NBER WORKING PAPER SERIES

THE REALITIES OF DISCLOSURE RISKS IN THE AGE OF DARK PATTERNS AND BIG DATA

Ramon Abraham A. Sarmiento

Working Paper 32926
http://www.nber.org/papers/w32926

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2024

The research presented in this article was conducted as part of the author's employment with the Bangko Sentral ng Pilipinas (Philippine Central Bank). Nonetheless, the views and opinions expressed in this paper are those of the author and not the Bangko Sentral ng Pilipinas. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

The author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w32926

The Realities Of Disclosure Risks In The Age Of Dark Patterns And Big Data
Ramon Abraham A. Sarmiento
NBER Working Paper No. 32926
September 2024
JEL No. K29, K39

ABSTRACT

This paper explores the realities of disclosure risks in the current big data landscape, examining implications for individuals, society, and the evolving ethical landscape. One these realities is pervasive use of dark patterns in data collection. Dark patterns are user interfaces crafted to trick users into doing things against their self-interest such as compromising their privacy. This has sparked significant ethical and legal debates on balancing confidentiality & privacy obligations with the need for precise research data. Thus, this paper seeks to shed light on the legal and ethical dilemmas of these disclosure risks arising from the intersection of data privacy, statistical data usage, and the employment of dark patterns, such as the possible insufficiency of existing data protection measures and their possible obsolesce in the age of big data.

Ramon Abraham A. Sarmiento
Bangko Sentral ng Pilipinas
Room 514, 5-Storey Bldg, BSP Main Complex
A Mabini St., Malate, Manila
Philippines
SarmientoRA@bsp.gov.ph

1. Introduction:

In 2012, Steve Lohr of the New York Times welcomed the general public to the age of Big Data (Lohr 2012). He touted many of the advances and changes, that we barely notice today, such as the rapid gathering of information by social networks and their processing of said information in new ways to gain increasingly valuable insights. Lohr did not mention that some of this gathering of data may be harmful such as emergence of dark patterns for the massive collection of personal information. Prior to the emergence of big data into the public eye, Solove already pointed out that we leaves traces of information whatever we do and wherever we go (Solove 2004). The emergence of big data, means that it is more likely that these information traces will be collected, shared, processed and in a sense monetized, which makes preserving privacy much harder.

2. Literature Review

2.1 Dark Patterns

Harry Brignull coined the term dark patterns in 2010. He defined dark patterns as "*A Dark Pattern is a manipulative or deceptive trick in software that gets users to complete an action that they would not otherwise have done, if they had understood it or had a choice at the time. For example if you have a button that functions as a "Yes" when clicked, but through the use of placement, colour and trick wording, it appears to say "No", then many users are going to be* caught out*."* Mathur et. al. (2021) in their review of literature on dark patterns have identified at least 19 different definitions of dark patterns. Beyond, these definitions and conceptual frameworks, the main purpose of a dark pattern is to manipulate people to pay for something they otherwise would not purchase or surrender personal information they would otherwise keep confidential (Luguri & Strahilevitz, 2021). For this paper, we focus on the surrender of personal information.

Kelly & Burkell (2023) have found that privacy dark patterns prevent users from making conscious, informed decisions about the management of their personal data which in turn exposes them to risks and

harms. A large-scale experiment conducted by on census-weighted samples of American adults showed that mild dark patterns more than doubled the percentage of consumers who signed up for a dubious identity theft protection service, and aggressive dark patterns nearly quadrupled the percentage of consumers signing up (Luguri & Strahilevitz 2021).

Numerous other studies have shown the effectiveness of dark patterns in manipulating individuals into surrendering their personal information such as the study concluded that dark patterns and privacy intrusive defaults, would nudge users of Facebook and Google, and to a lesser degree Windows 10, toward the least privacy friendly options to an unethical degree (Forbrukerrådet, 2018a). Utz et. al. (2019) in their analysis of a random sample of 1,000 Consent Management Platforms (CMPs) found that at least 57.4% used dark patterns to nudge users to select privacy-unfriendly options, and that 95.8% provide either no consent choice or confirmation only. Nouwens et. al, (2020) also found less than 12% of the design and text of the CMPs used by the top 10,000 websites in the UK to be compliant with EU law. Tahaei et. al (2022) showed that 77% of free apps have an ad library, which in turn allows advertisers to collect and track user data across a wide variety of applications. The defaults configurations of these ad networks are set to maximize data collection and may have dark patterns (Mhaidli et. al. 2019). In fact, a review of 4 popular ad networks showed the use of dark patterns to nudge developers to choose personalized ads and share more data, rather than giving developers a chance to make informed choices (Tahaei and Vaniea 2021). Di Geronimo et. al. (2020) found that 95% of 240 popular mobile applications contained one or more dark patterns and, on average. Popular applications contained an average 7.4 dark patterns. Finally, Hidaka et. al. (2023) found that 93% of 200 popular Japanese mobile applications contained one or more dark patterns.

Dark patterns pervasiveness and impacts may be still understated. There has been little attention paid to their presence on the internet of things (IOT) as most of the studies on dark patterns have focused on

websites and applications. Kowalczyk et. al. (2023) found at least 3 unique dark patterns in all 57 Internet-of-Things (IoT) devices studied and on each IoT devices contained an average of over 25 dark patterns. The pervasiveness of these dark patterns is illustrated by a recent study conducted by Marti et. al. (2024), which that each volunteer had an average of 2,230 companies sharing their data with Facebook and some volunteers had over 7,000 companies providing their data. The relationship between dark patterns and disclosure risks needs to be closely examined given the prevalence of large private databases containing personal information coupled with growing sophistication of algorithms to match those databases to data released by the government and research institutions (Hotz et. al. 2022).

2.2 Disclosure Risks

Disclosure refers to identification which is the association of a known individual with a particular microdata record or attribution which is the association of information in a micro-dataset with a particular individual. Although disclosure may seem difficult for an isolated dataset with supposedly "anonymized" data, re-identification can be performed by linking public records such as voter registration records, with de-identified data (Dwork et. al. 2017). In a highly cited study, Rocher et. al. (2019) found that 99.98% of Americans would be correctly re-identified in any dataset using 15 demographic attributes. These demographic attributes are contained in thousands, possibly tens of thousands of databases, many of which can be easily purchased or leased.

Previous disclosure avoid measures relied on protecting data with noise by imagining what 'sensitive' values an attacker would want to target, attack methods and databases that would reasonably be used (Oberski & Kreuter 2020). Groshen & Goroff (2023) have pointed out that many previously adequate disclosure avoidance procedures now leave people at risk for re-identification due to the ease of finding personal information due to the internet.

