As per UGC guidelines an electronic bar code is provided to seure your paper

International Journal for Modern Trends in Science and Technology, 8(04):380-388, 2022

ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0804064
Available online at: http://www.ijmtst.com/vol8issue04.html

Perception of privacy in a data driven world

Sakshi Kulkarni | Dr. Mangesh Bedekar

School of Computer Engineering & Technology, Dr. Viswanath Karad MIT World Peace University, Pune, Maharashtra, India
*Corresponding Author Email ID: kulkarnisakshi26@gmail.com

To Cite this Article
Sakshi Kulkarni and Dr. Mangesh Bedekar. Perception of privacy in a data driven world. International Journal for Modern Trends in Science and Technology 2022, 8(04), pp. 380-388. https://doi.org/10.46501/IJMTST0804064

Article Info
Received: 18 March 2022; Accepted: 15 April 2022; Published: 21 April 2022.

ABSTRACT

The volume of private data captured from users has increased drastically in the last few years. User data, particularly personal data, these days is digitally integrated and evaluated owing to the significant advancements in database, networking, machine learning & artificial intelligence. On one hand, this has resulted in the creation of data mining technologies aimed at inferring meaningful trends from this data. On the other hand, easy access to personal data puts individuals' privacy at risk.
Data Anonymization is a useful privacy protection approach that is utilized in a variety of technology domains, including data mining, cloud computing, and big data, to safeguard highly sensitive data from third parties. The paper covers the privacy paradox, the relationship between personalization, privacy, and anonymization.
This paper highlights the importance of personal data, its privacy, and implications of leaking of personal data. It highlights a few key techniques, tools and methods that users can use to safeguard their privacy. It also throws light on the legal frameworks for protection of user's data and maintaining data privacy from an Indian context namely the personal data protection bill to be tabled in the Parliament soon.

KEYWORDS: ***Personalization, Privacy, Anonymization, Privacy Paradox, Data Privacy***

1. INTRODUCTION

The process of adapting products, services, and information to the interests and preferences of users is known as personalization. Organizations will inevitably need to collect rich user data profiles in order to deliver relevant personalized offerings. According to previous studies, tailored services result in favorable user reactions such as greater desire to share personal data and make purchases. Another line of research focuses on the negative consequences of personalization, such as privacy concerns about the use of personal data and resulting detrimental effects on behaviors. Users can get personalized lists of recommended products from Amazon.com, for example, based on previous transactions and aggregate user purchase tendencies [12]. The absence of identifying information about a person is referred to as anonymization [18]. User anonymity is vital in the digital age because computers may deduce users' lifestyles, whereabouts, habits and associations from data acquired in various everyday transactions. Nevertheless, simply deleting explicit identifiers may not be enough to ensure security. The primary reason is that when released information is paired with freely available data, it can reveal an individual's identify. The Netflix crowdsourcing challenge is a well-known example. Netflix made a data set of users and their movie ratings public in 2012. The data could be downloaded and analyzed for patterns. A phony user ID was included in the data, as well as the movie, the users rating of the movie, and the date of the rating. The information released would not violate user privacy, according to the argument, because user identification has been deleted.

When the Netflix dataset is paired with additional supplementary data, Narayanan and Shmatikov (2008) demonstrated how users can be recognized (such as data from IMDB) [13].

2. What is data privacy, exactly?

Data privacy is a type of data protection that focuses on correct data processing while conforming to data security requirements. Data privacy refers to how data should be captured, stored, shared and managed with third parties, as well as ensuring that privacy standards are followed (such as GDPR or CCPA).

3.How is Personal Data Safeguarded?

There are several laws, policies, acts, regulations indicating how data is to be protected. These laws are mandatory for organizations to follow. Some of the prominent ones are listed below:

1. The GDPR

The General Data Protection Regulation (GDPR) covers 6 general data protection standards (purpose limitation, fairness and lawfulness, accuracy, storage limitation, data minimization, and integrity and confidentiality), but data protection by design and default is at the heart of the regulation. On the one hand, transparency and accountability support it [8].

