Understanding How Countries Regulate Data Collection

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Jaden Carroll

Spring 2023

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

Kent A. Wayland, Department of Engineering and Society

Today, 2.5 quintillion bytes of data are generated every day by consumers using the internet (Price). Technology companies like to gather data to better understand their consumers so that they can improve their products and shape their marketing strategies. The knowledge gained from data collection can help businesses gain more customers which leads to more profit. Company data collection is here to stay as the big data industry "shot up from $169 billion in 2018 to $274 billion in 2022 — a 62% increase" (Albertson). With this prominent growth in the industry, concerns about the practice of data collection also increase.

Although data collection has its advantages for the typical web user, there are still many privacy and security concerns. If a consumer's personal data falls into the wrong hands they can face consequences such as financial losses and fraud. The risks of data breaches and identity theft cause consumers to lose trust in the businesses and companies they associate with, which can be detrimental to both consumers and companies. To address this, different countries have developed regulations to manage the way companies collect consumer data in hopes of improving the overall security of the Internet. It is essential to analyze the relationship between consumers and large businesses to fully understand the issue of how to best regulate big data collection. For my STS research, I will examine different cases of how countries regulate data collection and how consumers view the issue of data processing. This will allow me to analyze the results of such regulations and allow for a proper examination of whether these regulations accomplished their intended goal. The research will provide an opportunity to discuss necessary changes or additions to government data collection regulations to improve data security for consumers.

Background:

In order to better understand the issue of regulating data collection, we must first understand the different types of software used by companies to collect consumer data. Companies use web trackers to collect different information about consumers on their websites. In 1994, the first web tracking cookie was invented by Lou Montulli who while working at Netscape, designed the tool to make websites more commercially applicable. Tracking cookies are perhaps the most widely-used web tracker as they store the consumer's user session by storing a small packet of information on the user's browser which monitors their activity on that particular website. These cookies can help to make the website faster and more personal which likely improves the overall experience of the user. Tracking cookies are versatile and very effective for monitoring the different users that utilize the different web platforms of a company. Web beacons are another data collection tool used by companies which is essentially a transparent image that when clicked, signals the company that a user has accessed some particular function on their site. This can help companies track what particular content a user accessed during their time on a website. These are just a few of the different technologies that have helped companies collect data and insights on different consumers.

These technologies can be used to collect and store a variety of data about users such as how engaged they are with particular content or the different ways they interact with the services. However, personal data is the most important type of information that businesses collect online because it can be used to personally identify individuals. Common examples of personal information consist of names, phone numbers, and addresses which alone can easily identify a user. Depending on the services a site is offering, they may collect sensitive information about users whether it is health records, social security numbers, or credit card

information. If sensitive data is leaked to malicious individuals, they could use this information to steal people's identities and/or monetary assets. The data collected by companies online can provide some extremely valuable insights into the particular demographics and products that may interest these individuals. Although most companies like to keep this information to themselves, some may consult with third-party data brokers who specialize in processing user data and selling it to different businesses. When companies "sell" information they are often selling it to these brokers who then process it to gain insights that other companies may be interested in purchasing. This relationship between companies and data brokers has sparked the growth of the data collection industry. An issue arises because of this relationship, as consumers often feel exploited when their personal data is being sold for monetary profit.

To better understand the reason why many users have qualms about companies gathering data about them, Cristl and Spiekermann (2016), analyze the relationship between consumers and companies. They discuss a clear power imbalance between consumers and companies as it relates to data collection on the Internet. Users are required to consent to data tracking if they want access to features many websites offer. For example, a user who does not accept cookies on a shopping website will lose all their items if they refresh. Transparency is another critical part of the issue as consumers usually do not understand what their data is being used for, and companies often are usually not obligated to explain their reasons for data collection. This lack of transparency from corporations causes a rift between these two parties as much of the public has concern over how these companies may be threatening their data security and privacy. Perhaps the main issue discussed in the article are the potential societal implications it may have on consumers relating to employment, credit scores, and other risks. Depending on how companies process the information the algorithm they use may exhibit biases that discriminate

against specific demographics from information such as addresses, or ethnicity. Governments have a role to play in this relationship, as they act as a medium that seeks to even out the power imbalances. Government regulations are created to offer a solution to consumers by giving them the right to understand how and why companies use the data they collect.

