Issues in Information Systems
Volume 22, Issue 2, pp. 177-186, 2021

DOI: https://doi.org/10.48009/2_iis_2021_185-195

Data privacy in the age of big data analytics

Christopher B. Davison, *Ball State University, cbdavison@bsu.edu*
Edward J. Lazaros, *Ball State University, ejlazaros@bsu.edu*
Jensen J. Zhao, *Ball State University, jzhao@bsu.edu*
Allen D. Truell, *Ball State University, atruell@bsu.edu*
Brianna Bowles, *Ball State University, blbowles@bsu.edu*

Abstract

The world is generating exabytes of information daily. This vast explosion in the growth of data creates a need for data analysis technologies as well as effective storage and retrieval systems. Relatedly, this vast amount of data presents privacy challenges ranging from legal constraints to ethical information use. The authors of this paper discuss a framework that enables privacy infusion in all aspects of big data analytics. Additionally, the authors present TIPPERS: a testbed for privacy enhancing technologies.

Keywords: Privacy, Big Data, Data Analytics, Design.

Introduction

The concept of big data holds the idea that data collection and databases continuously grow. Relatedly, the privacy, legal and security concerns expand as well. The research question that arises is: What privacy, legal and security issues arise from the collection and analysis of vast amount of information? The authors of this paper will present discussions on a number of these concerns within the context of big data analytics. The purpose of this research article is to describe and explore the privacy implications inherent in big data and to introduce the TIPPERS tool as a testbed for privacy research.

To present privacy in the big data context, the authors begin this paper by providing a literature review. The literature review will set the context and provide definitions associated with data privacy and information privacy as well as discuss selected laws governing data privacy aspects.

Following that, the authors provide a discussion of the Privacy by Design (PbD) concept. This is the idea that privacy should be engineered into systems at every stage of development. This concept is presented as an approach to ameliorate privacy issues before they surface in general usage markets.

Finally, a section on a large-scale, experimental sensor and data management framework known as the Testbed for IoT-based Privacy-Preserving PERvasive Spaces (TIPPERS) will be presented. An additional discussion on TIPPERS as a COVID-19 mitigation tool is provided as well. In this section, a discussion of individual privacy concerns as well as information granularity tradeoffs is presented.

Literature Review

Big Data

Big data can be defined as, "the exponential increase and availability of data in our world" (*What is Big Data?: University of Wisconsin*, 2019, para. 3). This entails data from smartphones, social media posts, sensors (e.g., traffic signals), point-of-sale terminals, consumer wearables, electronic health records and many other sources. The information within these platforms provide opportunities for systems to aggregate such data and come up with actionable insight, improved decision making, and competitive advantage (*What is Big Data?: University of Wisconsin*, 2019). Favaretto et al. (2020) found agreement amongst researchers who defined big data as large amounts of digital data produced from technological devices that require specific algorithmic or computational processes to answer relevant research questions.

A framework for the term can be explained via the Three V's (Volume, Variety, and Velocity). Volume refers to the magnitude of data. Large data sizes are documented in multiple terabytes and petabytes. A single terabyte stores the same amount of data that would fit on 1500 CDs or 220 DVDs, or even 16 million Facebook photographs. Volume, in reference to data, is relative and varies by factors such as time and type of data. The term *Big Data* today may not meet the threshold for the future due to increase in capacity. Big data is also dependent on industry usage due to requiring various data management technologies based on the type of data at hand. Currently, there exists petabyte size data warehouses (e.g., Ebay), with one petabyte is equivalent to 1,000 terabytes. The challenge with data volume focuses on how to identify relevant data within large data sets and utilize it.

Variety indicates the structural heterogeneity within a dataset (Favaretto et al., 2020). Technology advances allow use of structured data, semi-structured data, and unstructured data. With text, images, audio and video being examples of unstructured data, organizations acquire unstructured data from internal sources like sensor data and external sources like social media.

The final V, velocity, indicates the rate at which data are developed and the speed at which it should be analyzed and acted upon (Favaretto et al., 2020). Due to the rapid increase of digital devices, there is a never before seen rate of data, indicating a continuous increase for real-time data analytics and evidence-based planning (*What Is Big Data? | University of Wisconsin*, 2019).