2. The CCPA

The CCPA California Consumer Privacy Act of 2018 gives users added power over personal data that administrations collect from users, and the CCPA regulations spell out how the law will be applied. Users in California now have new privacy rights, including the right to know what personal information a company collects and how it is used and shared; the right to have personal data about them deleted (with some exceptions); the right to opt-out of revealing their lives; and the right to be treated fairly if they exercise their CCPA rights. Users must receive clear notices describing their privacy practices from organizations. The CCPA covers a wide range of organizations [1].

4. How can an individual ensure personal data privacy?

Maintaining the security of your personal information is critical in today's digital world. Hackers and cybercriminals have a plethora of ways to gain access to and misuse your data, so be sure you're doing everything you can to prevent a breach.

Here are some helpful hints and methods for keeping your data private:

- Make a copy of your data - It's vital that you not just password-protect but also back up your data. Many people underestimate the need of backup in protecting themselves from a data intrusion.
- Activate Multi-Factor Authentication or Two-Step Authentication - Double-down on security features like multi-factor authentication for logins to avoid a tragic breach.
- Examine the Domain of the Email Address - Employees are being asked to do anything when thieves create bogus email addresses. We've received hundreds of emails purporting to be from me, requesting that our accounting team send money to a different account, for example. Check that the email address domain matches the domain from which the email should be sent (for example, if you work at Microsoft, the email should finish in @microsoft.com).
- Take a look at the EULA - We've given up our privacy in exchange for convenience, and we're unwittingly handing over our personal information for free. So, always read the End User License Agreement (EULA) and make sure you know what you're giving up. Who owns the data, what will be done with it, and where will it be kept? Understanding these things is important.
- Passwords should never be reused - The most typical blunder is using the same password for all of their logins and social media accounts. That is the simplest way for someone to break into your system and have access to all of your information. It's critical to keep passwords that aren't the same to ensure security [5].

5.Personalization and Privacy

Technology has the power to understand us better than we understand ourselves, allowing them to design the best user experiences and propel our lives to new heights. Personalization can take many forms, including demographic filtering, content-based filtering, collaborative filtering, social network tagging, query and click-through history, community search trails, implicit relevance feedback, and hybrid techniques. Privacy and customization, user attitudes about privacy, and technologies and architectures that can aid in the creation

of privacy-preserving systems, such as e-learning, are all concerns of the user modelling and online personalization community. There is no need to explore Netflix's library because it already knows what users wanted to watch based on anything else that would indicate your mood. We live in a time where reliance on technical tools is unavoidable. Simultaneously, privacy-related issues are surfacing to the point where users are on the verge of losing control over data. Data sharing on social media sites may lead to privacy and security breaches. Furthermore, being on the Internet necessitates the ongoing exchange of information, whether personal or not. Because users value personalization, the personalization–privacy conundrum persists because sellers' exploitation of users' personal data to achieve personalization raises concerns about privacy. Users might therefore refuse to share private data, limiting personalization efforts. Attempts to use information technology to address privacy concerns (e.g., anonymizing techniques, peer-to-peer communication) have largely failed. The personalization–privacy conundrum describes the ongoing conflict between a company's demand for user data in order to tailor user experiences and a user's need for privacy. With the advancement of modern technology, the tension has grown to the point where users may consent to companies accessing users' personal data through technology-based channels without fully comprehending the terms of their consent. Companies typically utilize the data to tailor users' online navigation and improve their online experience, but the potential for misuse raises privacy issues. A more thorough theoretical framework may be required to strike a compromise between acquiring personal data to deliver meaningful, speedy user targeting while also minimizing users' privacy concerns. That is, while the paradox has apparent implications for user and firm behaviour, the linked trade-offs also have ramifications for platforms, particularly those that rely on the FAANG model. Such FAANG-type organizations have introduced numerous user privacy protections in order to deal with the perplexing dilemma. Experts have also suggested ways to get around it, such as utilizing modern Information and Communication Technologies (ICT) (e.g., anonymizing techniques, peer-to-peer communication), which can be sophisticated and thus difficult to employ [3].

