10-8-2018

Exploratory Strategic Roadmapping Framework for Big Data Privacy Issues

Maoloud Dabab
Portland State University

Rebecca Craven
Portland State University

Husam Barham
Portland State University

Elizabeth Gibson
University of Colorado Boulder

Citation Details

M. Dabab, R. Craven, H. Barham and E. Gibson, "Exploratory Strategic Roadmapping Framework for Big Data Privacy Issues," 2018 Portland International Conference on Management of Engineering and Technology (PICMET), Honolulu, HI, 2018, pp. 1-9.

Exploratory Strategic Roadmapping Framework for Big Data Privacy Issues

Maoloud Dabab [1], Rebecca Craven [2], Husam Barham [1], Elizabeth Gibson [3]
[1] Dept. of Engineering and Technology Management, Portland State University, Portland, OR - USA
[2] Dept. of Political Science, Portland State University, Portland, OR – USA
[3] Engineering Management Program, University of Colorado Boulder, Boulder, CO – USA

***Abstract*—The applications of Big Data continue to expand, due to the many possibilities and unprecedented insights it offers to people, organizations, and communities. However, Big Data poses serious challenges as well, including challenges to the privacy and security of individuals and their data. This paper considers how to best address one concern related to Big Data: the social problems that the pervasiveness of data collection, analysis, and storage create with regard to individuals' ability to control their own data. The paper uses Quality Function Deployment (QFD) and Technology Roadmapping analysis methods to assess the social problems, technologies, resources, and industries that are most relevant to data privacy, and what should be done to address it. The findings indicated that the healthcare industry is one of the most important industries to consider concerning data privacy because of the nature of the data generated through medical processes and technologies. Furthermore, it was found that enforcement mechanisms, specifically in the form of federal enforcement agencies, are the most effective approach to ensure compliance by actors. It was also realized that there are extenuating political circumstances and increased costs that make the implementation of those policies challenging in the United States.**

I. INTRODUCTION

We live in the information age, where the advances in information and communication technologies in the last few decades, and especially in the previous decade, has spurred the generation and use of unprecedented amount of data about almost everything surrounding our modern life [28, 50]. In response to this phenomenon, Big Data, a field in information technology, emerged as a viable way to handle and make use of this influx of data.

Big Data does not have a single agreed-upon definition in the literature but is most often characterized, as an entity, by its volume, velocity, variety, veracity, and value. Other definitions addressed Big Data from a process perspective, referring to Big Data as being a holistic information management approach, to acquire, clean, integrate, store, and analyze data that comes from multiple internal and external sources, that can be structured or unstructured, to generate insights and analytics to support decision making [10, 23, 24, 25, 26, 27]. Unlike analog data, this native digital data allows faster access, processing, and analysis in greater quantities than previously possible. Big Data depends on four main steps of data processing: collection, storage, analysis, and usage. The technologies used in these steps are no longer novel and emergent; Gartner has not included Big Data from its annual hype cycle report since 2015 because Big Data has "gone mainstream" [11].

Big Data technologies depend on continuous streams of data, generated by individuals that may or may not be aware of their data's place in datasets or analysts' hands, which lead to raising concerns over the privacy and security of individuals' data. Such concerns echo those of previous eras of non-digital data, which are at this point well legislated and litigated in the American system [1]. As such, the PCAST report on Big Data and privacy notes that "…it is the use of data (including born-digital or born-analog data and the products of data fusion and analysis) that is the locus where consequences are produced" (xii).

The aim of this paper is to address the privacy concerns that arise from the increasing ubiquity of Big Data. First, the paper provides an overview of literature related to the policy background and social problems related to privacy and Big Data. It then proceeds by identifying the most critical industries where Big Data privacy concerns have a serious impact. Then, potential solutions to these concerns are analyzed by using Quality Function Deployment (QFD) and Technology Roadmapping analysis methods. Finally, recommendations are offered.

II. BACKGROUND AND RELATED WORK

Big Data can be used to help solve a range of complex problems [39]. However, using Big Data constitutes social challenges as well. Some of the main issues and challenges imposed by Big Data, as indicated by literature, are related to privacy and security, data access and sharing, storage and processing, and management and technical issues [37, 36, 41].

A. Big Data and Privacy

As Big Data increasingly becomes part of every industry and every sector, it makes new solutions to a plethora of challenges possible. Examples include better addressing customers' needs, more accurate human behavior analytics, more effective medical treatment, improving food security, and preventing human trafficking, just to name few [22, 55, 56]. However, it also brings new social challenges with it. These social problems can be grouped into five general categories: privacy and security, data reuse, data accuracy, data access,

978-1-890843-37-3 ©2018 PICMET

and archiving and preservation [1]. Each of these social problem areas represents a point in the Big Data process at which social externalities can occur, and are discussed in both the PCAST report and in the broader Big Data literature. Privacy and security refer primarily to the initial generation of data by, or about, individuals, including secondary generation through association with other existing data sets. Data reuse, by contrast, is concerned with the repurposing of data from its intended recipients and processes for other uses. Data accuracy issues can arise when multiple data sources with differing controls and verification processes influence the overall quality of the data, and the degree to which the data is correct. Data access concerns the individuals and organizations that have access to any data that is part of the Big Data process, including the archiving and preservation of data, which refers to the historical cataloging of data once its initial use has passed. In all of these cases, individuals that generate data have little to no control over that data once it comes into digital existence unless technology processes are constrained by social and political processes [1].

In the literature, concerns about Big Data privacy have been specifically singled out for additional consideration by industry, government, and research consortia [2, 29, 30]. Many aspects of technology have been party to the challenges to the bounds of the American right to privacy [1]. However, others have noted that information technology poses privacy issues that are unique to digital-native data [9], and that Big Data is particularly pertinent to technology privacy discussions [8].

Because Big Data has implications for both public and private uses, there are especially large implications for government if policies are not sufficient to address privacy issues with Big Data [7], especially in light of the historical precedent favoring the individual's right to privacy in the United States. Thus, while many aspects of Big Data pose potential risks and social problems, the issues concerning data privacy and the protection of individual data generators are the most important, and hence, will be the center of focus for the remainder of this report.

a) Data Privacy

According to Westin, information privacy is the ability of the individual to control the terms under which personal information is acquired and used. Information here refers to information identifiable to a person. In this context, data and information refers to the same thing, and this definition can be used for Big Data privacy as well [32, 33]. Another popular definition by the Generally Accepted Privacy Principles (GAPP) standard is "the rights and obligations of individuals and organizations with respect to the collection, use, retention, disclosure, and disposal of personal information" [34].