Most recently, the 3-V's have been expanded upon and also include veracity, variability, visualization and value. Veracity indicates the quality of data collected, with the emphasis and reliance on computers and automated decision making, there's a large amount of trust in the quality of data. Variability referring to consistent change, relies on scientists to invent these sophisticated programs to understand various contexts and meanings. Visualization refers to creating graphs to inform scientists, transform data into information, transform information into insight, then insight into knowledge and knowledge into advantage. Lastly, value indicates how these organizations can use this data to make improvements in decision-making (*What Is Big Data? | University of Wisconsin*, 2019).

Big Data Analytics

Big data analytics is "the process of examining large data sets containing a variety of data types to uncover hidden patterns, unknown correlations, market trends, customer preferences, and other useful information" (NGDATA | What is Big Data Analytics? Learn About Tools and Trends, 2016, para. 1). For companies implementing big data analytics often benefit greatly and make informed business decisions. They tend to have more effective marketing campaigns, discover new revenue opportunities, improve customer services, increase efficiency in operations and improve competitively. There are also various types of tools used for big data analytics that help companies save time and money, and aid in obtaining insights to guide business decisions. These tools include data storage and management, data cleaning, data mining, data analysis, data visualization, data integration, and data collection (NGDATA | What is Big Data Analytics? Learn About Tools and Trends, 2016). Companies are able to analyze their data fully and quickly and offer real-time analysis, allowing them to optimize machine learning and address their big data needs in novel ways. They can narrow down their big data into the most applicable ways and analyze it to influence decision making.

Legal Environment of Big Data

There are many laws in the U.S. concerning the Internet, data security, and privacy, of which the Privacy Act of 1974 creates a solid foundation (NortonOnline). The Privacy Act established control over the collection, maintenance, use, and dissemination of personal information by agencies in the executive branch of the U.S. government. The following laws currently in place to solidify rights as a consumer include: Electronic Communications Privacy Act (ECPA), Computer Fraud and Abuse Act (CFAA), Cyber Intelligence Sharing and Protection Act (CISPA), and Children's Online Privacy Protection Act (COPPA).

The ECPA (1986) allows the U.S. government to access digital communications from companies such as emails, social media messages, information on public cloud databases, etcetera, without a warrant if the information desired is at least 180 days old. In addition, the ECPA also determines when the government is allowed access to GPS tracking via cellular devices. The CFAA was passed in the late 1980's and revised nearly ten years later. This act allows for protection against unlawful access and sharing of protected information. CISPA was introduced as an amendment to the National Security Act of 1947, which does not cover cyber crime. CISPA looks at how to share information on possible cyber threats with the federal government. Finally, in 2013 the final amendment was made to COPPA, which requires websites that collect information on children under 13 years of age to comply with the Federal Trade Commission (FTC). The FTC is responsible for investigating a website's language, content, advertising, graphics and features, and intended audience to make sure it is child-appropriate.

Andrea et al. (2015) detail the relationship between security and privacy with regard to data. Additionally, they provide a discussion of trust as it relates to user perception. Those researchers state that data security and privacy refers to the protection of any collected or stored data. This can be achieved by utilizing authentication, access control, data encryption, and data availability and redundancy through back-ups. They define trust as the reinforcement of the security goals previously mentioned, consisting of further objectives as well.

Privacy/Security and Big Data

Privacy, in Constitutional Law, is defined as the right of people to make personal decisions regarding intimate matters. Privacy under the Common Law is defined as "the right of people to lead their lives in a manner that is reasonably secluded from public scrutiny" (Right of Privacy, 2013, para. 1). The origins of the right to privacy are traced back to the nineteenth century to Samuel Warren and Louis Brandeis. In an 1890 law review essay, Louis Brandeis and law partner Samuel Warren coined the idea "right to privacy," aiming to protect a person's "inviolate personality," more specifically from new technologies (Nelson, 2011). The concept came about because of the invasion of privacy occurring with photojournalists prying into personal information. This derived conflict for the First Amendment protections of speech and the press, however the right to privacy already existed within the constitution (protections of one's home, prohibitions of the publication of one's private papers, and prohibitions against slander and libel). The goal of their argument allowed individuals to control what is *theirs* even if it is not a property right traditionally, nor is it a physical possession or trespass (Nelson, 2011).