6.THE PRIVACY PARADOX

Although users frequently express privacy concerns, it has been seen that these worries do not always correspond to actual disclosure. This pattern has been identified as puzzling and warranting additional examination. Several examples of this conflict between privacy attitudes and conduct can be found in the literature. Users concerned about their privacy, for example, are willing to share personal information in order to customize their online purchasing experience, or for convenience or discounts. Consider the discrepancy in framing between stated worries and observed behavior to have a better understanding of this conundrum. "Situational conditions, such as those relating to a specific Web site, may override the influence of general privacy concerns." according to one possible explanation for the conundrum [2].

7.WHAT CAN WE DO TO GET AROUND THE PRIVACY PARADOX?

Change your laptop, phone, and other devices' default settings to more privacy-friendly options, and look into your app settings for similar alternatives.

- Use a browser that prioritizes privacy, such as Firefox or Brave.
- Use DuckDuckGo or another non-spying search engine to find what you're looking for.
- Browser plugins that block trackers, such as uBlock Origin or Privacy Badger, can be downloaded.
- Choose a VPN that shares your ideals.
- To protect your accounts, use a password manager like KeePass, KeePassX, or BitWarden.
- Wherever practical, use two-factor authentication or multi-factor authentication.
- Find and delete old accounts that you no longer use.
- Use alternatives instead of Google, the market leader in surveillance capitalism [11].

8.WHAT ARE THE WAYS TO ADDRESS THE PRIVACY CONUNDRUM?

To overcome the privacy paradox, a societal effort will be required, beginning with laws and government rules, as well as new tech sector standards. The law's function in restoring individual privacy is twofold: first, it expands individual rights and protections, making it easier for people to understand and assert their rights,

and second, it overhauls how the tech industry is controlled. With the introduction of the EU's GDPR and the CCPA in 2018, progress was made, but national law in the United States of America and many other nations has yet to follow [11].

9.Techniques for maintaining Data Privacy - Anonymization Techniques

1. K-anonymity

The eventual goal of many privacy-preserving technologies is maintaining anonymity of the end users. When taken at a face value, anonymity simply implies being nameless, but a closer examination quickly reveals that eliminating names from a dataset is insufficient to establish anonymity. Data that has been anonymized can also be re-identified by connecting it to any other dataset. Quasi-identifiers, which are pieces of information which are not unique identifiers can be used to identify people when paired with other datasets. For example, 87% of the population of the United States of America can be uniquely identified with only gender, zip code and DoB. In order to achieve k-anonymity, data should contain at least k individuals who share a set of qualities that could be used to identify each individual. K-anonymity can be thought of as a "hiding in the crowd" guarantee: if each person is a member of a bigger group, any of the records in that group could belong to a single person. The sensitive columns of interest must not expose information that was redacted in the generic columns, according to the K-anonymity technique.

K-anonymity technique comes with the following conditions:

1. Information suppressed in the generic columns must not be revealed in the sensitive columns of interest.

2. For a given group of k, the values in the sensitive columns are not all the same.

3. The data's dimensionality must be kept to a minimum [10].

2. Generalization and Suppression

The idea of k-anonymity, as well as its formalization and enforcement via generalization and suppression, two microdata protection mechanisms that preserve information veracity, has been introduced. The process of replacing an attribute's values with more general ones is referred to as "generalization" (e.g., The year of birth can be used instead of the date of birth.). Generalization is based on a domain generalization hierarchy and a related value generalization hierarchy on domain values in the original k-anonymity concept. The domain generalization hierarchy is a tree, whereas the value generalization hierarchy is a total order, with the parent/child relationship representing the direct generalization/specialization relationship. At various levels of granularity, generalization and suppression can be used.

3. Slicing

Slicing is a data partitioning technique that divides data into vertical and horizontal partitions. A vertical partition is a set of qualities in a column that are related to each other. A horizontal partition divides a collection of tuples into buckets, with each bucket containing a column of randomly permitted values.

4. L diversity

Another method known as L diversity has been devised to overcome homogeneity attacks. Each equivalence class must have “L” well represented values for the sensitive property, according to L diversity. Due to the variety of data, implementing L diversity is not always possible. Skewness assault can have a negative impact on L diversity. Attribute disclosure cannot be assured when the overall dispersal of data is slanted toward a few equivalence classes. For example, if all of the data is classified into only three equivalence classes, semantic similarity between these values may lead to attribute disclosure. L diversity may also induce similarity assaults.