Data privacy concerns revolve around individuals losing the ability to control information related to them, by either unauthorized access to the data due to security breaches, or using the data for purposes, the individual is not aware of, when initially providing consent. Regulations protecting privacy govern what can be revealed under different circumstances. However, such regulations vary by country and region. For example, in the US, health records have strict regulations about how they can be used and shared [30, 31, 33]. Implementing privacy within Big Data is not a straightforward task; Big Data depends on sharing and consuming information from multiple internal and external sources to generate real value [25, 35]. This sharing of information among entities exposes data to increased risks of unauthorized access and unauthorized use. Furthermore, there is the possibility of identifying individuals even if their identity is removed. For example, an attacker, with access to data representing traffic routes by individuals can analyze patterns of traffic routes to identify home and workplace of individuals and then cross-check that with targeted individuals, and then use that information in harmful or illegal ways, even if the source data is anonymized [33].

b) Data Privacy in Different Industries

Each industry has unique Big Data privacy concerns due to how Big Data is utilized in that sector. Following is a review of the main Big Data privacy concerns in several industries where Big Data is being widely adopted, including healthcare, telecommunication, retail, education, and utility.

- Healthcare: Patient medical records can be used by Big Data to offer tremendous insights, allowing for better diagnosis and treatment decisions. However, failing to properly address patients' privacy, could have serious impact on patients, related to their jobs, health insurance, and social life [53, 54].
- Telecommunication: Big Data is used to capture and analyze customers' communication and mobility data. Such Big Data systems help telecommunication companies to offer better customer service and to better target customers from marketing and sales points of view. However, failing to properly address customers' privacy could result in breaches that expose business and personal secrets and plans, leading to economic losses and personal embarrassments [38, 40].
- Retail: There is a tremendous growth in the amount of retail data being generated. Big Data offers valuable insights to retailers, like the ability to better target customer needs, as well as engage in more efficient supply chain, operations, and inventory management. However, failing to properly address customers' privacy could result in exposing customers' information and habits, which might have impact on their jobs and personal life's [42].
- Education: Big Data have several educational applications, including, offering better insights that can help in enhancing students' performance, and offering students with educational methods that are customized to their individual skills. Furthermore, Big Data can play an important role in addressing the higher education retention phenomenon. However, failing to properly address students' educational

records privacy could have serious impact on their future jobs and future graduate studies [57, 58].

- Utility: Big Data is used to enable smart grid initiatives by collecting and analyzing tremendous amounts of data about how power is being used by customers. Smart grid initiatives have many applications. For example, smart grid can enable better forecasting of power demands and how to efficiently respond to it. Also, smart grid allows for better use of renewable intermittent resources. However, like with other industries, failing to properly address customers records privacy could result in exposing personal information and make customers vulnerable for identity theft and other kinds of data attacks [58, 59, 60].

c) Examples of Big Data Privacy Cases

There are many known cases were Big Data resulted in legal or ethical violations of privacy that had severe impacts. Two well-known cases illustrate how Big-Data-related privacy breaches can have serious impact. Considered first is the Equifax case, which represents a data cybersecurity breach; then, the Target case that represents unauthorized use of data [51, 52]: Equifax is one of the biggest credit bureaus in the United States. One of Equifax's websites was hacked in 2017, resulting in the leak of more than a hundred million consumers' data that includes their social security numbers, birth dates, and addresses, among other kinds of sensitive data, which can be used for identity theft. Another case is related to Target, the giant retailer. In 2014, Target implemented a Big Data system that can analyze purchasing patterns by customers to make predictions about them. Because of this system, a man received pregnancy related promotions addressed to his teenage daughter, who had not yet disclosed her pregnancy. The system invaded the teenager's privacy as it used her purchase patterns without her consent to reach conclusions about her health, and exposed this information to her father without her approval.

B. Policy Background

The full policy background on Big Data in the United States includes aspects of cybersecurity and privacy policies that are not necessarily specific to Big Data. Cybersecurity-specific policies are particularly well documented in the literature [5]. Additionally, much of the analysis of Big Data is from a legal, and not technical, perspective [3]. As such, the policy infrastructure primarily refers to existing privacy protection policies that are in place, which may not deal with digital privacy at all; these policies also operate under outdated assumptions of data isolation to specific contexts and fields [63]. In general, ingrained policies and protections are neither Big Data nor even technology specific, and do not reflect the evolving policy concerns that accompany each advance in information technology's abilities to collect and analyze data [64]. In the United States, policies regarding privacy and Big Data can be implemented at two levels: federal and state. Treatments of US Big Data policy generally show a lack of coherent federal policy that crosses sectors [6] [3], with most policies specifically focusing on particular sectors like healthcare, education, and financial institutions. Data policies are contained within broader privacy legislation like HIPAA, FERPA, and GLBA/FCRA for those sectors respectively. There are some precedents for state-level policies regarding data protection and privacy issues in the United States, though the scope of these laws and the strength of enforcement mechanisms vary greatly [65]. California is especially active in this regard, with statewide, cross-sector legislation like the California Online Privacy Protection Act (CalOPPA) addressing digital data privacy [12]. However, this type of legislation remains uncommon. This process of distinguishing between states has the effect of further fragmenting data privacy policies and their reach. Just as there are no comprehensive Big Data regulations at the federal level, there are also no bodies tasked specifically with Big Data compliance, monitoring, or enforcement. Legislation concerning data privacy is full of recommendations and self-enforcement requirements, but little in the way of coercive inducements to follow the guidelines or coordination to ensure equal adherence to those recommendations and requirements [66]. Given the acknowledged tension between corporate and consumer interests regarding data collection and usage [67], self-enforcement seems to protect corporate big data use over consumer privacy by default. Therefore, this policy regime also fails to explicitly consider the privacy concerns of individuals. This is not merely a corporate problem; there are also federal agencies that are engaged in Big Data research in cooperation with private industry (OSTP and NITRD are particularly important in this respect), though the research is actually carried out by legislatively-created agencies that have much broader missions (NASA, EPA, NOAA, etc. all fit this characterization). Thus, there are a multitude of agencies and laws that impact data privacy, though the effects are inconsistent across jurisdictions and industries.