The tabloid industry began profiting on distasteful interests disregarding modern morals and the harms of privacy fell into four main types; (1) intrusion into one's private life and affairs, (2) public disclosure of embarrassing private facts, (3) unwanted publicity of private individuals and (4) misappropriate of a name or likeness for financial advantage (Bycer, 2021). Warren and Brandeis (1890) went on to establish the injuries, potential remedies and basis for a right to privacy. The rule would protect individuals from publication of one's private matters with few exceptions; (1) privileged communications are the domain of libel/slander, (2) speaking gossip and oral communication are outside the purview of privacy rights, (3) consent to publication is an outright defense, (4) truth and (5) malic are irrelevant to a breach of privacy action (Bycer, 2021).

In today's age, companies are legally collecting information about their consumers such as, what they purchase, the medications they ingest, what websites they're visiting and credit histories. Computer software is then used to organize the data and prepare for sale to marketing companies, lending institutions, insurance companies and credit bureaus (Bycer, 2021). Examples as to how companies collect such data include shoppers cards, financial data, and motor vehicle data. Shoppers cards are used by grocery stores and retail businesses that offer discounts or premiums when shoppers utilize their cards. The purchases are scanned into the computer and the buyer's habits are then stored. The companies then use this information to aim marketing at certain customers or sell it to other companies seeking specific customers. Financial data collection is done by credit bureaus that collect credit histories compiled with personal information that is turned around and sold to anyone. Motor vehicle data is collected because an individual's motor vehicle registration is public information in the majority of states, as well as driver's license data. Allowing automobile dealers and insurance companies to have access to such data (Bycer, 2021).

The data trail of every individual leads to the term "data legacy" (Davison et al., 2020, p.34). All of the data collected about a person, anywhere from their debit card transactions, to the car they drive and the websites they visit are stored in the data world for extended periods of time, or even for eternity. Technology and social advances have created a flood of new digital applications and devices such as cell phones, websites, social media, smart household appliances, business software, industrial machines and smart cars, all of which generate an influx of big data. Because of this, developments in computational, storage and analytical technologies, appliances for handling and applying the data are also increasing in accessibility (Klievink et al., 2017).

Companies, governments and academia are three sectors largely benefiting from big data. Retailers like Walmart, Sears, and Amazon are using big data in order to better understand their customers and buying decisions. Financial institutions are using data analytics to predict market behavior and investment performance. Google, eBay, Twitter and Facebook revolve business models around volumes of digital data on individual behavior, information requests and preferences (Klievink et al., 2017). With these powerful analytical capabilities across the board, organizations are leading to an exponential increase in collection, storage, and use of personal data, putting pressure on traditional measures of privacy protection (Altman et al., 2018). Over time, commercial and government data begin to expose fairly detailed information regarding peoples' lives. The data is then made readily available to researchers, policy makers, and entrepreneurs in order to continue scientific research, public policy and innovation. Commercial data is used to provide goods and services to customers and enable analytics to improve services. While telecommunications providers, mobile operating systems, social media platforms and retailers often collect, store, and analyze large quantities of data about customers locations, transactions, usage patterns, interests, demographics, etc. (Altman et al., 2018). With that being said, the collection, storage, and use of personal data for prolonged periods of time is becoming the highlight of privacy, legal and policy issues.

Big Data and Data Analytics Privacy Exploits

One side effect of data analytics is the large masses of data that is squired (Volume in the 3-Vs model). Data sets are larger and growing larger still (Gantz & Reinsel, 2012). As larger repositories become the norm, privacy exploits become more of a threat. In the most benign case of privacy exploitation, data is repurposed by organizations beyond the original intent of the data collection (Walker, 2014). In more sinister cases of privacy exploitation, stalking and personal injury are possible.

Dinur and Nissim's (2003) work on the Fundamental Law of Information Recovery demonstrates that privacy can be lost with a few well designed queries into large data sets. This leads to the observation that *noise* can be injected into the data for privacy preservation purposes. Dwork et al. (2006) provided a formalized framework, the concept of ε-differential privacy, to quantify the amount of noise required to provide privacy. More recent work in the field (Wagh, et al., 2020) investigates privacy (injection of noise) versus utility (accuracy of information from a DP data set) and sensitivity of information release mechanisms (Laud et al., 2020).

