The Realities of Disclosure Risks in the Age of Dark Patterns and Big Data

Ramon Abraham A. Sarmiento

The prevalence of dark patterns in the age of big data has led to unprecedented collection of personal information which has been shown to erode the privacy rights of individuals. These dark patterns can manipulate individuals into the erosion of their privacy rights and in turn the heightening of disclosure risks. Given the scant and uneven application of regulations on dark patterns, out of the box solutions to disclosure risks should be considered in light of the almost unregulated market for personal information.

1. Dark Patterns

Harry Brignull 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….*" 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 people into surrendering their personal information such as a study that concluded that dark patterns 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, 2018).

2. Disclosure Risks

Previous disclosure avoidance 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. Dark patterns heighten disclosure risks as they have made personal information easier to gather and easier to purchase.

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) successfully identifying the known Netflix users in a publicly released dataset. 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.

The Risk of Re-identification or Reidentification Risk (RRI) refers to the potential that supposedly anonymous or pseudonymous datasets could be de-anonymized to recover the identities of users. We estimated the RII using microdata from: A. 2022 Philippine National Demographic and Health Survey (NDHS) by United States Agency for International Development Aid (USAID); B. 2022 Annual Poverty Indicators Survey (APIS) by the Philippine Statistics Authority (APIS); C. Consumer Expectations Survey (CES) by the Bangko Sentral ng Pilipinas. Larger equivalence classes have lower probabilities of re-identification as they have more data subjects to deal with, and smaller equivalence classes have higher probabilities of re-identification. Once the risk for each row is known using the below equation, the RRI of the whole dataset can be calculated by getting the mean of all per-row RRIs.

$$Risk\ of\ Re-identification = \frac{1}{Size\ of\ Equivalence\ Class}$$

Additionally, sensitive rows violating the k-anonymity condition are consolidated to gauge the level of disclosure risk. k-anonymity is a data anonymization technique used for reducing the risk associated with releasing individual-level data by ensuring that each entry in the dataset is indistinguishable from at least k-1 other entries, based on a set of quasi-identifiers.

The RRI and count of sensitive rows of the three sample datasets are computed and listed in Table 1. Masked data were also generated using data masking techniques, such as generalization and suppression, and underwent RRI estimation.

Table 1: Description of the three sample datasets

	No. of columns	No. of rows	No. of PIIs	% of Sensitive Rows (k=3)
NDHS	619	129,724	53	18.79%
APIS	75	179,947	9	25.11%
CES	1,361	7,468	26	5.58%

Table 2: Risk Re-Identification Indices (RRI) of the three sample datasets

	No Masking	With Data Masking
NDHS	0.3459	0.0635
APIS	0.2223	0.1149
CES	0.6613	0.2248

Data masking was applied to the three datasets by binning and generalizing the PIIs and compared the RRI before and after this procedure. For all datasets, it can be observed that the risk is lower for masked datasets, which increases the size of the equivalency classes. The proliferation of publicly available datasets and the emergence of unconventional access methods such as dark patterns, there arises the capability to interconnect multiple datasets by purchasing or securing datasets from numerous brokers.

3. Disclosure Avoidance

Many researchers 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, subject to onerous conditions for release or 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).

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. Similarly, the European Union (EU) 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. These obligations to keep personal information confidential run in parallel to the various laws to protect personal information and privacy rights.

4. Dark Pattern Regulation

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 in states such as California. Prior to specific legislation enacted to combat dark patterns, the FTC has used the FTC Act's prohibition on 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.

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 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 February 17, 2024, adds to the regulatory framework for dark patterns.

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.

Complaints by the FTC against GoldenShores Technologies, LLC., PaymentsMD, and X-Mode Social, among others, 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 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.

5. 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. 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.

6. Regulatory Gaps and Possible Solutions

Ruohonen & Mickelsson (2023) have voiced the concern that state-of-the-art privacy-preserving methods mentioned in the EU Data Governance Act 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.

Disclosure risks could be significantly reduced if one could 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. 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.

Regulatory action has been sparse at best. 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, Anomaly Six, and Safeguard, whose similar actions do not seem to have merited investigation let alone enforcement from regulators.

Companies using dark patterns are 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 California Consumer Privacy Act by registered data brokers.

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. 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. 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. 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.

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. 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.

7. 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.

References

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.

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

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

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

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

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

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

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

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

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.

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

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

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