Some countries and regions are ahead of the United States in terms of the data privacy policies that they have either implemented or are in the process of implementing. Greenleaf [4] offers a comparative look at privacy laws, though most are not actually Big Data specific. Perhaps the best example of data privacy policies from which the US might learn is the European Union, which has had a binding, cross-sectoral Directive on the Protection of Personal Data since 1995 [68]. The new General Data Protection Regulation (GDPR), which takes effect in May 2018, represents further industry-spanning comprehensive Big Data policies that affect all data within the geographical territory and clarify existing conceptual understandings around matters like "consent" that would affect anyone generating, transmitting, or storing data in the EU [69]. Other countries are also considering implementing similar laws, which indicates that these types of policies might be possible to consider in the US context.

The gaps in the current US policies and mechanisms regarding Big Data privacy are identified in the PCAST report and what is implemented elsewhere. The US has only focused on industry-specific privacy and data protection legislation and regulation (HIPAA, FERPA, etc.) and not the comprehensive approaches like those in the EU. Given the recommendations

of the PCAST report [1], especially recommendations 2 and 5, it is apparent that in order to meet increasing social needs regarding technology in the United States, a comprehensive policy is needed to address privacy issues that exist across jurisdictions and sectors.

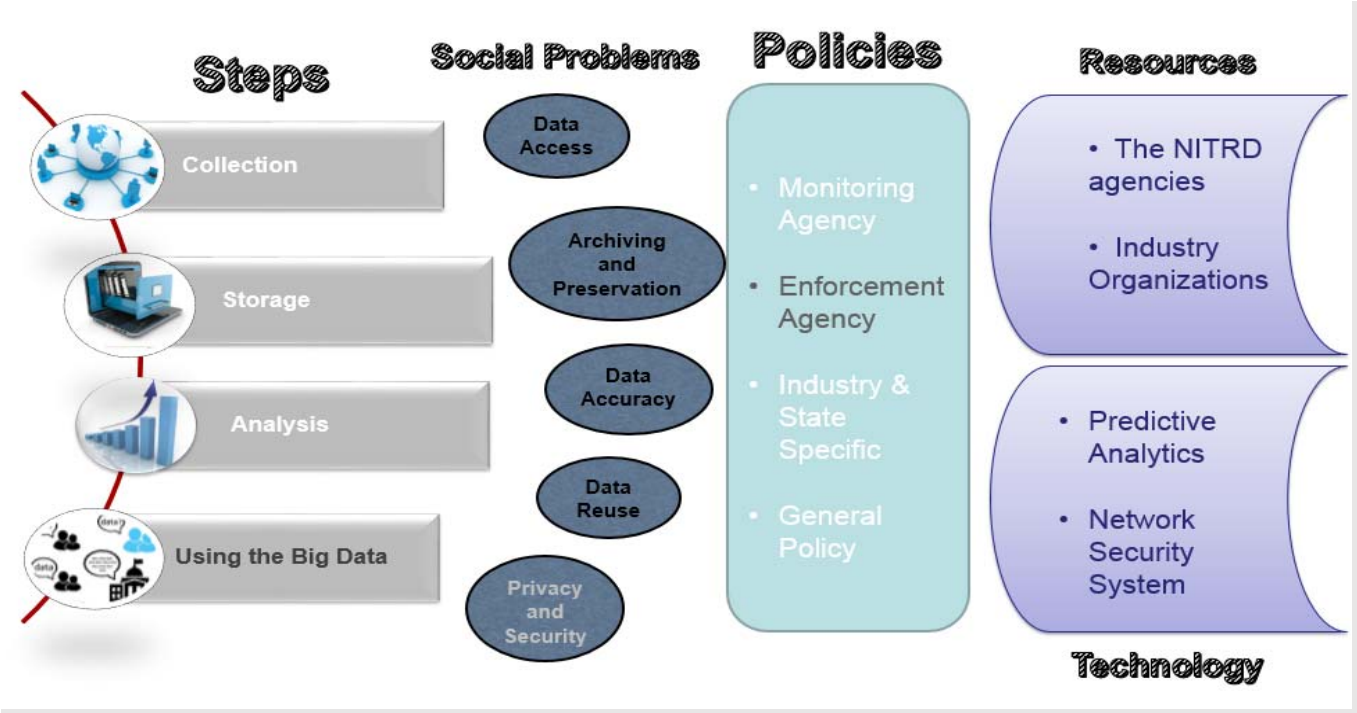
III. METHODOLOGY

To analyze the policy solutions that could be applied to the social problems created by Big Data, this paper utilized the Technology Roadmapping methodology (TRM), taking the multi-organizational approach as outlined in Phaal et al. [14]. Technology Roadmapping is a comprehensive approach for strategy planning to integrate science/technological considerations into business planning, and provide a way to identify new opportunities to achieve a desired objective from the development of new technologies [14, 45, 46].

Moreover, Technology Roadmapping plans has multiple related layers. Roadmaps are used to identify what need to be achieved, the barriers and shortcomings preventing its achievement (the gap), and what needs to be done to overcome those barriers and shortcomings. A technology roadmap of social problems and technologies can provide insights for policy makers, industry leaders, and private citizens regarding the social challenges related to a certain technology, what type of actions needed to address those challenges and what is missing, or need further development, to address in act those actions. This paper used Technology Roadmapping method to properly analyze and find out: What are the gaps in proper addressing of the privacy issues related to big data. In addition, what is needed to address them?