Methods:

Understanding the ever-changing relationships between the different involved parties is crucial to examine how and why the regulations of different countries have formed. To accomplish this, I have analyzed different articles demonstrating how India and the European Union have regulated data collection in their respective countries. The EU developed a revolutionary data protection regulation in the GDPR, and India was chosen as it adopts many of the ideas of the GDPR. To better understand the different actants in place for these situations, I have read articles that discuss the opinions of how the different actants such as consumers perceive the issue of data collection. I have also compiled different official government regulations such as the EU's GDPR to analyze the different policies these countries have established. Analyses of data protection regulations in the EU and India have been considered in order to understand how these regulations address the data protection issue. By utilizing these different types of sources which describe different unique aspects of the issue, I would be able to better comprehend the different aspects of the issue. These sources among others will not only be used to gain a better understanding of the issues surrounding company data collection but to hopefully offer a solution that will better serve all parties involved.

Results:

In order to provide some context about the consumers in both the EU and India, I have gathered surveys demonstrating how individuals in both markets view the data collection issue.

In 2019, a study was conducted by the European Union Agency for Fundamental Rights (FRA) to document people's experiences and opinions regarding their fundamental rights. The Fundamental Rights Survey provides excellent information about how EU citizens feel about data collection and regulations such as the GDPR. The first question asked in the document was for the subjects to express their levels of concern with their personal data being accessed by particular groups. The groups included: criminals, advertisers/businesses, their governments, foreign governments, and more. Not surprisingly consumers found that they were most concerned with criminals accessing their personal information as 55% felt concerned. Advertisers, businesses, and foreign governments were distant runner-ups as around 30% of consumers were concerned about such parties. It is important to note that consumers were the least concerned about their respective governments and law enforcement agencies accessing their personal information. Respondents with smartphones were then asked questions documenting their knowledge of privacy and location settings on their phones. It was found that 72% of people know how to access the privacy settings on their smartphones. However, 59% of users are aware of privacy settings for only some/none of the apps on their phones and 77% of people know where to turn off their location settings ("European Union," 2020). The respondents' knowledge of existing privacy regulations such as the GDPR was also measured in the study. 69% of respondents had previously heard about the GDPR which regulates data protection in the EU. Also, only 51% of respondents knew they had the right to access the personal information that businesses collect from them. The Fundamental Rights Survey presents an accurate representation of how citizens in the EU view data protection in their countries and their level of knowledge of existing laws and acts they can take to protect their personal data.

Surveys documenting Indian citizens' opinions concerning data collection have also been gathered to help compare between India and the EU. The Hindu Business Line published some survey findings in 2021 documenting how Indian consumers felt about certain data collection issues in their country. The questions asked in the survey were similar to the ones asked in the Fundamental Rights Survey in the EU. The research was done on 6,000 Indian respondents to record opinions and possible distrust in how organizations in India deal with consumer personal data. The report found that Indian consumers had an extreme preference(84%) towards companies that were committed to protecting consumer data privacy. In the EU, these rates are much lower as India surpasses the UK(49%), Germany(41%), Spain(36%), and France (17%). Although many of the respondents had some sort of trust in the different companies they interact with to store their personal information safely, 24% of the respondents do not trust third-party companies to keep their information confidential. Roughly 67% of the respondents knew how to keep their own personal data private data on their smartphones when managing different applications and services(Sheth, 2021). Those in India also seem to have a better understanding of the laws protecting their data where only 30% said they had little to no understanding of such laws which is less than in the UK(36%), Germany(32%), and Spain(40%). The survey findings provide a basic understanding of how EU and Indian consumers perceive data collection issues and their overall knowledge of how their data protection laws work.

Now that we have some insight about how the consumers themselves feel about data privacy we will discuss the data protection laws that these governments have established. First, we go over the General Data Protection Regulation(GDPR) adopted by the EU in 2016 which protects citizens in the EU from the improper processing of personal data. This regulation has since become the standard used by many countries outside of the EU to help protect consumers'

rights to data privacy. The GDPR attempts to empower consumers by requiring organizations to uphold three conditions: transparency, consumers' right to be informed, and informed consent. The regulation also provides standards to ensure that companies are keeping personal data private so that data breaches and outside attacks do not negatively affect consumers on their platforms. The regulation applies to any organizations that process any personal data of EU citizens which means that large companies like Google and Facebook which often process data from the EU must also abide by the GDPR guidelines. There are several important legal terms defined by the GDPR in order to make the laws clear and difficult to misinterpret. According to the GDPR, personal data "is any information that relates to an individual who can be directly or indirectly identified" ("What is the GDPR, " 2018). Personal data is often categorized into different categories that may be subjected to different compliance requirements depending on how sensitive the metric is. Three important parties involved in data collection processes are the data subjects, data controllers, and data processors. Data subjects are typically consumers on an organization's website whose information is being processed by either data controllers or processors. Data controllers are the individuals responsible for deciding how and why personal data is processed in a particular organization; anyone who handles data in an organization is a data controller. Data processors are defined as third parties that process personal data on behalf of a data controller, this includes website hosts, email service providers, or even big data analytics companies.