In the worst case, privacy exploitation of data can be used for stalking or violence. Yang et al. (2012) provided research detailing online privacy invasion. They studied privacy attacks from two types of online stalkers: tireless attackers and resourceful attackers. In their study, the researchers concluded that users' private information was highly identifiable by stalkers. They further warn that this can be the case even with little or possibly inaccurate information.

In cases of intimate partner violence, Southworth et al. (2007) detail the use of a number of different technologies being utilized and abused by the perpetrators of this violence. The authors warn that data warehouses and even court records are utilized for nefarious purposes. Often, the courts publicize records and even full case documents without the knowledge or consent of the victims (Blankley, 2002). The attackers can use this widely-available information to circumvent protective orders and find their victims' latest locations.

Privacy and Big Data Analytics

Privacy by Design

In the previous sections, a number of privacy issues and privacy exploits in data analytics. One possible solution to these issues is to adopt the Privacy by Design (PbD) principles introduced by Cavoukian (2011). Li and Palanisamy (2018) discuss the benefits of adopting privacy preserving methods. These benefits can include reputation enhancement as well as legal protections.

PbD is a beginning-to-end framework that infuses privacy and privacy mitigation strategies in every aspect of the technology design process. Instead of a post-de-facto retrofit of privacy-flawed technologies, Cavoukian (2011) is a proponent of building in privacy by design principles (see Figure 1) from inception. This alleviates the need to retrofit privacy into technology once there is a privacy breach.

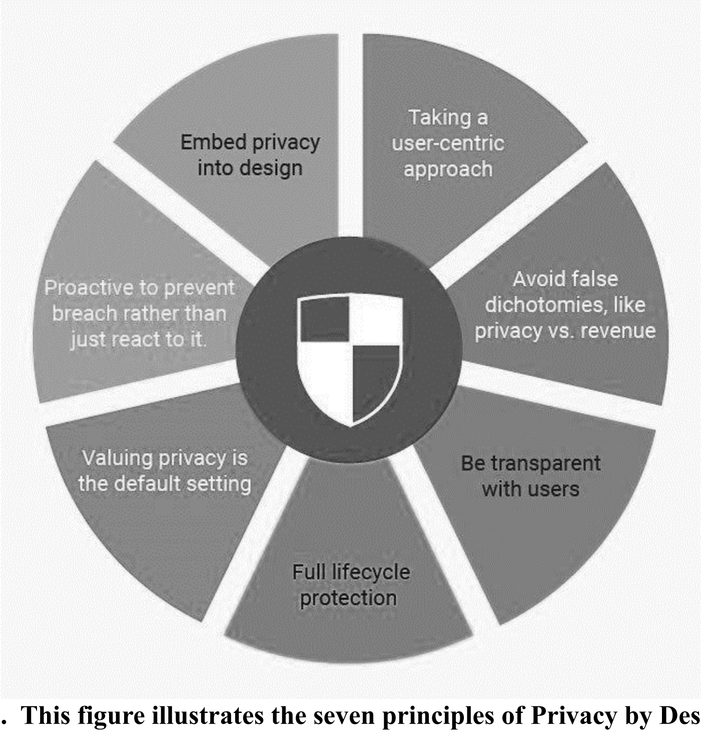
Embed privacy into design
Taking a user-centric approach
Avoid false dichotomies, like privacy vs. revenue
Be transparent with users
Full lifecycle protection
Valuing privacy is the default setting
Proactive to prevent breach rather than just react to it.

***Figure 1*. Privacy by Design. This figure illustrates the seven principles of Privacy by Design (Mackie, J., 2018).**

PbD is crucial in decreasing privacy risk and increasing trust. The creation of systems, products, processes and projects using this infrastructure upholds various beneficial aspects. First, identification of potential problems in the early stages and addressing them quickly. Second, it strengthens data protection and privacy across organizations. Lastly, organizations meet legal obligations instead of reacting to data privacy breaches.