There are numerous studies that have re-identified in whole or in part, supposedly anonymous data sets ranging from the New York Times revelation of the identity of AOL user no. 4417749 to Narayanan &

Shmatikov (2008) using a de-anonymization methodology for sparse micro-data to successfully identify the known Netflix users in a dataset publicly released by Netflix. Even datasets with strong anonymization measures are not immune. Abowd et. al. (2023) using only published data, found that an attacker could verify all records in 70% of all census blocks, equivalent to 97 million people for U.S. 2010 census.

3. The Ethical and Legal Frameworks

Hasan et. al (2020) have pointed out that privacy and protection of data is one the biggest critical issue of big data services. In addition to any legal obligations to do so, many researcher routinely promise anonymity to subjects who participate in studies or surveys (Heffetz & Ligett). These promise of anonymity hold immense value to respondents, as a survey by Hotz & Slanchev (2017) showed that 79.5% of respondents cited confidence in researchers to keep responses and information private" as the most important determinant for their participation in a hypothetical study.

As a recognition of disclosure risks, datasets may be so extensively altered to limit disclosure risks (e.g. noise added), subject to onerous conditions for release or simply data may not be released at all (Muralidhar & Palk 2020; Hotz et. al. 2022; and Groshen & Goroff 2023). These conditions or limitations, constrict the spread of the data and presumably prevent valuable innovations or insights from occurring, thus begging the question, why did we bother collecting the data at all, if we are not going to use it for societal progress (Oberski & Kreuter 2020). Thus, a close look at the interaction between current and proposed regulations on dark patterns and disclosure risks is in order.

3.1 Minimization of Disclosure Risks

The United Nations General Assembly adopted the Fundamental Principles of Official Statistics in January 2014. Principle 6 of which states: *"Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes."* To minimize disclosure risks for confidential data, a variety of obligations can be imposed on a researcher such as the requirement for special sworn status by the US

Census Bureau under Title 13, Section 23 of the U.S. Code. Ruggles et. al. (2018) critiques the process for this confidential access stating that among others, that most research topics ineligible since they are approved only if they benefit the Census Bureau, and the process is time-consuming and costlier than using public use data.

The European Union (EU) also imposes a legal obligation on European Statistical System members to protect confidential data as contained in Chapter V of Commission Regulation (EC) No 223/2009 on European statistics. Like the US Evidence Act, the EU has also implemented the European Data Governance Act (EU DGA). The EU DGA is a key pillar of the European Strategy for Data**.** A key feature of the EU DGA is the concept of data altruism, which is intended to allow easier sharing data based on their consent for purposes of general interest by individuals and businesses. Data altruism organizations will be monitored and supervised by the authorities competent for the registration of altruism organizations, established by each member EU state.

These obligations to keep personal information confidential run in parallel to the various laws to protect personal information and privacy rights. Oberski & Kreuter (2020) enumerated several examples of legislative measures to increase the protection of personal information such as the 2018 European General Data Protection Regulation (GDPR), the 2017 Japanese Amended Act on the Protection of Personal Information the 2020 Brazilian General Data Protection Law and the 2020 California Consumer Privacy Act (CCPA), among others. The list is growing and the United States which already has an arguably robust patchwork for the protection of personal information, is looking into enacting a comprehensive Federal law on data protection or the American Privacy Rights Act.

Philips et. al. (2017) notes that the recognition of the likelihood of reidentification, due in part to big data, has raised calls for the criminal penalties for unauthorized re-identification of anonymized information. At least two jurisdictions that impose criminal penalties on the re-identification of anonymized information. Section 48F of the Singapore's Personal Data Protection Act 2012 criminalizes the

unauthorized re-identification of anonymized information. Similarly, Section 171 of the United Kingdom's Data Protection Act 2018 makes it an offence for a person knowingly or recklessly to re-identify information that is de-identified personal data without the consent of the controller responsible for de-identifying the personal data.

3.2 Regulation of Dark Patterns

Dark patterns have been recognized as a distinct concept for over a decade, but it is only recently that legislation has been passed to regulate them. The Organisation for Economic Co-operation and Development (OECD) in its paper on dark commercial patterns (OECD 2022) doubts if market forces can address dark patterns alone and points out that at times the markets incentivize their use. In the U.S.A., states such as California regulate dark patterns. Section 1798.140.l of California Consumer Privacy Act (CCPA) defines dark patterns as a "user interface designed or manipulated with the substantial effect of subverting or impairing user autonomy, decision making, or choice, as further defined by regulation." The Connecticut Data Privacy Act, the Colorado Privacy Act, and the Vermont Data Privacy Act also regulate dark patterns. Prior to specific legislation enacted to combat dark patterns, the FTC has used the Federal Trade Commission Act, 15 U.S.C. § 57a(a)(1)(b) prohibiting unfair or deceptive trade practices, as the basis for penalizing entities that make use of dark patterns. The FTC's September 2022 report, *Bringing Dark Patterns to Light*, details many of these cases.[1]

In the EU, dark patterns are not specifically mentioned in the GDPR. However they may still violate the fairness and transparency principle in article 5(1)(a), the accountability principle in article.5(2), data protection by design and default in article 25, the requirement to provide transparent privacy notices to data subjects in articles 12(1), 13 & 14), and the data subject rights in articles 15 to 22 of the GDPR. In 2023, the European Data Protection Board ("EDPB") adopted Guidelines 03/2022 on Deceptive Design

[1] https://www.ftc.gov/reports/bringing-dark-patterns-light

Patterns in Social Media Platform Interfaces: How to Recognise and Avoid Them. The EDPB Guidelines 03/2022 only provide recommendations and guidance for the design of the interfaces of social media platforms, thus limiting both its scope and enforceability.

The recent EU Digital Services Act (EU DSA) which recently came into effect last 17 February 2024, adds to the regulatory framework for dark patterns. The EU DSA covers all intermediary services that offer their services to users based in the EU, including online platforms such as app stores, collaborative economy platforms, and social media platforms. Article 25 of the EU DSA stipulates that providers must not design online platforms in a deceitful manner that would impair recipients' ability to make free and informed decisions such as promoting certain user choices over others; repeated requests to the recipient to make a choice which has already been made; and termination of the service being made more difficult than initial subscription or sign-up.

