person would have considered such collection appropriate in the circumstances.[109]

What a reasonable person considers appropriate in the circumstances is the formula for determining commercial fairness in handling personal information set out in *PIPEDA*.[110] Further, Canadian constitutional protection for freedom of expression allows more restrictions regarding publication of sensitive personal information held by public authorities, even if it is "newsworthy" in the way that is understood in the US. The

---

107. Office of the Privacy Commissioner of Canada, *Google Inc WiFi Data Collection*, *PIPEDA* Report of Findings #2011-001 (Ottawa: OPC, 6 June 2011).
108. *Ibid* at para 17.
109. *Ibid* at paras 18, 21.
110. *PIPEDA*, *supra* note 14, ss 3, 5.

names of sexual assault victims, for instance, are routinely made subject to publication bans, even though their names are available to the public via court proceedings.[111] In other words, Canada does not place the onus only on data custodians to keep personal information from the public. Instead, Canada has mechanisms in place to impose obligations on publishers who have had access to that information where the sensitivity of the information warrants it, viewing such restrictions as justified in a free and democratic society: in other words, a fair restriction in grander terms.[112]

Many privacy scholars have expressed grave concerns about the ways that businesses are exploiting publicly accessible personal information, especially considering how little information these businesses make available about exactly how their information collection and packaging algorithms function.[113] Julie Cohen coined the term "biopolitical public domain", referring to the popular idea that all data are fair game and can be collected freely, which she sees as employing a skewed sense of the concept of "public domain" that operates in a more well-developed fashion in intellectual property law.[114] She argued that we need to develop a more robust notion of what belongs in the "data commons" with regard to the practices of information aggregators and processors, so as to better protect personal information even in the realm of publicly accessible raw or de-identified data.[115]

---

111. *Criminal Code*, RSC, 1985, c C-46, s 486.4.
112. *Canadian Newspapers Co v Canada (AG)*, [1988] 2 SCR 122.
113. Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Cambridge: Harvard University Press, 2015); Tauel Harper, "The big data public and its problems: Big Data and the structural transformation of the public sphere" (2017) 19:9 New Media & Society 1424.
114. Julie Cohen, "The Biopolitical Public Domain: the Legal Construction of the Surveillance Economy" (2017) Philosophy & Technology 1, online: Springer <link.springer.com/content/pdf/10.1007%2Fs13347-017-0258-2.pdf>.
115. *Ibid* at 12.

VII. Conclusion: Data Privacy in "Public"

Privacy scholars have begun to explore the various ways that publicly accessible information is being collected and used, both by public and private entities.[116] Unfair handling of publicly accessible personal information has a particularly potent adverse affect on historically or situationally vulnerable populations, further amplifying the urgency of a fairness-based approach to businesses that deal in such information. Public records that are easily accessible have the potential to be misused, disproportionally affecting the reputations and corresponding opportunities of members of historically marginalized groups, such as economically disadvantaged persons and historically persecuted ethnic minorities, as well as individuals who are vulnerable as a result of adverse life events. User-generated content is also more likely to disproportionately affect historically marginalized groups online, mainly due to the "subculture of humiliation" where users post derogatory, harassing information, often about disempowered groups (women, the poor, ethnic minorities, and persons with disabilities).[117] As Frank Pasquale wrote:

> [n]ew threats to reputation have seriously undermined the efficacy of health

116. Chris Jay Hoofnagle, "Big Brother's Little Helpers: How ChoicePoint and Other Commercial Data Brokers Collect and Package Your Data for Law Enforcement" (2003) 29:4 North Carolina Journal of International Law & Commercial Regulation 595; Danah Boyd & Kate Crawford, "Critical Questions For Big Data" (2012) 15:5 Information, Communication & Society 662; Kate Crawford & Jason Schultz, "Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms" (2014) 55:1 Boston College Law Review 93 at 101; Levin, *supra* note 29; Amy Conroy & Teresa Scassa, "Promoting Transparency While Protecting Privacy in Open Government in Canada" (2015) 53:1 Alberta Law Review 175; Ramona Pringle, "'Data is the new oil': Your personal information is now the world's most valuable commodity" *CBC News* (25 August 2017), online: CBC <cbc.ca/news/technology/data-is-the-new-oil-1.4259677>.