The PbD principle is gaining widespread attention and is becoming the dominant privacy protection paradigm for information engineers according to Bu et al. (2020). However, these researchers caution that adoption of PbD paradigms is lacking. They discuss the problem from two perspectives. The first is that PbD does not have specific implementation protocols defined within its principles. The second is that systems engineers are not adequately incentivized to adopt the design principles. The authors recommend a well-designed incentive plan to entice the technology industry to focus more on PbD concepts and incorporate the concepts into the product development process.

TIPPERS: A Testbed for Privacy Preservation

The Testbed for IoT-based Privacy-Preserving PERvasive Spaces (TIPPERS) is deployed in a number of places including University of California, Irvine (UCI); University of California, San Diego (UCSD); The California Institute for Telecommunications and Information Technology (CalIT2); and Ball State University (BSU). TIPPERS is a system that manages IoT smart spaces by collecting sensor data, inferring semantically meaningful information from it, and offering such inferences to developers to create smart applications.

Sensor data collection, processing, and sharing leads to potential violations of privacy given the sensitivity of the data and the inferences that can be extracted from it. TIPPERS integrates different Privacy Enhancing Technologies (e.g., policy-based access control, differential privacy, secure computing, etc.) to deal with privacy issues in IoT data management.

Recently, TIPPERS was deployed in the US Navy's Trident Warrior exercises. Military environments present unique circumstances for data analytics and privacy. As the US military moves toward the Internet of Battlefield Things (IoBT) and smart, sensorized environments (Castiglione et al., 2017), privacy and the perception of privacy may degrade. Archer et al. (2020) discuss the DARPA-funded TIPPERS system being deployed on board a Navy vessel. The researchers found that privacy was a concern across the entire chain of command spectrum. From the newly enlisted to the career officers, privacy and the perception of privacy was a continual discussion and topic of concern. This concern was especially prevalent within the domain of onboard IoT and human-sensing devices.

TIPPERS is designed as a testbed to test privacy-enhancing technologies. Often, the trade-off exists between data privacy and utility of information. Data analysts and other information consumers require actionable information. However, users of IoT-based technologies require a varying degree of privacy that is often contextually based. An example, tested within the TIPPERS testbed, would be emergency response.

Building occupants, during the normal course of their routine, usually have a high expectation of privacy and data privacy (e.g., location within the building, activity specifics, with whom to share schedule information, time spent in location, etc.). In the TIPPERS system, these privacy preferences can be set by the user. Additionally, there is an opt-in/opt-out model so that users may choose to not participate in the system at all.

Concomitant with privacy specifications, users can also specify emergency (e.g., disaster response) settings. These settings provide location and other situal awareness information during an emergency event. This information would be provided to first responders. building managers, or other response coordinators (having proper authorization and credentialing). The idea is that concerns for privacy are overridden by concerns for rescue during a crisis response event.

TIPPERS Privacy and Privacy Policy Enforcement Mechanisms

The TIPPERS system provides a number of Privacy Enhancing Technologies (PETs) as well as policy enforcement mechanisms. The system is designed to allow users to specify their privacy preferences (as policy) and the system makes use of a policy engine to provide privacy guarantees.

Users are provided with an interface in order to state their privacy policy and preferences. This information is incorporated into the TIPPERS policy engine. The policy engine consists of a policy manager and a policy enforcer. The policy manager tracks users policy updates or adds. The policy enforcer works with the TIPPERS data engine to ensure that policies in the policy nanger are operational on the data. Policies can be either specific (a device or user) or semantic (e.g., "do not video record me") in nature.

TIPPERS users have varying degrees of policy specification rights. Infrastructure managers, with the proper credentialing, can add or remove devices from the system. An example would be the facilities manager turning off the video camera when the building opens for the day. Another example would be an individual office tenant specifying policy on their personal cell phone's WiFi.

Data encryption is throughout the TIPPERS system. From the transmission of data over the network, to the storage of data within the repositories, encryption is specified and enforced as one form of PET.

Another PET integrated into TIPPERS is differential privacy. Occupancy heat -maps can be generated with ε-differential privacy to provide variable privacy yet some (variable) degree of utility. As a matter of policy, utility will increase during times of emergency or other need. Privacy vs. Utility analysis is an ongoing topic of research with the TIPPERS team.