EU regulators have already been acting even prior to the implementations of the EU DSA and the EDPB Guidelines on Dark Patterns. In C-252/21 Meta Platforms Ireland and Others v. Bundeskartellamt, the Court of Justice of the EU found that to sign up for Facebook, users have to consent to a user-agreement that allows the Facebook to track their movements across the web by using off-Facebook data or online data outside Facebook.[2] The CJEU also found that there is no mechanism to separately consent for the processing of data within the Facebook social network and for the off-Facebook data. Due to the absence of these separate consent mechanisms, the consent of users to the processing of off-Facebook data must be presumed not to be freely given. This is a form of bundled consent, and while the CJEU did not specifically mention dark patterns or deceptive design patterns in its decision, regulators in the UK have labelled it as a harmful online choice architecture practice.[3]

[2] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A62021CJ0252

[3] https://www.drcf.org.uk/__data/assets/pdf_file/0024/266226/Harmful-Design-in-Digital-Markets-ICO-CMA-joint-position-paper.pdf

Similarly, France's *Commission Nationale de l'Informatique et des Libertés* (CNIL) imposed a fine of €60,000,000 on Facebook[4] and €150,000,000 on Google[5] for making it more difficult for internet users to refuse cookies than to accept them. Tiktok was also fined by the CNIL for implementing advertising identifiers sans consent and for insufficiently informative cookie banners.[6] CNIL penalized Voodoo for still processing personal information despite user refusal, information linked to the browsing habits for advertising purposes in contradiction with what it indicates in the information screen it displays.[7] Ireland's Data Protection Commission also found that TikTok used dark patterns such as Preselection, Visual Interference, and Forced Action, resulting in a reprimand and fine of €345,000,000.[8]

While, most extensively studied in the United States and the Europe Union, regulators are noticing dark patterns worldwide. India has recently enacted guidelines prohibiting dark patterns. China has enacted the Personal Information Protection Law of the People's Republic of China which prohibits big data swindling among other dark patterns. The Australian Competition and Consumer Commission issued the Digital Platforms Inquiry which tackled dark patterns among other issues.[9] Even the Philippine National Privacy Commission has issued guidelines on deceptive design patterns.

Despite the more comprehensive scope of EU law as compared with the United States, complaints by the FTC best illustrate how dark patterns increase disclosure risks in a relatively straight forward manner. An application or website uses dark patterns such as deceptive omissions or deceptive presentations to manipulate its users into sharing a large amount of personal information, this personal information is then sold or shared to data brokers. As the means used to gather this information was a dark pattern, the

[4] https://www.cnil.fr/sites/cnil/files/atoms/files/deliberation_of_the_restricted_committee_no._san-2021-024_of_31_december_2021_concerning_facebook_ireland_limited.pdf
[5] https://www.cnil.fr/en/closure-injunction-issued-against-google
[6] https://www.cnil.fr/en/cookies-cnil-fines-tiktok-5-million-euros
[7] https://www.cnil.fr/en/mobile-games-closure-injunction-issued-against-voodoo
[8] https://www.edpb.europa.eu/system/files/2023-09/final_decision_tiktok_in-21-9-1_-_redacted_8_september_2023.pdf
[9] https://www.accc.gov.au/system/files/Digital platform services inquiry.pdf

manipulated individual often has little or no idea that they were a victim of this manipulation, and their data is being spread far and wide. Notable examples are the FTC complaints against GoldenShores Technologies, LLC,[10] PaymentsMD,[11] and X-Mode Social.[12] In the EU, for example, the CNIL has also fined other companies for failing to properly secure consent or other legitimate basis for processing personal information, such as Foriou,[13] and Hubside.Store.[14]

3.3 Sale of Personal Data and Data Brokers

The sale and exchange of data is global in scope, yet laws and regulations on the sale of personal data and data brokers vary widely if there are any specific laws at all in each jurisdiction. In the EU, the GDPR is very limited on provisions regarding the selling and trade of personal data. Generally, consent and explicit consent may be used as the legal basis for said sales. However, Ruschemeier (2023) argues that most data brokers are not compliant with the GDPR, due to the fundamental problems of obtaining informed consent in digital environments and the need to balance this against the legitimate interest as a basis for the lawful processing of personal data. In the United States, there is no federal regulation that comprehensively regulates the sale of personal data. Muralidhar & Palk (2020) previously highlighted the unusual discrepancy between strong privacy obligations for the government and nearly nonexistent privacy obligations for the data brokers in the United States. Nonetheless, regulations have come into force as California, Nevada, Virginia, Colorado, Connecticut, and Utah have laws that specifically regulate the sale of personal data mostly giving consumers the right to opt out of said sale and even sharing in the case of California.

To illustrate how this regulation of personal data sales works, section 1798.120 of the California Privacy Rights Act (CPRA) which amended the CCPA, expanded California users' right to opt-out expands to the

[10] https://www.ftc.gov/sites/default/files/documents/cases/131205goldenshorescmpt.pdf
[11] https://www.ftc.gov/system/files/documents/cases/141201paymentsmdcmpt.pdf
[12] https://www.ftc.gov/system/files/ftc_gov/pdf/X-Mode-D%26O.pdf
[13] https://www.cnil.fr/en/commercial-prospecting-foriou-fined-eu310000
[14] https://www.cnil.fr/en/commercial-prospecting-hubsidestore-fined-eu525000

"sharing" as well as the "selling" of their personal information. Section 1798.135 of the CPRA enumerates methods for limiting the sale, sharing, and use of personal information and use of sensitive personal information. In the EU, data brokers are not regulated differently from other personal information controllers. In the US, states such as Vermont, Nevada, Texas, Oregon and California have specific regulations on data brokers. This may change soon as Bureau of Consumer Financial Protection requested the public for Information Regarding Data Brokers and Other Business Practices Involving the Collection and Sale of Consumer Information last March 2023.

Despite these rules, Sherman et. al. (2021) found that U.S. data brokerage industry gathers data on virtually every American, by scraping public records, embedding code into mobile apps, and purchasing customers data from various companies. Sometimes, companies purchase data directly from each other as shown in the recent Federal Trade Commission (FTC) filings against various data brokers such as Kochava.[15]

Even companies that claim not to sell personal data are reported to engage in the bartering of said data for a valuable consideration such as increased engagement on a platform.[16] This distinction between selling and bartering would probably not assuage the concerns of many about the spread and use of their data by companies with whom there is no contract or even a contact. Such is the scope of the data gathered by these brokers that Muralidhar & Palk (2020) claim that it is harder it is to request such data from the government than it is to purchase detailed data about a population from data brokers. The collection of excessive amounts of personal information from users for so-called secondary uses such as selling the data as a commodity, has been well documented by researchers (Spiekermann et. al. 2015; Muralidhar & Palk 2020; Schauer & Schnurr 2023; and Ichihashi 2020).

[15] https://www.ftc.gov/legal-library/browse/cases-proceedings/ftc-v-kochava-inc; The amended complaint against Kochava illustrates what can be described as dark patterns however the FTC uses the broader term "unfair or deceptive acts or practices in or affecting commerce".