Furthermore, to provide a prioritised list that is most related to help build a relevant technology roadmap, the paper utilized Quality Function Deployment (QFD) method. QFD is defined as "a systematic way of ensuring that the development of product features, characteristics, and specifications, as well as the selection and development of process equipment, methods, and controls, are driven by the demands of the customer or marketplace" [48, p. 2]. This method was originally developed by Akao at the Tokyo Institute of Technology in the late 1960s. Under this method, a house of quality (HoQ) is built, which consists of a matrix with alternatives presented in one dimension and the desired characteristics in another dimension. The cells of the matrix in the QFD method represent the weight of each alternative in achieving each desired characteristic in by experts. QFD is effective in transferring the qualitative judgment of experts into quantitative parameters. In this paper, using HoQ, experts determined how well each alternative meets the requirements on each of four levels. QFD is used at each level of the Roadmap, starting from policies against social problems, then Industry fields against each policy, then resources against each industry, and finally, technologies against each resource [16, 17, 49, 70].

The technology roadmap was divided into four phases. As Big Data has vast effects on privacy across all sectors, and the literature varies considerably regarding projections and levels of certainty, it would be proper to have a technology roadmap with several relevant phases. Additionally, as outlined above, the groundwork is already laid for sector-specific approaches to data and privacy issues. Pavolotsky indicated that "… Because no two businesses are the same, if privacy policies are the same or substantially similar, at least one of the privacy policies is not on point" [15]. Therefore, we divide our Roadmapping into four phases, which include social problem identification, policy analysis, industry assessment, technology assessment, and resource allocation. This leads us to identify an initial industry in which the need for policy change is most pertinent, and a prioritized list of technologies and resource



needs to address this industry's needs when it comes to big data privacy.

The cells of the matrix in the QFD method were filled based on the experience of the authors since some of them have long experience with technology policy, and others have strong technical background and experience related to big data. In addition, the information that they gained from the literature review helped them give more accurate ratings. The weights represent how well each alternative addresses the requirements, and the index of weights is illustrated in the figures. Then, total scores for each alternative are calculated to identify the best alternatives. This analysis method was used throughout the four TRM phases to narrow down the aspects of Big Data and the policy environment that are most pertinent to a roadmap.

Therefore, QFD helped in building a more accurate technology roadmap, by identifying the most important alternatives for each phase in the map, and hence shedding light on the industries that are most affected by big data privacy challenges, and what actions should be done to address those challenges.

IV. ANALYSIS

Four phases of QFD analysis were used to generate the findings. The purpose of these phases of analysis is to identify in which industry, big data privacy issues are more pertinent, and what policies are needs to address it. Then to identify resources necessary for policy change to occur. The categories included in each analysis are based on the literature review and authors experience (see above), as well as, the PCAST report [1]. In each stage, between four and seven characteristics were ranked on a scale of strong, medium, and weak, with strong indicating the most pronounced effects based on each pairing.

The analysis tables for all four phases are shown above. In the first phase, policies that include enforcement mechanisms or enforcement agencies were selected as the most effective in addressing privacy and security issues. In the second phase, the medical industry was selected as the industry in which the effects of data policies are most felt. In the third phase, in term of policy resources, industry organizations and NITRD agencies were selected as having the strongest role to play in policymaking that will affect data privacy and security in the medical field. Finally, in the fourth phase, the specific applications of Big Data technology that are require the most consideration in regard of privacy, were identified as Internet of Things, Real Time Analytics, and Predictive Analytics. Literature bolsters these findings, for example, Roski et al [18], identified these technologies as providing the greatest opportunities for advancement, but also the greatest threats to privacy and security for individuals, in the healthcare context.