TIPPERS Deployments for COVID-19 Mitigation

The TIPPERS system deployments (see above) are being utilized to mitigate COVID-19. This idea was first deployed and tested by the US Navy as part of their COVID-19 response. The TIPPERS system received a 100% accuracy rating during the COVID-19 mitigation testing by the Navy. As a result, the other institutions realigned TIPPERS for the same purpose. As TIPPERS can be deployed in a passive monitoring fashion (completely privacy preserving) and still provide accurate information, the deployments are being utilized extensively by the respective user communities. Preserving privacy and yet providing utility in information is an ongoing area of research.

Features of the TIPPERS deployments utilized for COVID-19 mitigation include a custom dashboard, privacy aware social distancing monitoring, and occupancy counts. The dashboard is created as the first stop for the user community. The building administrators provide maps, AP locations, and any regions of interest. They are then fed into the TIPPERS system by virtue of the concomitant T-Mapper tool. From that point, a dashboard is created.

From the dashboard, users can scroll through graphs and other time series data (up to and including real-time data) for social distancing adherence. Any particular location or region of interest can be observed If there is an issue with social distancing in a region, it is tagged and highlighted.

The dashboard also provides the authorized user occupancy counts of any region or location. This informs the user of the current occupancy count and the user can decide if they wish to go to that location. For instance, if a lab on the first floor of Building A shows a high level of occupancy, the user can select another lab in which to do their work.

At the universities mentioned above as well as the CalIT2 research institute, TIPPERS is used as a Covid-19 mitigation tool. Users check common spaces such as the libraries or lab spaces to ensure social distancing adherence. On average, the TIPPERS dashboards are receiving several hundred hits per day.

Conclusion

In this paper, privacy issues related to big data analytics were discussed. Most scholars estimate that the world produces exabytes of information on a daily basis. Related to that vast amount of information are privacy concerns.

The authors began the paper with a literature review providing a contextual discussion that covered the nascent concepts of big data and big data analytics. From that point, the idea of privacy from legal definition to current issues was presented. The literature review concluded with a discussion on privacy and security with regard to big data analytics.

In order to address privacy consideration in the area of big data analytics, the authors suggest Cavoukian's (2011) PbD approach. This approach is such that privacy is an integral aspect of systems engineering and is prevalent in all design stages. This serves to put privacy as a design principle and not as an after-thought.

The authors finished this paper with a discussion of TIPPERS in both a military environment as well as a tool for COVID-19 mitigation. In both cases, the privacy-preserving functionality of the system was discussed.

Big data analytics presents a number of privacy concerns. Identity theft to actual physical harm can result from data privacy breaches. These serious issues make privacy research an important topic for scholars to pursue.

References

Altman, M., Wood, A., O'Brien, D.R,, and Gasser, U. (2018). Practical Approaches to Big Data Privacy over Time. *International Data Privacy Law*. doi:10.1093/idpl/ipx027.

Andrea, I., Chrysostomou, C., Hadjichristofi, G. (2015). Internet of Things: Security vulnerabilities and challenges. *IEEE Symposium on Computers and Communication (ISCC)*, Larnaca, 2015, pp. 180-187

Archer, D., August, M.A., Bouloukakis, G., Davison, C. B., Diallo, M. H., Ghosh, D., Graves, C. T., Hay, M., He, X., Laud, P., Lu, S., Machanavajjhala, A., Mehrotra, S., Miklau, G., Pankova, A., Sharma, S., Venkatasubramanian, N., Wang, G., & Yus, R. (2020). Transitioning from testbeds to ships: an experience study in deploying the TIPPERS Internet of Things platform to the US Navy. *The Journal of Defense Modeling and Simulation, 1548512920956383.* https://doi.org/10.1177/1548512920956383

Blankley, K. M. (2004). Are public records too public-why personally identifying information should be removed from both online and print versions of court documents. *Ohio St. LJ, 65*, 413.

Bu, F., Wang, N., Jiang, B., & Liang, H. (2020). "Privacy by Design" Implementation: Information system engineers' perspective. *International Journal of Information Management,* 53, 102124.