[16] See https://apiacademy.co/2018/06/how-the-facebook-api-led-to-the-cambridge-analytica-fiasco/; See also https://www.dli.tech.cornell.edu/post/facebook-and-google-are-the-new-data-brokers; See also

Data brokers such as DB to Data, a Philippine based data broker openly sell call and email lists online.[17] DB to Data offers Philippine call lists with prices ranging from USD 150.00 for a list of 10,000.00 to USD 4,000.00 for a list of 3,000,000.00. The lists include the names, cellphone numbers, address, relationship status, age and work. Lists are also openly available for other countries including the Germany (USD 6,000.00 for a list of 6,000.000.00) and United States (USD 50,000.00 for a list of 90,000,000.00). Special categories such as Binance (USD 4,000.00 for a list of 3,000,000.00) and Chinese Overseas British (USD 6,000.00 for a list of 100,000.00) are also available. There does not seem to be any restrictions on the purchase of the data. There is also no ready explanation as to the source of this data except that it has been verified by both a human and a computer. DB to Data is by no means unique. Buymailmarketinglists.com,[18]. AllEmailList,[19] DBShop,[20]and probably other companies offer comparable products with a similar lack of restrictions on purchase.

The prices and variety of the microdata available from these presumably legitimate companies is dwarfed by the microdata that is available for sale on social media networks. Facebook Groups such as "USA, UK, AUS, CAN, Email Data Buy Sell for Tech Support"[21], "High Ticket Closers and Setters (JOB OPPORTUNITIES),[22] "E-mail Blasting Data-Buyers and Sellers"[23] are among the numerous venues where the sale of microdata can take place. One online seller in these groups offered email lists with 450,000,000 entries on individuals based in the United States for the price of USD 1,500.00 and email lists of 10,000,000

[17] Per the Philippine National Privacy Commission, there is no record of DB to Data being registered as a personal information controller which is required by Philippine law for entities processing the sensitive personal information of more than 1,000 individuals.
[18] This is a Sri Lanka based broker, it offers a list of over 500,000 individuals in the Philippines for USD 89. See https://www.buyemailmarketinglists.com/
[19] The Philippine database for sale by AllEmailList is alleged to be update monthly and contains no usage limitations. See https://www.allemaillist.com/philippines-cell-mobile-phone-number-list-database.html
[20] Includes Philippine data per province. No prices indicated. See https://dbshop.in/database/states/philippines/173
[21] Over 2,100 members, see https://www.facebook.com/share/hpNRfAz5P7AxPjQ9/?mibextid=A7sQZp
[22] Over 42,000 members, see https://www.facebook.com/share/FePtfvt2kend4dzW/?mibextid=A7sQZp
[23] Over 4,700 members, see https://www.facebook.com/share/qen3HbUKdzGRiPCa/?mibextid=A7sQZp

Filipinos for USD 400.00.[24] The number of sellers in these groups dwarfs the 540 data brokers that are registered in both or either California and Vermont, as of 2021, as required by their respective state laws.[25] Data brokers appearing in Google searches and Facebook are only 2 of the very many possible channels for which individual data is bought and sold. These are just the small-fry, in 2020, Axciom, one of the biggest data brokers, declared that is has "*the largest depth and breadth of demographic segmentation and predictive data, identifying over 260 million US consumer portraits*.".[26] Axciom claims that it has the *"largest catalog of over 12,000 global data attributes specifically focused on providing personalized experiences.*"[27] Axciom's data is sourced among others from "*other commercial entities where consumers have been provided notice of how their data will be used, and offered a choice about whether or not to allow those uses*."[28]

Simply put, for the price of a few dollars, one can possibly re-identify a person out of a government released data set. For the price of a few hundred dollars, one can possibly link all the information in said dataset with additional information, producing a very in-depth profile of all individuals.

Those re-identified in publicly available de-identified data sets have also suffered actual harm. In 2021, a newsletter covering the Catholic Church used 24 months' worth of commercially available records of app signal data covering portions of 2018, 2019, and 2020, which included records of Grindr usage and locations where the app was used, to identify and reveal the sexual orientation of a high ranking member of the clergy, who later resigned.[29] Thus, the concerns of preserving privacy and anonymity in the big data context raised by Heffetz and Ligett (2013), a decade ago, still remain given that there is a risk of more harm due to both the often-sensitive content and the vastly larger numbers of people affected.

---

[24] Person will not be named due to possible privacy implications. However, from the profile, it is unclear if this is an individual or a group working together as a data brokerage.

[25] https://privacyrights.org/resources/registered-data-brokers-united-states-2021

[26] https://business.linkedin.com/content/dam/me/business/en-us/amp/marketing-solutions/images/lms-partner/partner-dedicated-acxiom/pdf/top-data-and-models-for-digital-by-industry-vertical-flyer.pdf

[27] https://www.acxiom.com/customer-data/

[28] https://www.acxiom.com/wp-content/uploads/2013/09/Acxiom-Marketing-Products.pdf

[29] The Pillar, "Pillar Investigates: USCCB gen sec Burrill resigns after sexual misconduct allegations" 21 July 2021, available at https://www.pillarcatholic.com/p/pillar-investigates-usccb-gen-sec

4. Regulatory Gaps and Possible Solutions

Ruohonen & Mickelsson (2023) have voiced the concern that state-of-the-art privacy-preserving methods mentioned in the EU DGA cannot prevent de-anonymization and re-identification by efficient algorithms for de-anonymization and re-identification of data subjects. Moreover, the efficiency of these algorithms seems to be increasing with advances in machine learning and artificial intelligence. Oberski & Kreuter (2020) state that the regulatory goal should be ensuring the widest number of data custodians reduce such risks to the lowest possible degree, while still retaining the utility of sharing data. Oberski & Kreuter (2020) have proposed that de-identification policy should be like cybersecurity policy, since perfect, impregnable security is impossible, it follows that de-identification and data security policy should minimize the risk of breaches and other failures down to acceptably low levels. They have pointed out that sharing even a de-identified dataset involves a non-zero chance of re-identifying or learning something new about people within a de-identified dataset. Closing the regulatory gaps regarding dark patterns would contribute to this regulatory goal. Disclosure risks could be significantly reduced if individuals knew where their personal data was going, who was using it, and what purposes it was being used for. Equally important would be the ability to easily control one's personal data without being the subject of manipulation via dark patterns. This is easier said than done as research conducted by Di Geronimo et. al. (2020) shows that 55% of users did not spot malicious designs in applications containing dark patterns, 20% were unsure, and the remaining 25% found a malicious design.

Dark patterns sharpen the criticism of the current notice and consent based framework for the protection of privacy (Cate & Mayer-Schönberger 2013, Hull 2014, Nehf 2019, and Nouwens et. al, 2020). It has been close to a decade since the NITDRP (2016) stated that it is unfair to place the burden individuals with understanding all the legal, privacy, or ethical implications the of sharing data that being used in new and unanticipated ways in the big data era. Better and systematic regulation of dark patterns would help in

reducing disclosure risk for the simple reason that individuals would be able to exercise more control on the spread of their personal information.

Many surveys show people have a strong distaste for the sale of their data and given the openness of the sale of these data sets, it would be fair to conclude that a sizable portion was sourced using dark patterns.[30] Absent more detailed studies on the prevalence of data sets assembled by privacy infringing dark patterns, commercially data sets that can be purchased with little or no limitations, are a reasonable proxy for illustrating disclosure risks in relation to the prevalence of dark patterns.