5. T-closeness

Other enhancement to L diversity is the T-closeness, which determines whether an equivalence class has “T-closeness” if the distance between distributions of sensitive characteristics in the class is less than a threshold, and whether all equivalence classes have T closeness. With respect to sensitive attributes, T-closeness can be calculated on any attribute. T closeness may assure attribute disclosure, although it is possible that T-closeness will not always result in proper data distribution.

6. Differential Privacy

Adaptive attacks that exploit auxiliary information are resistant to differentially private algorithms. This technique incorporates random noise into the mix, making everything an adversary (attacker) receives chaotic and imprecise, making privacy breaches

considerably more difficult. Differential privacy promises to protect individuals from any additional harm they may suffer as a result of their data being in a private database that they would not have suffered if their data had not been included. Differential privacy does not guarantee that one's secrets remain concealed. It promises to keep the data differentially secret and not divulge it, but it does not guarantee that it will be protected from attacks. One of the most popular types of privacy attack is the differential attack. It just ensures that user's participation in a survey will not result in the disclosure of any specifics that one has contributed to the survey if kept differentially private, and that participation will not result in the disclosure of any specifics that one has contributed to the survey if kept differentially private. Differential privacy is a query property rather than a database property. The amount of noise introduced to make the data secret is important since the more noise added, the worse the model utility or accuracy is.

Limitations of Differential privacy

Differential privacy is a notion that has piqued the interest of academia and the research community, but it has gotten less attention from industry due to its strong privacy guarantee. Noise must be injected k times if the initial assurance is to be maintained across k queries. When k is big, the output's usefulness is reduced. More noise is required for a succession of searches, which eventually exhausts the privacy budget and may lead to the user's death. Meaning that once a user's privacy budget is depleted, user is no longer allowed to ask any more questions, and if you allow for user cooperation, you'll run into problems with what this privacy budget signifies on a per-user basis.

8.Techniques for maintaining Data Privacy - Online Data Anonymization Tools

1. ARX

https://arx.deidentifier.org/

ARX is a huge open-source application that encrypts sensitive personal information. ARX covers a wide variety of privacy and risk models, as well as methods for modifying data and evaluating the utility of output data. The application has been used in a variety of settings, including commercial big data analytics systems, academic initiatives, clinical trial data interchange, and training. ARX can handle large datasets on commodity hardware and includes a cross-platform graphical user interface.

2. Amnesia

https://amnesia.openaire.eu/

The Athena Research Center developed Amnesia, a data anonymization tool. km-anonymity and k-anonymity are supported. Amnesia's hierarchy generator and editor allow users to customize anonymization to strike the balance between privacy and data utility. Users can install Amnesia from the website.

3.μ-ARGUS https://research.cbs.nl/casc/mu.htm

ARGUS is a tool for creating secure microdata files that is based on the R programming language, which was created primarily for statistical analysis. 'Anti Re-identification General Utility System' is the acronym for 'Anti Re-identification General Utility System.' The program employs a variety of statistical anonymization techniques including randomization, noise addition, local suppression, global recoding (category grouping), micro aggregation, and top- and bottom coding. Can also be used to make up fake information.

4. sdcMicro

https://cran.r-project.org/package=sdcMicro

sdcMicro is a package which comes under R programming language which allows users to create anonymized (micro)data. Various risk estimation approaches are also presented. It should be noted that package contains a GUI that allows user to utilize the program's numerous methods. In May of 2018, sdcMicro was released [15].

5. Anonimatron

https://github.com/realrolfje/anonimatron

Anonimatron is a collection pseudonymization technique that may be used to generate pseudonymized production data and undertake performance testing outside of the user's environment. Along with the implementation of the GDPR (General Data Protection Regulation), a new feature allowing users to anonymize files was added: "Users can now designate a column to be anonymized without preserving the created synonyms for future runs." [15]

6. TOR Browser https://www.torproject.org/download/

The Tor Project is a charitable organization. committed to the study and advancement of online privacy and anonymity. Its goal is to keep anyone, including government organizations from tracking users browsing history. It proposes a system based on that study that bounces internet users' and websites' traffic flow around "relays" operated by volunteers all over the world, making it tremendously impossible for anyone to pinpoint user's location [19].