Bycer, M. (2014). Understanding the 1890 Warren and Brandeis "The Right to Privacy" Article. Nationalparalegal.edu; National Juris University. https://nationalparalegal.edu/UnderstandingWarrenBrandeis.aspx

Castiglione, A., Choo, R. K., Nappi, M., & Ricciardi, S. (2017). Context Aware Ubiquitous Biometrics in Edge of Military Things. *IEEE Cloud Computing, 4*(6), 16-20.

Cavoukian, A. (2011). Privacy by design: origins, meaning, and prospects for assuring privacy and trust in the information era. In *Privacy protection measures and technologies in business organizations: aspects and standards* (pp. 170-208). IGI Global.

Davison, C., Hua, D., Sheridan, B., Shimer, Z., & Bowles, B. (2020). Privacy presentation and data legacy. *International Journal of Latest Research in Engineering and Technology*, *6*(2), 34-39.

Dinur, I., & Nissim, K. (2003, June). Revealing information while preserving privacy. In *Proceedings of the twenty-second ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems* (pp. 202-210).

Dwork, C., McSherry, F., Nissim, K., & Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In *Proceedings of the Third conference on Theory of Cryptography (TCC'06)*, Shai Halevi and Tal Rabin (Eds.). Springer-Verlag, Berlin, Heidelberg, 265–284.

Favaretto, M., De Clercq, E., Schneble, C. O., & Elger, B. S. (2020). What is your definition of Big Data? Researchers' understanding of the phenomenon of the decade. *PloS one*, *15*(2), e0228987. https://doi.org/10.1371/journal.pone.0228987

Gantz, J., & Reinsel, D. (2012). The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east. *IDC iView: IDC Analyze the future, 2007* (2012), 1-16.

Klievink, B., Romijn, BJ., Cunningham, S. *et al.* Big data in the public sector: Uncertainties and readiness. *Inf Syst Front* 19, 267–283 (2017). https://doi.org/10.1007/s10796-016-9686-2

Laud, P., Pankova, A., & Pettai, M. (2020). A Framework of Metrics for Differential Privacy from Local Sensitivity. *Proceedings on Privacy Enhancing Technologies, 2,* 175-208.

Li, C., & Palanisamy, B. (2019). Privacy in the Internet of Things: From Principles to Technologies. *IEEE Internet of Things Journal*, 6(1), 488-505.

Mackie, J. (2018). "Privacy by Design." *TermsFeed*. www.termsfeed.com/blog/privacy-design/.

Nelson, K. A. (2011, December). *"The Right to Privacy" by Warren and Brandeis*. Inpropriapersona.com. https://inpropriapersona.com/articles/the-right-to-privacy-by-warren-and-brandeis/

NGDATA | What is Big Data Analytics? Learn About Tools and Trends. (2016, February 9). NGDATA. https://www.ngdata.com/what-is-big-data-analytics/

NortonOnline. (n.d.). *What Are Some of the Laws Regarding Internet and Data Security?*. Retrieved from https://us.norton.com/internetsecurity-privacy-laws-regarding-internet-data-security.html

Right of Privacy. (2013). TheFreeDictionary.com. https://legal-dictionary.thefreedictionary.com/right+of+privacy

Southworth, C., Finn, J., Dawson, S., Fraser, C., & Tucker, S. (2007). Intimate partner violence, technology, and stalking. *Violence against women, 13*(8), 842-856.

Wagh, S., He, X., Machanavajjhala, A., & Mittal, P. (2020). DP-Cryptography: Marrying Differential Privacy and Cryptography in Emerging Applications. arXiv preprint arXiv:2004.08887.

Walker, K. (2014, July 16). *The legal considerations of the Internet of Things.* https://www.computerweekly.com/opinion/The-legal-considerations-of-the-internet-of-things

Warren, S. D., Brandeis, L. D. (1890). The Right to Privacy. *Harvard Law Review, 4*(5), 193-220.

What Is Big Data? | University of Wisconsin. (2019, October 7). University of Wisconsin Data Science Degree. https://datasciencedegree.wisconsin.edu/data-science/what-is-big-data/

Yang, Y., Lutes, J., Li, F., Luo, B., & Liu, P. (2012, February). Stalking online: on user privacy in social networks. In *Proceedings of the second ACM conference on Data and Application Security and Privacy* (pp. 37-48).