Regulatory action has been sparse at best, even cursory searches show the vast amount of personal information for sale, a lot of appears to have been gathered by dark patterns. There seems to be a whack-a-mole application of regulations since for every action taken against companies and brokers for using dark patterns such as PaymentsMD, there are numerous other entities such as Life360,[31] Anomaly Six,[32] and Safeguard,[33] whose similar actions have been reported in media but do not seem to have merited investigation let alone enforcement from regulators.

---

[30] See a 2018 survey conducted on behalf of the Australian Competition and Consumer Commission showed that 86% of respondents consider sharing information with unknown third parties to be a misuse of personal information (https://www.accc.gov.au/system/files/ACCC%20consumer%20survey%20Consumer%20views%20and%20behaviours%20on%20digital%20platforms%2C%20Roy%20Morgan%20Research.pdf); See also A 2019 survey conducted on behalf of the United Kingdom Information Commissioner's Office (UK ICO) which showed that 67% of respondents found it unacceptable for their details to be sold to or shared with other organizations (https://ico.org.uk/media/for-organisations/documents/2618466/ico-data-brokers-research-presentation-2903191.pdf); A 2023 survey by Pew Research showed that 67% respondents say they understand little to nothing about what companies are doing with their personal data. 81% of these respondents claim that they are very or somewhat concerned about how companies are using the data they collect about them. 77% of these respondents also claimed to have little or no trust that companies will admit mistakes and take responsibility when they misuse or compromise data. 76% of these respondents also claimed to have little or no trust that social media companies will not sell users' personal data to others without their consent. 84% were very worried or somewhat worried that companies selling your information to others without you knowing.(https://www.pewresearch.org/internet/2023/10/18/how-americans-view-data-privacy/)

[31] https://theintercept.com/2022/04/22/anomaly-six-phone-tracking-zignal-surveillance-cia-nsa/

[32] https://themarkup.org/privacy/2021/12/06/the-popular-family-safety-app-life360-is-selling-precise-location-data-on-its-tens-of-millions-of-user

[33] https://www.eff.org/deeplinks/2022/05/safegraphs-disingenuous-claims-about-location-data-mask-dangerous-industry

Companies using dark patterns seem to be incentivized by the market to do so (OECD 2022 & Runge et. al. 2023). Thus, even if there were more comprehensive regulations, push back from companies and data brokers that have thrived in the current big data landscape is expected. It is one thing for a research team to declare the existence of a dark pattern. It is an altogether different animal for a regulator to determine its existence and implement corrective measures. Understanding that new laws or regulatory frameworks are needed is different from marshaling the power to enact them, let alone the will to adequately enforce them. This is best demonstrated by the light enforcement found by Mahoney (2020) of the Do Not Sell provisions of the CCPA by registered data brokers in California. More troubling was that this study showed some brokers seemingly resorted to dark patterns to get consumers not to exercise their right not to have their data sold (Mahoney 2020).

Given the non-rival nature of personal information, even the current regulations do not seem to be a viable solution to all the personal data that was already gathered due to dark patterns. Assuming that it can be proven that the personal data was sourced due to a dark pattern, it seems untenable for regulators to going after every buyer of this personal data. Many of these buyers relied on the current lack of regulation and would probably be considered to have acted in good faith in arms-length commercial transactions. The dark pattern sourced data in these datasets were allowed to build up over years with scant regulation. It would be wishful thinking that a few laws and occasional regulatory action would instantly remove them from circulation. Nonetheless, data owners need to know where and how much dark pattern sourced data there is, as there is a risk that it could dirty their otherwise clean de-identified data sets. Knowing how much dark pattern sourced data is out there is especially important for research where current protection methods such as differential privacy are not appropriate such as demographers, redistricting analysts or immigration (Ruggles et. al. 2019, and Groshen &Goroff 2022).

From the perspective of individuals, Hotz et. al. (2022) posit that the individual will care a great deal about small increases in the disclosure risk if the probability of disclosure is already high but may not be bothered

by even a large relative increase in risk from data release if the probability of disclosure remains low in absolute terms after release. Laws such as the EU DSA mandate detailed transparency and privacy reporting of VLOPs and VLOSEs with third party audits. The California Delete Act also mandates more transparency from data brokers regarding the sale of personal information. These reports in addition to regulatory sweeps should give data owners and individuals a better idea of where their data is going and appropriate steps to take if these data flows are contrary to their best interests. Hopefully this would keep the probability of disclosure low in absolute terms. Ultimately these mandated reports can show policy makers the best possible use of limited regulatory resources, and individuals if they are at a substantial risk of being re-identified.

Going beyond closing regulatory gaps relative to dark patterns, it might be best to frame the balancing of the utility of sharing data with researchers and the privacy rights of individuals within the context of the Belmont Principle of beneficence.[34] Beneficence is an obligation: (1) to do no harm and (2) maximize possible benefits and minimize possible harms. Integrating the principle of beneficence to a potential remodeling de-identification law and policy to minimize the risk of breaches and other failures down to acceptably low levels, should give individuals redress in case of privacy breaches due to re-identification in released datasets. It can be argued it is much harder to achieve beneficence due to individuals lack of awareness of privacy risks due to unregulated or lightly regulated dark patterns in the context of big data. In this regard, while much of the research has focused on minimization of disclosure risks, government agencies and research institutions can explore minimization or rectification of the harms in cases in cases that result in unwarranted disclosure of personal information. This could be similar to how subjects in medical research are referred to treatment centers after the conclusion of research participation can help avoid unintended harm.

[34] https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html

Questions on how best to deal with dark patterns and their accompanying aggravation of disclosure risks will vary from country to country. Nonetheless, researchers and policy makers have a clear duty to continue to explore and publicly discuss how to approach these questions in the spirit of transparency and the furtherance of beneficence. Scientific and economic research must not be unduly hindered even if we have yet to find the perfect answers to these questions. We may never find perfect answers rather we can hopefully adopt workable solutions that are most appropriate for their age and context, but most importantly, . Technological progress and societal norms are by their very nature dynamic, and thus discussions on balancing these competing needs must be done on a regular basis.

5. Conclusion

This exploratory study highlights the pervasiveness of dark patterns and its effect on disclosure risks. Dark patterns are by no means the only factor in the ongoing debate between protecting privacy and utilizing datasets in the age of big data, but they should still be considered in view of their prevalence and scant regulation. Understanding the gaps in the regulations of these dark patterns should be a key subject for researchers and policy makers as such knowledge will enable society at large better utilize the enormous and varied data sets that fuel big data.

Disclosure Statement

This research was supported by Bangko Sentral ng Pilipinas as part of the author's employment. The views expressed herein are those of the author and do not necessarily reflect the views of the Bangko Sentral ng Pilipinas.