117. OPC, "Online Reputation", *supra* note 78, citing Nicolaus Mills, "Television and the Politics of Humiliation" (2004) 51:3 Dissent 79 at 79; Emily B Laidlaw, "Online Shaming and the Right to Privacy" (2017) 6:1 Laws 1.

> privacy law, credit reporting, and expungement. The common thread is automated, algorithmic arrangements of information, which could render a data point removed or obscured in one records system, and highly visible or dominant in other, more important ones … [it] is not much good for an ex-convict to expunge his juvenile record, if the fact of his conviction is the top Google result for searches on his name for the rest of his life. Nor is the removal of a bankruptcy judgment from a credit report of much use to an individual if it influences lead generators' or social networks' assessments of creditworthiness, and would-be lenders are in some way privy to those or similar reputational reports.[118]

However, some scholars do not draw a parallel between business use of publicly accessible information and the kind of activities that search engines or other information location and packaging services do. For example, Neil Richards and Woodrow Hartzog noted that "[m]ost people are vastly less powerful than the government and corporate institutions that create and control digital technologies and the personal data on which those technologies run".[119] However both see the EU's "right to be forgotten" as a serious threat to online freedom of expression and access to information, which could create "[a]n internet that could be edited like Wikipedia by individuals who do not like the facts reported about them in newspapers".[120]

Freedom of expression remains an important component to determining when the privacy interests of data subjects should or should not prevail over public interest in access to an individual's personal information, whether commercial or not. But it is worth remembering that Google is a huge and diverse company, and that while *Google Spain*

---

118. Frank Pasquale, "Reforming the Law of Reputation" (2015) 47:2 Loyola University of Chicago Law Journal 515 at 516.
119. Neil Richards & Woodrow Hartzog, "Privacy's Trust Gap: A Review", Book Review of *Obfuscation: A User's Guide for Privacy and Protest* by Finn Brunton & Helen Nissenbaum, (2017) 126:4 Yale Law Journal 1181 at 1183.
120. *Ibid* at 1185; see also Neil M Richards, *Intellectual Privacy: Rethinking Civil Liberties In The Digital Age* (New York: Oxford University Press, 2015) at 90–92; Woodrow Hartzog, "A Stronger 'Online Eraser' Law Would Be a Mistake" *New Scientist* (6 November 2013), online: New Scientist <newscientist.com/article/mg22029420-200-a-stronger-online-eraser-law-would-be-a-mistake>.

is a decision that only affects its public search engine business, its parent company, Alphabet, is rapidly diversifying in a way that will make it increasingly difficult to separate out revenue derived from advertising linked to search results and revenue derived from data analytics more generally (*e.g.* connections between AdWords, AdSense, and YouTube, Google Maps, Gmail, Google Drive, and Google Play). What we decide to do in terms of characterizing information location and packaging services as either first and foremost business ventures, or as guardians of publicly available information, will affect regulations about the big data analytics industry and privacy going forward. Algorithms and other forms of machine learning and processing have inherent errors and biases. Therefore, imposing data protection obligations on businesses that use them to collect and package publicly accessible personal information can serve as a useful, if limited, means of addressing one variant of the machinations of informational power online.

Overall, however, all personal data collection, processing and packaging should be subject to an analysis rooted in fairness, regardless of whether that information is publicly accessible. That is, fairness requires an appropriate balance between competing interests, where the sensitivity of the information must be taken into account, including disproportionate impact on vulnerable populations, in order to determine what is a fair business practice in the ever-changing information marketplace.