The GDPR requires data controllers to demonstrate that they are GDPR compliant which involves documenting the different types of data stored and how they use and store this information. Data controllers are required to train their employees and designate data protection responsibilities to ensure that the data is secure from unauthorized access. Data protection

agreements are also needed between any third-party data processors that deal with any personal information. GDPR compliance helps to promote accountability for companies dealing with any form of personal data in the EU. The GDPR also has certain guidelines for when it is acceptable to collect personal data and rules for informed unambiguous consent from data subjects to process their information. Furthermore, there are rules ordering organizations to consider consumer data protection during the design and implementation of their applications in order to prevent insecure design flaws. The GDPR enforces its rules by issuing serious fines and punishments for not following its protocols. Since being established, the GDPR has forced many companies to be more proactive in improving the data privacy and security of their applications with the risk of these incurring fines (Jones, 2022).

Many countries have adopted many ideas from the GDPR as it has done a good job regulating data controllers and giving consumers on the internet more rights to access the data being processed. India's current data protection regime consists of a collection of dated laws that do not hold up in the growing technological economy. To address the issue, India's Ministry of electronics and information technology issued the 2022 Digital Personal Data Protection Bill (PDPB) for public consultation("India," 2023). The bill has not yet been passed, but it was designed with the intent to help regulate company data processing similar to foreign laws like the GDPR. The table below(Figure 1) depicts key differences and similarities between the European Union's GDPR and India's PDPB:

Similarities:	**Differences:**
Purpose: Both regulations share the aim at protecting their consumer's privacy and security online. They establish different laws to promote transparency, confidentiality, and accessibility for the data that is collected.	**Scope**: The GDPR has much more of a global impact as any company in any part of the world that processes EU citizen data must comply with its regulations. Whereas, the PDPB only affects Indian businesses.
Data Subject Rights: Both regulations give their citizens more access to their personal data online. Such as the rights to access, modify, and remove any personal information that a company collects.	**Penalties:** The GDPR imposes more significant fines, with penalties of up to 4% of a companies annual income. The PDPB also fines for non-compliance, but the penalty amount is lower.
Informed Consent: Both regulations require companies to acquire informed consent from their users if they are to process their personal information	**Data Protection Officers:** Both regulations may require companies to appoint officers in charge of helping them comply with various laws. Under the GDPR, companies are only forced to appoint one if they deal with large amounts of sensitive information.
Data Breaches: Both regulations require companies to notify consumers and authorities in the event of a data breach.	**Privacy Notices**: Privacy notices under the PDPB only need to describe the types of personal information being collected.

Figure 1. Data Protection Regulation Comparison Table

India’s PDPB has taken inspiration from the GDPR as it aims to protect its citizens' right to privacy by imposing certain rules businesses must follow. Although both the GDPR and the PDPB share a primary focus on increasing data protection, the PDPB is laxer in its regulations as companies are not forced to reveal as much information to the consumers when compared to the GDPR. Another difference is that the PDPB refers to data controllers as data fiduciaries and data subjects as data principles; which leans into the idea that organizations managing personal data are maintaining the trust of their customers. Unlike the GDPR, the PDPB classifies different types of data fiduciaries depending on the scale of the personal data they process(Mahawar, 2022). These ‘significant data fiduciaries’ are forced to oblige towards increased compliance requirements such as appointing different types of data protection auditors and officers to help ensure their data is secure. The PDPB has a higher threshold for reporting data breaches than the GDPR as the breach must be “likely” to cause harm to the data subjects. The penalties for non-compliance under the PDPB are tough and companies could pay 31 million dollars or even more if the contravention was extreme. Although the PDPB and GDPR take different approaches to protect the privacy of their citizens, they both have established sufficient laws that force companies to be more considerate in their data processing.

Discussion:

One of the most important aspects of any complicated social issue is to analyze the individuals who face the most risk. Data subjects or consumers are often at risk of getting critical personal information exposed to those that are not authorized to view this information. So when looking at the different consumer surveys from the EU and India it is not shocking to see that there are existing fears of personal data getting into the hands of criminals or third-party organizations. Users on the internet often have little to no say on what companies do with their

personal information which illustrates the need for proper data protection regulations to reduce the power large corporations have over consumer data. The surveys done in India and in the EU also demonstrated a need for educating the public on strategies to better manage their data online. In the EU, there is a significant portion of the population who do not know how to properly access privacy or location-sharing settings on their devices. India was only marginally better in this regard, so educating consumers in both the EU and India can help to limit the risks they face. Citizens in the EU and India also suffer from lacking much knowledge about their own data protection regime, as there are significant portions of the population that have no knowledge of the different privacy rights they hold online. Providing more education about the specific rights users have in the EU and India is essential to increasing the power these citizens have over their data.