References

Abowd, J. M., & Schmutte, I. M. (2019). An Economic Analysis Of Privacy Protection And Statistical Accuracy As Social Choices. the American Economic Review, 109(1), 171–202. https://doi.org/10.1257/aer.20170627

Abowd, J.M., Adams, T., Ashmead, R., Darais, D., Dey, S., Garfinkel, S.L., Goldschlag, N., Kifer, D., Leclerc, P., Lew, E., Moore, S., Rodr'iguez, R.A., Tadros, R.N., & Vilhuber, L. (2023). The 2010 Census Confidentiality Protections Failed, Here's How and Why. ArXiv, abs/2312.11283.

Acemoglu, D., Makhdoumi, A., Malekian, A., & Ozdaglar, A. (2022).Too Much Data: Prices and Inefficiencies in Data Markets. American Economic Journal: Microeconomics, 14 (4): 218-56. DOI: 10.1257/mic.20200200

Acquisti, A., Adjerid, I., & Brandimarte, L. (2013). Gone in 15 seconds: The limits of privacy transparency and control. IEEE Security & Privacy, 11(4), 72–74. https://doi.org/10.1109/msp.2013.86.

Acquisti, A., Taylor, C., & Wagman, L. (2016). The Economics of Privacy (March 8, 2016). Journal of Economic Literature, 54(2): 442-492 (2016) http://doi.org/10.1257/jel.54.2.442, Sloan Foundation Economics Research Paper No. 2580411, http://dx.doi.org/10.2139/ssrn.2580411

Athey, S., Catalini C., and Tucker, C. (2017). The digital privacy paradox: Small money, small costs, small talk. Technical report, National Bureau of Economic Research https://www.nber.org/system/files/working_papers/w23488/w23488.pdf

Auxier, B., Rainie, L., Anderson, M., Perrin, A., Kumar, M., and Turner, E. (2019). Americans and Privacy: Concerned, Confused and Feeling Lack of Control over their Personal Information. Pew Research Center. https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/

Barth-Jones, D. C. (2012). The “Re-Identification” of Governor William Weld’s Medical Information: A Critical Re-Examination of Health Data Identification Risks and Privacy Protections, Then and Now. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2076397

Bösch, C., Erb, B., Kargl, F., Kopp, H., & Pfattheicher, S. (2016). Tales from the Dark Side: Privacy Dark Strategies and Privacy Dark Patterns. *Proceedings on Privacy Enhancing Technologies*, *2016*(4), 237–254. https://doi.org/10.1515/popets-2016-0038

Cate, F. H. and Mayer-Schönberger, V. (2013) Notice and Consent in a World of Big Data". *Articles by Maurer Faculty*. 2662. https://www.repository.law.indiana.edu/facpub/2662

Culnane, C., Rubinstein, B. & Teague, V. (2017). Health Data in an Open World. https://doi.org/10.48550/arXiv.1712.05627

Culnane, C., Rubinstein, B.I., & Teague, V. (2019). Stop the Open Data Bus, We Want to Get Off. ArXiv, abs/1908.05004.

Culnane, Chris & Rubinstein, Benjamin & Watts, David. (2020). Not fit for Purpose: A critical analysis of the 'Five Safes'. https://doi.org/10.48550/arXiv.2011.02142

Diebold, F. X. (2012). On the Origin(s) and Development of the Term “Big Data.” PIER Working Paper No. 12-037, http://dx.doi.org/10.2139/ssrn.2152421

Di Geronimo, L., Braz, L., Fregnan, E., Palomba, F., & Bacchelli, A. (2020. UI Dark Patterns and Where to Find Them: A Study on Mobile Applications and User Perception. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376600

Dwork, C., Smith, A., Steinke, T., & Ullman, J. (2017, March 7). Exposed! A Survey of Attacks on Private Data. *Annual Review of Statistics and Its Application*, *4*(1), 61–84. https://doi.org/10.1146/annurev-statistics-060116-054123

European Commission, Directorate-General for Justice and Consumers, Lupiáñez-Villanueva, F., Boluda, A., Bogliacino, F. et al. (2022) *Behavioural study on unfair commercial practices in the digital environment : dark patterns and manipulative personalisation : final report*. Publications Office of the European Union. https://data.europa.eu/doi/10.2838/859030

Farzanehfar, A., Houssiau, F., & De Montjoye, Y. (2021). The risk of re-identification remains high even in country-scale location datasets. *Patterns*, *2*(3), 100204. https://doi.org/10.1016/j.patter.2021.100204

Forbrukerrådet. (2018a). Deceived by Design. https://fil.forbrukerradet.no/wp-content/uploads/2018/06/2018-06-27-deceived-by-design-final.pdf

Forbrukerrådet. (2018b). Every Step you Take. https://storage02.forbrukerradet.no/media/2018/11/27-11-18-every-step-you-take.pdf

Gray, C. M., Santos, C., Bielova, N., & Mildner, T. (2023). An Ontology of Dark Patterns Knowledge: Foundations, Definitions, and a Pathway for Shared Knowledge-Building. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2309.09640

Groshen, E. L., & Goroff, D. (2022). Disclosure Avoidance and the 2020 Census: What Do Researchers Need to Know? *Harvard Data Science Review*, (Special Issue 2). https://doi.org/10.1162/99608f92.aed7f34f

Hasan, M., Popp, J., & Oláh, J. (2020). Current landscape and influence of big data on finance. *Journal of Big Data*, *7*(1). https://doi.org/10.1186/s40537-020-00291-z

Henriksen-Bulmer, J., & Jeary, S. (2016, December). Re-identification attacks—A systematic literature review. *International Journal of Information Management*, *36*(6), 1184–1192. https://doi.org/10.1016/j.ijinfomgt.2016.08.002

Heffetz, O. (2022). What Will It Take to Get to Acceptable Privacy-Accuracy Combinations? . *Harvard Data Science Review*, (Special Issue 2). https://doi.org/10.1162/99608f92.5d9b1a8d

Heffetz, O., & Ligett, K. (2013). Privacy and Data-Based Research. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2324830

Hidaka, S., Kobuki, S. Watanabe, M., and Seaborn, K. (2023). Linguistic Dead-Ends and Alphabet Soup: Finding Dark Patterns in Japanese Apps. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 3, 1–13. https://doi.org/10.1145/3544548.3580942

Hotz, V. J. & Slanchev, V. (2017). "Designing Consent Protocols to Link Sensitive Health and Administrative Records in Social Science Surveys: Phase I" Accessed at https://public.econ.duke.edu/~vjh3/working_papers/ConsentProject.pdf

Hotz, V. J., Bollinger, C. R., Komarova, T., Manski, C. F., Moffitt, R. A., Nekipelov, D., Sojourner, A., & Spencer, B. D. (2022). Balancing data privacy and usability in the federal statistical system. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(31). https://doi.org/10.1073/pnas.2104906119