7. BRAVE Browser https://brave.com/download/

The privacy-focused browser has a slew of useful features that make it a strong competitor. Brave by default bans tracking scripts and resources, as well as advertising-related tracking scripts and resources. Brave protects users from monitoring based on your fingerprint in the browser by randomizing it, giving users a very good tracking protection. Brave is the only popular browser that by default disables both third-party cookies and third-party storage (such as storage in iframes).

9.Personal Assistants and Data Privacy

Voice-activated personal assistants are becoming increasingly popular, to the point where they are frequently users first point of entry into the home Internet of Things (IoT) ecosystem. Siri from Apple, Cortana from Microsoft, Alexa from Amazon, and Assistant from Google are all software agents that run on purpose-built speaker gadgets or smartphones [9].

10.Things you can do with your Voice Assistants

Asking your digital assistant—such as Google Home—for today's weather, a witty joke, or a reminder using your voice. Using your voice assistant, Amazon Alexa, to ask your Nest smart thermostat to lower the temperature. Inquiring about today's news with Apple's Siri [20]. Even the most private talks can be recorded and stored in the server for later analysis. Homebased assistant gadgets have a large number of sensor nodes and other attributes that may collect a great amount of information which was once thought to be private. In addition to voice activity, smart homebased gadgets are related to cloud-based services which can capture other behaviors facilitated through such systems, such as news or online purchase behavior and entertainment preferences. Data collecting occurs passively (as a background operation) due to the "always on" listening component that identifies the device's initiation keyword ("Alexa", "OK Google") often without individuals' knowledge or understanding [7][17]. PwC questioned 1,000 Americans aged 18 to 64 in 2018 to discover more about their understanding of speech technologies. The results were astounding: 90% of respondents stated they were already familiar with voice assistant technology, with 57 percent using it on their smartphone and 20-30% using it on other devices such as tablets, laptops, speakers, and TV remotes.

11.Why should a privacy-conscious person not have an assistant device in their home?

- Always on - The fact that these gadgets are always on and listening should terrify internet users who cherishes privacy. An assistant, lacks an "offline" switch, and the only way to turn it off is to unplug it - and even that won't work if it has a backup battery [20].
- Always Listening - It is not an exaggeration to say that an assistant listens everything said and must process that speech in order to function. Because the device is programmed to begin listening when it hears a trigger keyword, it should also upload voice recordings to its server on a regular basis in order to figure out what is being said. While the company promises that voice recordings are not stored on users' computer systems or used for any other intent, there have been numerous amounts of reports from users who claim to have received targeted ads purely based on what they've discussed with someone else while near an assistant device [20].

12.What measures are these companies taking to provide users data privacy from virtual assistants?

"User trust is at the heart of everything that we do at Amazon," an Amazon spokesperson said, "and we take security and privacy extremely seriously." We've always believed that privacy should be a fundamental component of every technology, product, or service we develop. We've always prioritized users, and we're continually exploring for new ways to make it easier for them to manage their Alexa experience. Among the new privacy features include the ability to have voice recordings automatically destroyed after three or 18 months on a regular basis, the ability to ask Alexa to

"delete what I've said" and "delete what I said today," as well as the Alexa Privacy Hub, a global resource.

Apple's Intelligent Tracking Prevention (ITP)

ITP is a function of the Safari browser, which in countries like the United States and the United Kingdom has a 50% mobile share of the market. It enhances the limitations on JavaScript and now Local Storage cookies, limiting them to a seven-day lifespan. As a result, any data on the visitor stored in cookies (and now Local Storage) is immediately erased after seven days [16].

Google Chrome and Mozilla Firefox

Google has said that third-party cookies would be banned in Chrome by 2022, but this will have no effect on experimentation platforms. Third-party cookies were blocked in June 2019 by Mozilla Firefox's Enhanced Tracking Protection (ETP) technology, which works similarly to adblockers [16].

13.Personal Assistants - Can you protect yourself?