When it comes to the regulations themselves the GDPR was a revolutionary data law that not only changed the way companies in the EU process data but also the whole world. The GDPR gave more power to the consumers, and the government could financially punish companies for insecure and unauthorized usage of consumer data. With the threat of larger fines, companies are forced to take measures in protecting sensitive data and being more transparent to the public. Transparency is essential to managing the data subject-controller relationship as consumers often do not believe companies hold their best interest in mind. The GDPR forces websites to provide different notices to their users to inform them of the types of data they may be collecting and the purpose for which they are collecting this information. A major fear consumers have is that of the unknown and what companies may be doing with the personal data they collect. Companies practicing transparency can help mitigate these concerns as users will have more knowledge and assurance that their data is being managed securely. The PDPB took a

more trust-focused approach in their regulation as instead of terms like data subject and controller they use data principle and fiduciary. These terms describe companies as trustees who are given permission and trust from consumers online to use their personal information in a way that would benefit them. The benefit could be easier ordering, product recommendations, etc. The PDPB does require organizations to notify users of the types of personal data they may be collecting but they are not obligated to tell users how they plan to use this information. The PDPB likely does not require this because users may not understand the purpose and it could cause potential customers to deny a service.

Since being passed into law, the GDPR has forced businesses not only in the EU but around the world to put forth more effort into protecting their consumers' privacy rights. As in a study documenting how the GDPR has affected businesses, roughly 78% of U.S. companies in the study had conducted a GDPR gap assessment to assess their different privacy policies. 27% of companies in the same study contributed over half a million dollars to become GDPR compliant (Bonderud, 2022). This is due to the fact that any company that processes any sort of EU citizen personal information must comply with the GDPR no matter if they are an EU company or not. India's PDPB has a much smaller scope of enforcement as they have not passed their bill into law yet as it is still being evaluated by the public for any modifications. It is also interesting to see that the PDPB is notably less strict in its non-compliance penalties when compared to the GDPR. Statistics have actually shown that European businesses exposed to the GDPR saw their profits shrink by an average of 8.1 percent(Mueller, 2022). Indian legislators may have considered this fact when drafting the PDPB. The loss of profits had a large effect on the smaller to mid-sized companies in the EU. By being a little more lax in their enforcement, India may be protecting their smaller companies that could suffer from such harsh fines.

Ultimately, the PDPB has taken a lot of inspiration from the GDPR in trying to enforce data protection. However, the designers felt that being too controlling of companies could impact the success of businesses in India. In all, both of these approaches would help to close the gap in power between large organizations and consumers by limiting the amount of control companies have over user data.

Conclusion:

The issue of data collection has been a large controversial issue over the years with the growth of technology. Finding a way to properly manage the complex process of data collection is a difficult task that requires reducing the amount of power organizations have over consumers. Different regulations such as the GDPR and PDPB have attempted to solve the issue by instituting large penalties for improper data security measures and adding more information on the types of data being collected. The PDPB may be less transparent in the types of information companies are forced to communicate to their users but both regulations help to protect their resident's data privacy. In order to properly manage data processing in the future, governments must work to educate consumers on the rights they have and the necessary actions they should take to better protect their information. Improving the trust between businesses and consumers is necessary for data collection, as users need to be comfortable with lending their trust and information to organizations. The GDPR and PDPB try their best to build this trust with their rules designed to increase transparency. Overall these two regulations are a good step toward proper and secure data processing. Later on, when the GDPR has been in action for longer and India has finalized the PDPB these regulations can be better analyzed to discover what parts may need to be amended in order to better handle data processing. The data collection industry is rapidly developing and this evolving issue is not going to be resolved with one universal

solution. Organizations, consumers, and governments need to work together to find solutions that would benefit all parties without leaving one in a dominant position over the other.

References:

Albertson, M. (2018, March 9). Software, not hardware, will catapult big data into a $103B business by 2027. SiliconANGLE. https://siliconangle.com/2018/03/09/big-data-market-hit-103b-2027-services-key-say-analysts-bigdatasv/

Bonderud, D. (n.d.). *General Data Protection Regulation (GDPR) Statistics*. Retrieved April 28, 2023, from https://withpersona.com/blog/top-gdpr-statistics-businesses-must-know

Cristl, W. Spiekermann, S. (2016). Networks of Control – A Report on Corporate Surveillance, Digital Tracking (pp. 118–130). Vienna, Austria: Facultas.