Hotz, V. J., & Salvo, J. (2022). A Chronicle of the Application of Differential Privacy to the 2020 Census. *Harvard Data Science Review*, (Special Issue 2). https://doi.org/10.1162/99608f92.ff891fe5

Hull, G. (2015). Successful failure: what Foucault can teach us about privacy self-management in a world of Facebook and big data. *Ethics and Information Technology*, *17*(2), 89–101. https://doi.org/10.1007/s10676-015-9363-z

Ichihashi, S. (2020). Non-competing data intermediaries. Staff Working Papers 20-28, Bank of Canada. https://doi.org/10.34989/swp-2020-28

Kallioniemi, P. (2022). Facebook's dark pattern design, public relations and internal work culture. *DOAJ (DOAJ: Directory of Open Access Journals)*. https://doi.org/10.34624/jdmi.v5i12.28378

Keller, S. A., & Abowd, J. M. (2023, March 15). Database reconstruction does compromise confidentiality. *Proceedings of the National Academy of Sciences*, *120*(12). https://doi.org/10.1073/pnas.2300976120

Kelly, D., & Burkell, J. (n.d.). *Documenting Privacy Dark Patterns: How social networking sites influence users' privacy choices*. Scholarship@Western. https://ir.lib.uwo.ca/fimspub/376

Kitkowska, A. (2023). The Hows and Whys of Dark Patterns: Categorizations and Privacy. In: Gerber, N., Stöver, A., Marky, K. (eds) Human Factors in Privacy Research. Springer, Cham. https://doi.org/10.1007/978-3-031-28643-8_9

Kollmer, T. & Eckhardt, A. (2022). Dark Patterns: Conceptualization and Future Research Directions. Business & Information Systems Engineering. 65. 10.1007/s12599-022-00783-7.

Kowalczyk, M., Gunawan, J. T., Choffnes, D., Dubois, D. J., Hartzog, W., & Wilson, C. (2023, April 19). Understanding Dark Patterns in Home IoT Devices. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3544548.3581432

Leiser, D. M. (2020). "Dark Patterns": The Case for Regulatory Pluralism. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3625637

Lohr, S. (2012, February 11) The Age of Big Data, *New York Times*, https://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html

Luguri, J. B., & Strahilevitz, L. (2021). Shining a light on dark patterns. *the Journal of Legal Analysis*, *13*(1), 43–109. https://doi.org/10.1093/jla/laaa006

Mahoney, M., (2020) California Consumer Privacy Act: Are Consumers' Digital Rights Protected? Consumer Reports Digital Lab, http://advocacy.consumerreports.org/wp-content/uploads/2020/09/CR_CCPA-Are-Consumers-Digital-Rights-Protected_092020_vf.pdf

Marti, D., Lin, F. Scharwtz, M. and Fahs, J. (2024). Who shares your information with Facebook: Sampling the Surveillance Economy in 2023. https://innovation.consumerreports.org/wp-content/uploads/2024/01/CR_Who-Shares-Your-Information-With-Facebook.pdf

Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer, J., Chetty, M., & Narayanan, A. (2019). Dark patterns at scale. *Proceedings of the ACM on Human-computer Interaction*, *3*(CSCW), 1–32. https://doi.org/10.1145/3359183

Mathur, A., Kshirsagar, M., & Mayer, J. (2021). What makes a dark pattern. . . dark? *arXiv (Cornell University)*. http://arxiv.org/abs/2101.04843

McGeeney, K., Kriz, B., Mullenax, S., Kail, L., Walejko, G., Vines, M., Bates, N., & García Trejo, Y. (2019). *2020 Census Barriers, Attitudes, and Motivators Study survey report: A new design for the 21st century*. U.S. Census Bureau. https://www2.census.gov/programs-surveys/decennial/2020/program-management/finalanalysis-reports/2020-report-cbams-study-survey.pdf

Mhaidli, A.H., Zou, Y., & Schaub, F. (2019). *"We Can't Live Without Them!" App Developers' Adoption of Ad Networks and Their Considerations of Consumer Risks.* In Proceedings of the Fifteenth USENIX Conference on Usable Privacy and Security (SOUPS'19). USENIX Association, USA, 225–244. https://dl.acm.org/doi/10.5555/3361476.3361493

Muralidhar, K., & Palk, L. (2018). A free ride: Data Brokers' Rent-Seeking Behavior and the Future of data Inequality. *Vanderbilt Journal of Entertainment & Technology Law*, *20*(3), 779. https://scholarship.law.vanderbilt.edu/cgi/viewcontent.cgi?article=1097&context=jetlaw

Narayanan, A., & Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. *Proceedings - IEEE Symposium on Security and Privacy/Proceedings of the . . . IEEE Symposium on Security and Privacy*. https://doi.org/10.1109/sp.2008.33

Narayanan, A., Mathur, A., Chetty, M., & Kshirsagar, M. (2020, April 30). Dark Patterns: Past, Present, and Future. *Queue*, *18*(2), 67–92. https://doi.org/10.1145/3400899.3400901

Nehf, J.P. (2019). The Failure of 'Notice and Consent' as Effective Consumer Policy. LSN: Regulation of Contracting Private Parties (Topic). http://dx.doi.org/10.2139/ssrn.3440816

Nissenbaum, H. (2009). *Privacy in context, technology, policy, and the integrity of social life*. Stanford University Press.

Nouwens, M., Liccardi, I., Veale, M., Karger, D., & Kagal, L. (2020, April 21). Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3313831.3376321

Nycyk, M. (2021). Ethical Use of Informant Internet Data: Scholarly Concerns and Conflicts. *Journal of Digital Social Research*, *4*(1), 1-22. https://doi.org/10.33621/jdsr.v4i1.88

Oberski, D. L., & Kreuter, F. (2020). Differential Privacy and Social Science: An Urgent Puzzle. *Harvard Data Science Review*, *2*(1). https://doi.org/10.1162/99608f92.63a22079

OECD (2022), "Dark commercial patterns", *OECD Digital Economy Papers*, No. 336, OECD Publishing, Paris, https://doi.org/10.1787/44f5e846-en.

The Office of the Victorian Information Commissioner (2022). The Limitations of De-Identification – Protecting Unit-Record Level Personal Information. https://ovic.vic.gov.au/privacy/resources-for-organisations/the-limitations-of-de-identification-protecting-unit-record-level-personal-information/

Ohm, P. (2009). Broken Promises of privacy: Responding to the surprising failure of anonymization. UCLA Law Review, Vol. 57, p. 1701, 2010, U of Colorado Law Legal Studies Research Paper No. 9-12, Available at SSRN: https://ssrn.com/abstract=1450006

Owens, K., Gunawan, J., Choffnes, D., Emami-Naeini, P., Kohno, T., & Roesner, F. (2022, September 29). Exploring Deceptive Design Patterns in Voice Interfaces. *Proceedings of the 2022 European Symposium on Usable Security*. https://doi.org/10.1145/3549015.3554213

Press, G. (2019, July 17). A very short history of big data. Forbes. https://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/?sh=116b2a9865a1

Reilly, B. C. (2015). Doing More with More: The Efficacy of Big Data in the Intelligence Community. *American Intelligence Journal*, *32*(1), 18–24. http://www.jstor.org/stable/26202099

Rocher, L., Hendrickx, J. M., & De Montjoye, Y. (2019). Estimating The Success Of Re-Identifications In Incomplete Datasets Using Generative Models. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-10933-3

Ruohonen, J. & Mickelsson, S. (2023). Reflections on the Data Governance Act. Digital Society. 2. 10.1007/s44206-023-00041-7.