Unfortunately, the only way to protect yourself from smart speaker data breaches is to avoid using them. If users are truly concerned about the risks, one should probably avoid using smart assistants and devices altogether. You'll be safer than the typical user if you keep a tight check on which gadgets you use, know what you may opt into and out of, and understand how your data is utilized [6].

14.What comes first? What is more important, personalization or privacy?

Use or abuse of data?

People are becoming more aware of the value of their personal information. Users have taken to heart the adage that "if you aren't paying for a product, you are the product," particularly in relation to social media. Brands that want to leverage user data to produce more tailored services are running up against a tide of opinion that claims, whether or wrongly, that they acquire personal data primarily to sell to third parties or to spam users with more marketing. Real-time data from a variety of sources, such as social media platforms, point-of-sale systems, and financial transactions, is critical in helping brands define their users' behaviors and motives. According to recent reports, Facebook shares user data in ways that make people feel vulnerable and exploited.

Do Users Really Want as Much Personalization as We Believe?

According to certain studies, personalization isn't as crucial to users as many experts believe. Understanding what your users desire in a comfortable, non-intrusive approach is the key to personalizing without going overboard. Build a loyalty language with your users so they understand why you're asking for particular information and how you'll protect their privacy. Personalization is crucial, but in many circumstances, users value security and professionalism more. As a general guideline, if the information was not willingly provided by the user, it should not be used for personalization. What is the point of personalization techniques if they jeopardize user trust? How would this enable individuals to feel at ease or establish loyalty? Why would users suggest your brand over one that requires them to supply less information in exchange for a similar experience? Here are some pointers to help you achieve a balance between personalization and privacy protection for your users:

- Ask for the information up front and explain why you require it as well as what you intend to do with it.
- Use trust badges on web site to demonstrate information will be stored securely and carefully.
- Allow people to choose whether or not to answer questions and to have complete control over their personal privacy.
- Do not assume what kind of information users are willing to offer.
- Inform users of the benefits they would receive in exchange for their information, such as better products, a better experience, and so on. This deal must be skewed in favor of the payoff that will be received by users. In life, it's a good rule of thumb to give more than you take [4].

15.The legal perspective in India – The Personal Data Protection Bill

No one can be trusted on the Internet in this digital age. The main challenge is determining when and how an individual or organization can exchange personal data with others for processing, as well as how that data's privacy will be preserved. These concerns necessitate the creation of a legal framework to preserve individual

privacy, protect personal data, and prevent data breaches. To ensure this, the government of India formed a team led by Justice B. N. Srikrishna to examine the current difficulties and develop a legislative data protection framework to address individual data privacy. The "Personal Data Protection Bill, 2018" is a comprehensive data protection bill that's been proposed. Both from a policy and an intellectual standpoint, the progress of the Bill and its potential ramifications are crucial. This would be a significant step forward in India's digital rights regulatory framework. In the Indian context, the PDP Bill introduces the notion of Data Principals and Data Fiduciaries. Personal data is defined in a broader sense in the PDP Bill, and it is not limited to sensitive personal data [14].

16.CONCLUSION

Users are increasingly appreciating personalized service. Users, of course, want to preserve their privacy, but we also want to take full advantage of technology's benefits. Personalization is usually at the expense of user privacy. As a result, for companies working in the information-centric, global networked economy, privacy has become a strategic concern. It is critical for each and every user to understand how information about them is captured, saved, processed and managed. If users are aware about how data about them is processed, as an individual the user may be able to safeguard their personal data privacy.

Users and Businesses must modify and accept responsibility for their activities in order to coexist in the Privacy Paradox. Users have to pay attention to what they are committing to and shift their thinking to consider the long-term ramifications of sharing their personal information on the Internet, not just the short-term benefits. The t-closeness principle has been accepted as an upgraded concept that fixes the primary disadvantages of k-anonymity and l-diversity, which include weaknesses against homogeneity, background knowledge, skewness, and similarity assaults, as a result of the comparison of the anonymity models.