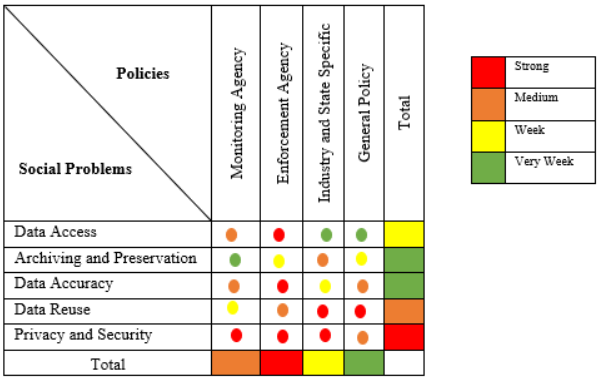


Fig. 2. The first phase

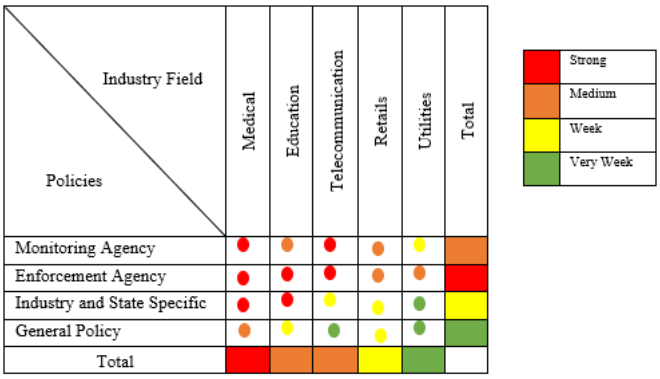


Fig. 3. The second phase

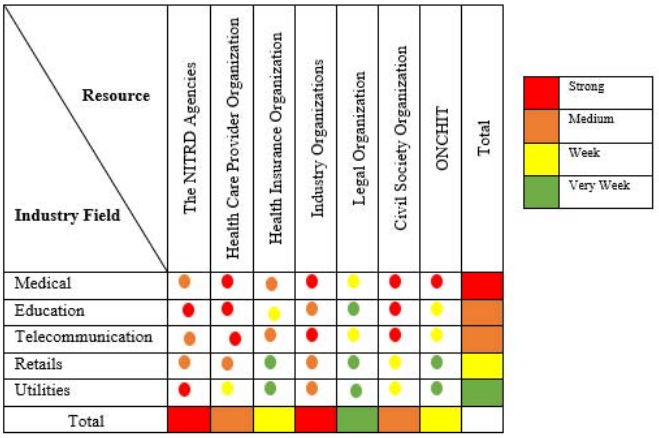


Fig. 4. The thired phase

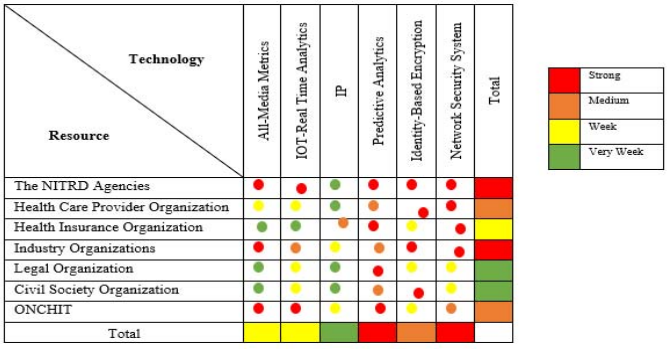


Fig. 5. The fourth phase

The information from the QFD analysis were then compiled into the technology roadmap framework to provide a pictorial representation of the analysis. Using the typology of roadmaps articulated by Phaal et al [14], these are best described as strategic planning roadmaps that use four main elements (social problem, policy, technology, and resources) that were then extended with elements specific to this analysis: addressing big data related privacy issues. Because there is a variety of roadmap types, several depictions of technology roadmaps are included here. Each is informed by the QFD analysis and simply organizes the resulting information slightly differently. These roadmaps provide insight into the feedback loops that exist between technology, industry, and policy as Big Data becomes more ingrained in the processes of the healthcare industry. The steps in the development and adoption of any technology necessarily inform each other, and our analysis of relevant literature revealed the linkages shown in these Roadmapping graphics (see Figure 6 and Figure 7).

V. DISCUSSION

There are several interesting points that arise from this analysis. First, while many fields are affected by Big Data, the healthcare industry poses both one of the biggest opportunities for Big Data to add value, as well as, one of the biggest threats to individual privacy. Personal health data is considered among the most private and protected by strict laws, especially in the USA. The compilation of diverse data via typical processes and advancements like IoT-connected devices through processes that may or may not be evident to individuals generating data, are new challenges to data that do not have an analog counterpart. The use of these data streams for predictive analytics purposes also is a bigger concern given the sensitivity of this data. As the history of privacy in the United States shows, this is an evolving legal process that will continually be challenged as technologies develop and new uses for data are found. Data governance is a relatively new concept in the public sector in particular, this will undoubtedly result in the emergence of additional social problems around the related values, risks, and costs of big data that need to be considered while developing and implementing more Big Data governance policies [19].

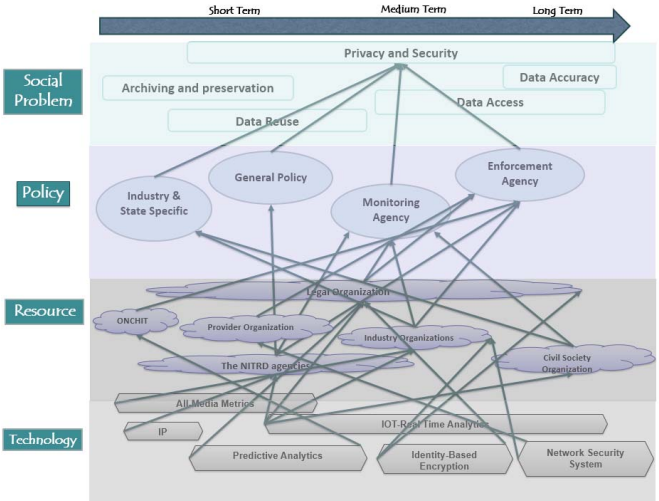


Fig. 6. Technology Roadmapping

The policy tool with the most support for effectively protecting individual privacy and providing data security is an enforcement agency with the ability to impose penalties on corporations, organizations, and industries that fail to adequately protect information. This is similar to the enforcement mechanisms that are in the process of being implemented in the European Union [13]. Indeed, the enforcement agencies play a powerful role in Europe to ensure movement of goods, persons, services, and capital. Beyond EU-level enforcement, some countries also empower their own state agencies to enact and carry out control of domestic enforcement; one example of such an agency is the Sweden National Tax Board, which controls both civil and public matters [43]. At the same time, the European Union has taken several steps to create data protection laws and enforcement agencies for protecting security and achieving justice. An excellent example of this is the European Union Agency for Law Enforcement Cooperation (Europol), which was created in 1998 to coordinate criminal intelligence and combat terrorism. The main goal behind Europol is to obtain a safer environment for the benefit of all the EU citizens through the coordination of information from numerous organizations [44]. This is a major departure from current US policies at all levels, which provide guidelines to industries but ultimately rely on self-reporting and the voluntarily provision of compensation, or civil litigation in the absence of this compliance. Criminalization of negligent data handling, like that proposed in the European Union, provides greater incentives to comply but also requires political capital to facilitate such a major policy shift. Thus, while this policy tool emerged from the QFD analysis as the one with the greatest potential benefit, it also may face political feasibility issues that could derail it in its entirety. Nevertheless, there are examples of enforcement agencies in other sectors in the United States. One such example is the Consumer Financial Protection Bureau (CFPB). In the context of the 2008 economic crisis, the CFPB was established under the Dodd-Frank Act with the objective of protecting consumers' financial interests from "unfair, deceptive, or abusive acts or practices" of financial entities. The CFPB supervises financial entities' practices and has the authority to enforce policies by taking actions against violating financial entities. Over the years, CFPB issued fines totaling hundreds of millions of dollars and returned billions of dollars to harmed consumers; its enforcement prerogative allowed for better consumer protection [61, 62]. However, even this agency remains susceptible to political feasibility issues, as legislation to repeal the Dodd-Frank Act and essentially remove all enforcement mechanisms from the CFPB gains momentum in Congress under the tutelage of the Trump administration.

VI. CONCLUSIONS AND RECOMMENDATIONS

Big Data poses new security and privacy risks as greater quantities of data are generated, processed, and used, often without the knowledge of the individuals creating the data streams. The United States has a complex history with privacy that is only exasperated by the speed at which Big Data technology is adopted in industries that already collect and compile sensitive data. Big Data policies that provide

consistency across industries are desirable, as they create consistent expectations for individuals that are generating the data that is used in datasets and predictive analytics. For this reason, progress in the healthcare sector can be beneficial in other industries as well, as healthcare often provides a blueprint for other sectors seeking to protect individuals' data. The use of Big Data is a primary concern for consumers and those managing data privacy in general, and modernization of privacy policies and data protection measures is needed to ensure individuals that their data is safe. The summary of the background information, QFD analysis, and key aspects of the Technology Roadmapping are shown in the following figure.