Department of International Law: Data Protection. (2021). Retrieved March 30, 2023, from http://www.oas.org/dil/data_protection.htm

European Union Agency for Fundamental Rights. (2020). *Your rights matter: Data protection and privacy : fundamental rights survey.* Publications Office. https://data.europa.eu/doi/10.2811/292617

General Data Protection Regulation. (2023). Retrieved March 30, 2023, In Wikipedia. https://en.wikipedia.org/w/index.php?title=General_Data_Protection_Regulation&oldid=1146591268

General Data Protection Regulation (GDPR) – Official Legal Text. (n.d.). General Data Protection Regulation (GDPR). Retrieved February 3, 2023, from https://gdpr-info.eu/

Govindarajan, V., Srivastava, A., & Enache, L. (2019, December 18). How India Plans to Protect Consumer Data. Harvard Business Review. https://hbr.org/2019/12/how-india-plans-to-protect-consumer-data

Herault, G. *What You Need to Know About Cookies and Web Beacons*. Monster.Com. Retrieved March 31, 2023, from https://www.monster.com/career-advice/article/cookies-and-web-beacons

India: Comparing the Digital Personal Data Protection Bill, 2022 and the GDPR. (2023, January 24). DataGuidance. https://www.dataguidance.com/opinion/india-comparing-digital-personal-data-protection-0

Jones, A. (2022). *GDPR Three Years Later: What Impact Has It Made? | I.S. Partners*. Retrieved March 29, 2023, from https://www.ispartnersllc.com/blog/gdpr-one-year-later-impact/

Kulkarni, S., Konde, K., & Bedekar, M. (2022). Analyzing Data Privacy Concerns in Young Adults—Apprehensions of Engineering Students. Grenze International Journal of Engineering & Technology (GIJET), 8(2), 411–418.

Deva Prasad M, Suchithra Menon C. (2020). The Personal Data Protection Bill, 2018: India's regulatory journey towards a comprehensive data protection law. International Journal of Law & Information Technology, 28(1), 1–19. https://doi.org/10.1093/ijlit/eaaa003

Mahawar, S. (2022, May 21). Is there an Indian equivalent for GDPR and what are the laws in India concerning personal data. IPleaders. https://blog.ipleaders.in/is-there-an-indian-equivalent-for-gdpr-and-what-are-the-laws-in-india-concerning-personal-data/

Mueller, B. (2022, April 9). A New Study Lays Bare the Cost of the GDPR to Europe's Economy: Will the AI Act Repeat History? Center for Data Innovation. https://datainnovation.org/2022/04/a-new-study-lays-bare-the-cost-of-the-gdpr-to-europes-economy-will-the-ai-act-repeat-history/

Pandey, A. S., Dixit, N., & Sagar, M. (2020). Data Protection Framework for India. Telecom Business Review, 13(1), 36–46.

Price, D. (2015, March 17). Infographic: How much data is produced every day? CloudTweaks. https://cloudtweaks.com/2015/03/how-much-data-is-produced-every-day/

Schufrin, M., Reynolds, S. L., Kuijper, A., & Kohlhammer, J. (2021). A Visualization Interface to Improve the Transparency of Collected Personal Data on the Internet. IEEE Transactions on Visualization & Computer Graphics, 27(2), 1840–1849. https://doi.org/10.1109/TVCG.2020.3028946

Selvadurai, N., Kisswani, N., & Khalaileh, Y. (2019). Strengthening data privacy: the obligation of organisations to notify affected individuals of data breaches. International Review of Law, Computers & Technology, 33(3), 271–284. https://doi.org/10.1080/13600869.2017.1379368

Sheth, H. (2021, March 14). 84% Indian consumers willing to pay more to do business with organisations committed to protecting data privacy: Report. https://www.thehindubusinessline.com/info-tech/84-indian-consumers-willing-to-pay-more-to-do-business-with-organisations-committed-to-protecting-data-privacy-report/article34066225.ece

Weigl, M. (2016). The EU General Data Protection Regulation's Impact on Website Operators and eCommerce. Computer Law Review International, 17(4), 102–108. https://doi.org/10.9785/cri-2016-0403

What is the GDPR, the EU's new data protection law? (2018, November 7). GDPR.Eu. https://gdpr.eu/what-is-gdpr/