Rubinstein, I. (2013). Big data: the end of privacy or a new beginning? *International Data Privacy Law*, *3*(2), 74–87. https://doi.org/10.1093/idpl/ips036

Ruggles, S., (2018). "Implications of Differential Privacy for Census Bureau Data and Scientific Research." Minnesota Population Center Working Paper 2018-6 https://assets.ipums.org/_files/mpc/wp2018-06.pdf#%5B%7B%22num%22%3A370%2C%22gen%22%3A0%7D%2C%7B%22name%22%3A%22FitH%22%7D%2C792%5D

Ruggles, S., Fitch, C. A., Magnuson, D. L., & Schroeder, J. P. (2019). Differential Privacy and census Data: Implications for social and economic research. *AEA Papers and Proceedings*, *109*, 403–408. https://doi.org/10.1257/pandp.20191107

Runge, J., Wentzel, D., Huh, J.Y. & Chaney, A. *(2023).* "Dark patterns" in online services: a motivating study and agenda for future research. *Mark Lett* 34, 155–160. https://doi.org/10.1007/s11002-022-09629-4

Ruschmeier, H. (2023). Data brokers and European digital legislation. *European Data Protection Law Review*, *9*(1), 27–38. https://doi.org/10.21552/edpl/2023/1/7

Sala, E., Knies, G., & Burton, J. (2014). Propensity to consent to data linkage: experimental evidence on the role of three survey design features in a UK longitudinal panel. *International Journal of Social Research Methodology*, *17*(5), 455–473. https://doi.org/10.1080/13645579.2014.899101

Schmutte, I., and Vilhuber, L. (2020). "Balancing Privacy and Data Usability: An Overview of Disclosure Avoidance Methods." In: Cole, Dhaliwal, Sautmann, and Vilhuber (eds), *Handbook on Using Administrative Data for Research and Evidence-based Policy*. Accessed at https://admindatahandbook.mit.edu/book/v1.0/discavoid.html on 2024-04-05.

Sherman, J., (2021) Data Brokers and Sensitive Data on U.S. Individuals, available at https://techpolicy.sanford.duke.edu/wp-content/uploads/sites/4/2021/08/Data-Brokers-and-Sensitive-Data-on-US-Individuals-Sherman-2021.pdf

Soe, T. H., Nordberg, O. E., Guribye, F., & Slavkovik, M. (2020, October 25). Circumvention by design - dark patterns in cookie consent for online news outlets. *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*. https://doi.org/10.1145/3419249.3420132

Solove, D. J. (2004). *The Digital Person: Technology and Privacy In The Information Age*. New York: New York University Press.; available at https://scholarship.law.gwu.edu/cgi/viewcontent.cgi?article=2501&context=faculty_publications;

Spiekermann, S., Acquisti, A., Böhme, R., & Hui, K. L. (2015, April 29). The challenges of personal data markets and privacy. *Electronic Markets*, *25*(2), 161–167. https://doi.org/10.1007/s12525-015-0191-0

Stoughton, J. W., Whelan, T. J., & Thompson, L. F. (2015). Perceptions of confidentiality in survey research: Development of a scale. ResearchGate. https://www.researchgate.net/publication/269104204_Perceptions_of_confidentiality_in_survey_research_Development_of_a_scale

Susser, D., Roessler, B., & Nissenbaum, H. F. (2018). Online Manipulation: Hidden Influences in a Digital World. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3306006

Susser, D., Roessler, B., & Nissenbaum, H. (2019). Technology, autonomy, and manipulation. *Internet Policy Review*, *8*(2). https://doi.org/10.14763/2019.2.1410

Sweeney, L. (2011). Patient identifiability in pharmaceutical marketing data. *Harvard University, Cambridge, MA, WP--1015*. https://dataprivacylab.org/projects/identifiability/pharma1.pdf

Sweeney, L. *(2015). Only You, Your Doctor, and Many Others May Know. Technology Science*. 2015092903. https://techscience.org/a/2015092903/

Sweeney, L. (2018). Patient Identifiability in Pharmaceutical Marketing Data. Carnegie Mellon University. Journal contribution. https://doi.org/10.1184/R1/6625193.v1

Tahaei, M. and Vaniea, K. (2021). "Developers Are Responsible":
What Ad Networks Tell Developers About Privacy. In Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21). Association for Computing Machinery, New York, NY, USA, Article 253, 1–11. https://doi.org/10.1145/3411763.3451805

Tahaei, M., Ramokapane, K. M., Li, T., Hong, J. I., & Rashid, A. (2022). Charting App Developers' Journey Through Privacy Regulation Features in Ad Networks. *Proceedings on Privacy Enhancing Technologies*, *2022*(3), 33–56. https://doi.org/10.56553/popets-2022-0061

Tan, X., Qin, L., Kim, Y. and Hsu, J. (2012) Impact of Privacy Concern in Social Networking Web Sites. Internet Research, 22, 211-233. https://doi.org/10.1108/10662241211214575

Utz, C., Degeling, M., Fahl, S., Schaub, F., & Holz, T. (2019). (Un)informed Consent. Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, 973–990. https://doi.org/10.1145/3319535.3354212

Utz, C., Amft, S., Degeling, M., Holz, T., Fahl, S., & Schaub, F. (2023). Privacy rarely considered: Exploring considerations in the adoption of Third-Party services by websites. *Proceedings on Privacy Enhancing Technologies*, *2023*(1), 5–28. https://doi.org/10.56553/popets-2023-0002

Wolf LE. Risks and Legal Protections in the World of Big-Data. Asia Pac J Health Law Ethics. 2018 Mar;11(2):1-15. Epub 2018 Mar 31. PMID: 31745539; PMCID: PMC6863510.

Yoo J., Thaler A., Sweeney L., and Zang J. (2018) Risks to Patient Privacy: A Re-identification of Patients in Maine and Vermont Statewide Hospital Data. *Technology Science*. https://techscience.org/a/2018100901/

Zimmer, M. (2010). "But the data is already public": on the ethics of research in Facebook. *Ethics and Information Technology*, *12*(4), 313–325. https://doi.org/10.1007/s10676-010-9227-5