There are three primary recommendations that arise from the analysis. First, the new presidential administration in the United States should renew the funding of those agencies that participate in Big Data research and policy formation, including the NITRD in particular. NITRD is an organization that is created through executive order, so it is subject to presidential renewal with each new administration. NITRD also created a 5-year research plan for Big Data, much of which focuses explicitly on the privacy and security of data. Continuing to research the impacts and possible policies associated with Big Data in a collaborative way is imperative to protecting current and future data streams, and ensuring the continued existence and funding of agencies that are committed to this mission seems the best way to do this.

Additional research on Big Data's relationship to data privacy and security should also continue. Future research should incorporate the costs of technologies and policies that can ensure data security, which are acknowledged but not a part of this analysis. The legal and political costs of policy change in this area also warrants further analysis, as even the PCAST report acknowledges that "privacy protection cannot be achieved by technical measures alone" (xii). The complexity of the policy environment surrounding Big Data governance is a factor that makes a single analysis method unlikely to capture the full scope of the opportunities and threats that are facing all actors. We therefore recommend that additional research continue combining methods of analysis to attempt to better understand the extent of the interacting aspects of the ever-changing technology environment. While the opportunities that come from Big Data analytics are well documented, the particular technology and policy methods that can be used to best ensure data security and privacy remain unclear. This analysis shows that enforcement mechanisms in the form of agencies that are capable of leveraging civil and criminal penalties against those that fail to adequately guard the data generated by individuals provide the strongest protections. However, the social, political, and technical complexity of the policy environment continues to increase, and only time will tell if this analysis provides insight that is either possible or feasible to implement in the United States.

REFERENCES

[1] "Report To The President Big Data And Privacy: A Technological Perspective," May-2014. [Online]. Available: https://bigdatawg.nist.gov/pdf/pcast_big_data_and_privacy_-_may_2014.pdf. [Accessed: 03-Mar-2018].

[2] "The Federal Big Data Research and Development Strategic Plan," May 2016. [Online]. Available: https://www.nitrd.gov/pubs/bigdatardstrategicplan.pdf. [Accessed: 03-Mar-2018].

[3] F. Jahanian, "The Policy Infrastructure for Big Data: From Data to Knowledge to Action," *ISJLP*, vol. 10, p. 865, 2014.

[4] G. Greenleaf, "Sheherezade and the 101 data privacy laws: Origins, significance and global trajectories," *JL Inf. & Sci.*, vol. 23, p. 4, 2014.

[5] C. Hart and A. Feenberg, "The Insecurity of Innovation: A Critical Analysis of Cybersecurity in the United States," *International Journal of Communication*, vol. 8, p. 19, 2014.

[6] R. Stough and D. McBride, "Big Data and US Public Policy," *Review of Policy Research*, vol. 31, no. 4, pp. 339–342, 2014.

[7] A. Williamson, "Big Data and the Implications for Government," *Legal Information Management*, vol. 14, no. 4, pp. 253–257, 2014.

[8] J. Van Dijck, "Datafication, Dataism and Dataveillance: Big Data Between Scientific Paradigm and Ideology," *Surveillance & Society*, vol. 12, no. 2, p. 197, 2014.

[9] H. Nissenbaum, "Toward an approach to privacy in public: Challenges of information technology," *Ethics & Behavior*, vol. 7, no. 3, pp. 207–219, 1997.

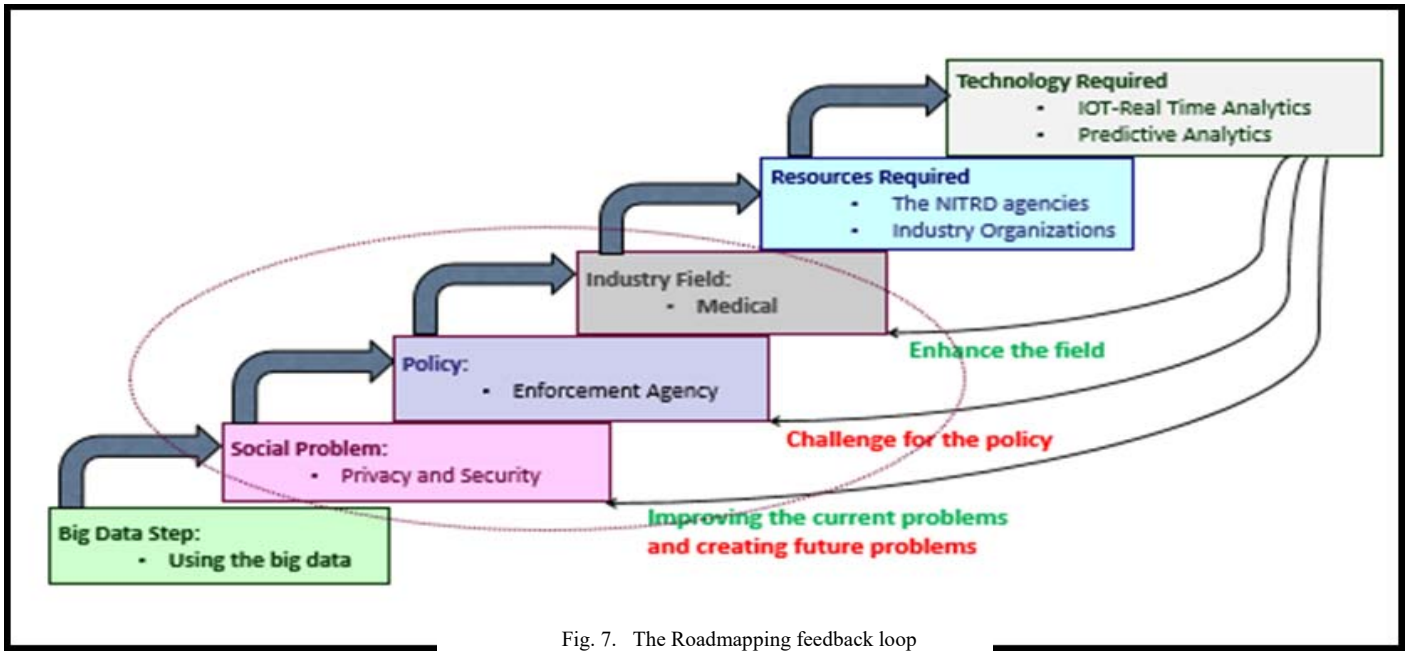


Fig. 7. The Roadmapping feedback loop

[10] N. Kshetri, "Big data′ s impact on privacy, security and consumer welfare," *Telecommunications Policy*, vol. 38, no. 11, pp. 1134–1145, 2014.