The future of data anonymization includes new strategies for systems that are universal for any type of data and that also address all types of privacy attacks. Privacy cannot be a late addition it has to be incorporated right from the design of systems and applications. Privacy by design is becoming a key concept in privacy laws; it suggests that every product or plan that is developed should address privacy right from the beginning, not as an afterthought.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] California Consumer Privacy Act (CCPA). State of California - Department of Justice - Office of the Attorney General. (2022, March 28). https://www.oag.ca.gov/privacy/ccpa

[2] CiteSeerX — the ultimate display - Pennsylvania State University https://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.136.3720

[3] Cloarec, J. (2020). The personalization–privacy paradox in the attention economy. Technological Forecasting and Social Change, 161. https://doi.org/10.1016/j.techfore.2020.120299

[4] Dudharejia, M. (2022, January 20). Which is more important to consumers: Personalization or privacy? A/B Testing Software. https://www.convert.com/blog/personalization/personalization-too-much/

[5] Editors, F. T. C. (2020, May 11). Council post: 14 personal data security tips for everyday users. Forbes. https://www.forbes.com/sites/forbestechcouncil/2020/05/11/14-personal-data-security-tips-for-everyday-users/?sh=675264e54663

[6] English, T. (2020, January 31). What data are voice assistants collecting and how to protect yourself. Interesting Engineering. https://interestingengineering.com/what-data-are-voice-assistants-collecting-and-how-to-protect-yourself

[7] Fruchter, N., & Liccardi, I. (2018). Consumer attitudes towards privacy and security in home assistants. Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. https://doi.org/10.1145/3170427.3188448

[8] Goddard, M. (2017). The EU general data protection regulation (GDPR): European regulation that has a global impact. International Journal of Market Research, 59(6), 703–705. https://doi.org/10.2501/ijmr-2017-050

[9] Hoy, M. B. (2018). Alexa, Siri, Cortana, and more: An introduction to voice assistants. Medical Reference Services Quarterly, 37(1), 81–88. https://doi.org/10.1080/02763869.2018.1404391

[10] K - anonymity: An introduction. Privitar. (2021, January 3). https://www.privitar.com/blog/k-anonymity-an-introduction/

[11] Ker, A. D. Decoding the privacy paradox. The Privacy Issue. https://theprivacyissue.com/privacy-and-society/decoding-privacy-paradox

[12] Lee, C. H., & Cranage, D. A. (2011). Personalisation–privacy paradox: The effects of personalisation and privacy assurance on customer responses to Travel Web Sites. Tourism Management, 32(5), 987–994. https://doi.org/10.1016/j.tourman.2010.08.011

[13] Lu, X., & Au, M. (2016). [PDF] an introduction to various privacy models: Semantic scholar. [PDF] An Introduction to Various Privacy Models | Semantic Scholar. https://www.semanticscholar.org/paper/An-Introduction-to-Various-Privacy-Models-Lu-Au/ae9c0a86daec5a5d0d62aab322feefbc4d0309b2

[14] Prasad M, D., & Menon C, S. (2020). The Personal Data Protection bill, 2018: India's regulatory journey towards a comprehensive

data protection law. International Journal of Law and Information Technology, 28(1), 1–19. https://doi.org/10.1093/ijlit/eaaa003

[15] Sartor, N. (2019, September 25). Top 5 free data anonymization tools. Aircloak. https://aircloak.com/top-5-free-data-anonymization-tools/

[16] Todaro, F. D. (2020, May 12). Balancing privacy and personalization - how can marketers stay compliant. Kameleoon. https://www.kameleoon.com/en/blog/balancing-privacy-personalization

[17] Voice assistants - PWC. https://www.pwc.com/us/en/advisory-services/publications/consumer-intelligence-series/voice-assistants.pdf

[18] What is Data Anonymization: Pros, Cons & Common Techniques: Imperva. Learning Center. (2020, December 31) https://www.imperva.com/learn/data-security/anonymization/

[19] What is tor? A beginner's guide to the privacy tool - the Guardian. theguardian.com. https://www.theguardian.com/technology/2013/nov/05/tor-beginners-guide-nsa-browser

[20] Why home assistant devices are a privacy nightmare. hide.me. (2019, March 18). https://hide.me/en/blog/assistant-devices-are-a-privacy-nightmare/