[11] N. D. Kho, "The State of Big Data 2018," *EContent Magazine*, 2018. [Online]. Available: http://www.econtentmag.com/Articles/Editorial/Feature/The-State-of-Big-Data-2018-122572.htm.[Accessed: 01-Mar-2018].

[12] "California Online Privacy Protection Act (CalOPPA)," July 2015. [Online]. Available: https://consumercal.org/about-cfc/cfc-education-foundation/california-online-privacy-protection-act-caloppa-3/. [Accessed: 03-Mar-2018].

[13] A. H. Raymond, "Data management regulation: Your company needs an up-to-date data/information management policy," *Business Horizons*, vol. 56, no. 4, pp. 513–520, 2013.

[14] R. Phaal, C. J. Farrukh, and D. R. Probert, "Technology Roadmapping—a Planning Framework for Evolution and Revolution," *Technological forecasting and social change*, vol. 71, no. 1–2, pp. 5–26, 2004.

[15] J. Pavolotsky, "Privacy in the Age of Big Data," *The Business Lawyer*, vol. 69, no. 1, pp. 217–225, 2013.

[16] C. Adiano and A. V. Roth, "Beyond the House of Quality: Dynamic QFD," *Benchmarking for Quality Management & Technology*, vol. 1, no. 1, pp. 25–37, 1994.

[17] L.-K. Chan and M.-L. Wu, "Quality Function Deployment: A Literature Review," *European journal of operational research*, vol. 143, no. 3, pp. 463–497, 2002.

[18] J. Roski, G. W. Bo-Linn, and T. A. Andrews, "Creating Value in Health Care through Big Data: Opportunities and Policy Implications," *Health affairs*, vol. 33, no. 7, pp. 1115–1122, 2014.

[19] P. P. Tallon, "Corporate Governance of Big Data: Perspectives on Value, Risk, and Cost," *Computer*, vol. 46, no. 6, pp. 32–38, 2013.

[20] S. Kaisler, F. Armour, J. A. Espinosa, and W. Money, "Big Data: Issues and Challenges Moving Forward," presented at the System sciences (HICSS), 2013 46th Hawaii international conference on, 2013, pp. 995–1004.

[21] M. D. Assunção, R. N. Calheiros, S. Bianchi, M. A. Netto, and R. Buyya, "Big Data Computing and Clouds: Trends and Future Directions," *Journal of Parallel and Distributed Computing*, vol. 79, pp. 3–15, 2015.

[22] K. C. Desouza and K. L. Smith, "Big Data for Social Innovation," *Stanf Soc Innov Rev*, vol. 2014, pp. 39–43, 2014.

[23] D. Laney, "3D data management: Controlling data volume, velocity and variety," META Group Research Note, vol. 6, p. 70, 2001.

[24] R. Bean, "Just Using Big Data Isn't Enough Anymore," Harvard Business Review, Feb. 2016.

[25] C.-W. Tsai, C.-F. Lai, H.-C. Chao, and A. V. Vasilakos, "Big Data Analytics," in Big Data Technologies and Applications, Cham: Springer International Publishing, 2016, pp. 13–52.

[26] H. Barham, "Achieving Competitive Advantage Through Big Data: A Literature Review," 2017.

[27] A. Shaygan, D. O. Gungor, H. Kutgun, and A. Daneshi, "Adoption Criteria Evaluation of Activity Tracking Wristbands for University Students," 2017, pp. 1–7.

[28] A. Acquisti, L. Brandimarte, and G. Loewenstein, "Privacy and Human Behavior in the Age of Information," Science, vol. 347, no. 6221, pp. 509–514, Jan. 2015.

[29] E. Bertino and E. Ferrari, "Big Data Security and Privacy," in A Comprehensive Guide Through the Italian Database Research Over the Last 25 Years, vol. 31, S. Flesca, S. Greco, E. Masciari, and D. Saccà, Eds. Cham: Springer International Publishing, 2018, pp. 425–439.

[30] D. Nunan and M. Di Domenico, "Big Data: a Normal Accident Waiting to Happen?," Journal of Business Ethics, vol. 145, no. 3, pp. 481–491, 2017.

[31] H. V. Jagadish et al., "Big data and its technical challenges," Communications of the ACM, vol. 57, no. 7, pp. 86–94, Jul. 2014.

[32] A. F. Westin, "Privacy and freedom," Washington and Lee Law Review, vol. 25, no. 1, p. 166, 1968.

[33] M. J. Culnan and P. K. Armstrong, "Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation," Organization Science, vol. 10, no. 1, pp. 104–115, Feb. 1999.

[34] D. Chen and H. Zhao, "Data Security and Privacy Protection Issues in Cloud Computing," 2012, pp. 647–651.

[35] U. Sivarajah, M. M. Kamal, Z. Irani, and V. Weerakkody, "Critical analysis of Big Data challenges and analytical methods," Journal of Business Research, vol. 70, pp. 263–286, Jan. 2017.

[36] S. Kaisler, F. Armour, J. A. Espinosa, and W. Money, "Big Data: Issues and Challenges Moving Forward," 2013 46th Hawaii International Conference on System Sciences, 2013.

[37] A. Katal, M. Wazid, and R. H. Goudar, "Big data: Issues, challenges, tools and Good practices," 2013 Sixth International Conference on Contemporary Computing (IC3), 2013.

[38] J. Bughin, "Reaping the benefits of big data in telecom," Journal of Big Data, vol. 3, no. 1, 2016

[39] L. Patterson, "What Social Issues Will Big Data Solve in 2017?," https://www.technology. org/2017/07/04/what-social-issues-will-big-data-solve-in-2017/, 04-Jul-2017.

[40] N. Kshetri, "Big data′s impact on privacy, security and consumer welfare," Telecommunications Policy, vol. 38, no. 11, pp. 1134–1145, 2014.

[41] S. Sagiroglu and D. Sinanc, "Big data: A review," 2013 International Conference on Collaboration Technologies and Systems (CTS), 2013.

[42] B. Nelson and T. Olovsson, "Security and privacy for big data: A systematic literature review," 2016 IEEE International Conference on Big Data (Big Data), 2016.

[43] W. A. Kennett, Enforcement of judgments in Europe. Oxford: Oxford University Press, 2005.

[44] C. Kaunert, Léonard Sarah, and J. D. Occhipinti, Justice and home affairs agencies in the European Union. London: Routledge, 2015.

[45] M. Amer and T. U. Daim, "Application of Technology Roadmaps for Renewable Energy Sector," Technological Forecasting and Social Change, vol. 77, pp. 1355–1370, Oct. 2010.

[46] K. Vishnevskiy, O. Karasev, and D. Meissner, "Integrated Roadmaps for Strategic Management and Planning," Technological Forecasting and Social Change, vol. 110, pp. 153–166, Sep. 2016.

[47] J. V. Hillegas-Elting, T. Oliver, J. Binus, T. Daim, J. Estep, and J. Kim, "Opening the Door to Breakthroughs that Address Strategic Organizational Needs: Applying Technology Roadmapping Tools and Techniques at an Electric Utility," in PICMET '2015, 2015, pp. 2564–2573.

[48] W. E. Eureka and N. E. Ryan, The customer-driven company: managerial perspective on quality function deployment, 2nd ed. Dearborn, Mich. : Burr Ridge, Ill: ASI Press ; Irwin, 1994.

[49] S. Mizuno, Y. Akao, and K. Ishihara, Eds., QFD, the customer-driven approach to quality planning and deployment. Tokyo, Japan: Asian Productivity Organization, 1994.

[50] D. R. Kinder, "Communication and politics in the age of information.," in Oxford handbook of political psychology, New York: Oxford University Press, 2003, pp. 357–393.

[51] T. Reddy, "7 Big Data Blunders You're Thankful Your Company Didn't Make," Umbel, 22-Oct-2014. [Online]. Available: https://www.umbel.com/blog/big-data/7-big-data-blunders/. [Accessed: 10-Jan-2018].

[52] T. Armerding, "The 16 biggest data breaches of the 21st century," CSO Online, 11-Oct-2017. [Online]. Available: https://www.csoonline.com/article/2130877/data-breach/the-16-biggest-data-breaches-of-the-21st-century.html. [Accessed: 17-Jan-2018].

[53] A. Shaygan, "Landscape Analysis: What Are the Forefronts of Change in the US Hospitals?," in Infrastructure and Technology Management, T. U. Daim, L. Chan, and J. Estep, Eds. Cham: Springer International Publishing, 2018, pp. 213–243.

[54] T. B. Murdoch and A. S. Detsky, "The Inevitable Application of Big Data to Health Care," JAMA, vol. 309, no. 13, p. 1351, Apr. 2013.

[55] S. LaValle, E. Lesser, R. Shockley, M. S. Hopkins, and N. Kruschwitz, "Big data, analytics and the path from insights to value," MIT sloan management review, vol. 52, no. 2, p. 21, 2011.

[56] H. Chen, R. H. L. Chiang, and V. C. Storey, “Business intelligence and analytics: From big data to big impact,” MIS Quarterly: Management Information Systems, vol. 36, no. 4, pp. 1165–1188, 2012.

[57] D. M. West, “Big data for education: Data mining, data analytics, and web dashboards,” Governance studies at Brookings, vol. 4, pp. 1–0, 2012.

[58] M. Savin-Baden, “Education and Big Data,” in Encyclopedia of Educational Philosophy and Theory, M. Peters, Ed. Singapore: Springer Singapore, 2015, pp. 1–7.

[59] N. Chaichi, J. Lavoie, S. Zarrin, R. Khalifa, and F. Sie, “A comprehensive assessment of cloud computing for smart grid applications: A multi-perspectives framework,” presented at the Portland International Conference on Management of Engineering and Technology (PICMET), Portland, USA, 2015, pp. 2541–2547.

[60] Y. Song, G. Zhou, and Y. Zhu, “Present status and challenges of big data processing in smart grid,” Power System Technology, vol. 37, no. 4, pp. 927–935, 2013.

[61] A. M. Townsend, Smart cities: Big data, civic hackers, and the quest for a new utopia. WW Norton & Company, 2013.

[62] CFPB.gov, “Consumer Financial Protection Bureau official website,” Consumer Financial Protection Bureau. [Online]. Available: https://www.consumerfinance.gov/. [Accessed: 22-Jan-2018].

[63] E. Horvitz and D. Mulligan, “Data, Privacy, and the Greater Good,” *Science*, vol. 349, no. 6245, pp. 253–255, 2015.

[64] H. Jeff Smith, Tamara Dinev & Heng Xu. (2011). "Information Privacy Research: An Interdisciplinary Review." MIS Quarterly, 35(4): 989-1015.

[65] F. Spidalieri, “State of the States on Cybersecurity,” *Pell Center for International Relations. Google Scholar*, 2015.

[66] A. L. Newman and D. Bach, “Self-Regulatory Trajectories in the Shadow of Public Power: Resolving Digital Dilemmas in Europe and the United States,” *Governance*, vol. 17, no. 3, pp. 387–413, 2004.

[67] M. J. Culnan and R. J. Bies, “Consumer privacy: Balancing economic and justice considerations,” *Journal of social issues*, vol. 59, no. 2, pp. 323–342, 2003.

[68] N. E. Bowie and K. Jamal, “Privacy Rights on the Internet: Self-regulation or Government Regulation?,” *Business Ethics Quarterly*, vol. 16, no. 3, pp. 323–342, 2006.

[69] P. Elias, “A European perspective on research and big data analysis,” *Privacy, Big Data, and the Public Good: Frameworks for Engagement*, vol. 1, pp. 173–191, 2014.

[70] N. J. Sheikh, “Developing a Strategic Roadmap for Policy and Decision Making: Case Study of ICT and Disaster Risk Reduction in Public Safety Networks,” 2017, pp. 1–7.