



Analyzing Iranian Public Sector Big Data System Requirements Based on System Design Thinking

Soheil Paydar Fard^{a*}, Ali Rajabzadeh Ghatari^a, Mahmoud Dehghan Nayeri^a

^a Department of Industrial Management, Faculty of Management and Economics, Tarbiat Modares University, Tehran, Iran.

How to cite this article

Paydar Fard, S., Rajabzadeh Ghatari, A., Dehghan Nayeri, M., 2024. Analyzing Iranian Public Sector Big Data System Requirements Based on System Design Thinking, *Journal of Systems Thinking in Practice*, 3(3), pp. 1-17. doi: 10.22067/jstinp.2024.84259.1093.
URL: https://jstinp.um.ac.ir/article_45297.html

ABSTRACT

The contemporary world is marked by generation and consumption of vast volume, high velocity, and considerable diverse data, leading us to the concept of big data. In this study, a system design thinking approach was employed to identify the requirements of Iran's public sector big data system. National big data systems would help governments to support their decisions by data and answer to national problems faster. Given the complexity and time-intensive nature of traditional system requirement analysis methods, their practical application in the industry has been declined. Therefore, in this research, system design thinking as an agile alternative for identifying system requirements has been discussed. To accomplish this, the LDA machine learning method has been utilized to analyze approximately 88,000 articles, a thematic analysis on around 600 Instagram and Twitter posts has been conducted, and six experts representing targeted problem persona were interviewed. The objective of this research is to extract insights to serve as a foundation for formulating big data policies in Iran. Findings reveal that Iran big data system requirements can be classified into four categories which indicate on increasing managed access to data while considering security and privacy, encouraging private and public sectors cooperation, transformation to smart governance, and establishing national data organization which would be responsible of data ID documents.

Keywords

Big data, System design thinking, System thinking, Policy, Latent dirichlet allocation (LDA), Topic modelling.

Article history

Received: 2024-02-12
Revised: 2024-05-18
Accepted: 2024-06-02
Published (Online): 2024-09-12

Number of Figures: 3

Number of Tables: 9

Number of Pages: 17

Number of References: 27



1. Introduction

Our contemporary world can be characterized by the continuous generation and consumption of data. Nowadays, all of us contribute to the production of data, whether through direct content creation or engagement on various platforms. Additionally, development of the Internet of Things (IoT) technology and the increasing usage of smart devices have further amplified the generation of vast amounts of data. The storage, analysis, processing, sharing, transmission, and utilization of this massive amount of data have become a very important concern for governments, organizations, companies, and even individuals in the present era (Fredriksson, et.al., 2017)

In light of the above, a new field of knowledge has emerged that concentrates into applications and concerns related to large-scale, diverse, and high-velocity data. This field of knowledge is big data (Ivanov, et.al, 2013). Governments are among the most significant entities dealing with the creation and utilization of big data. Such data typically holds considerable value and can greatly enhance governance decision-making process (Desouza and Jacob, 2017). However, there are many concerns about the preservation of privacy and security associated with these data (Fredriksson, et.al., 2017).

This study aims to determine requirements of Iran public sector big data system. The lack of this system, caused many problems for Iran governance system including disability to use national data for policy and decision making, data conflicts and cooperation issues between public sector organizations or between them and private sector companies.

In this paper, system design thinking approach has been used to analyse these requirements. In past decades, system thinking approaches and tools have played a crucial role in the development of software systems and system analysis has been a fundamental step in developing a new software system. However, due to the high dynamics and complexity of this field, traditional system approaches and tools are no longer widely applicable, and software industry practitioners generally utilize alternative methods for designing their products (Nerur, et al., 2005; Paetsch, et al., 2003).

In this study, design thinking as a new system approach was used to identify Iran public sector system requirements. To achieve this goal, first an analysis of Iranian opinions on big data on Persian-language social media and news websites was conducted. For this purpose, content from Persian news websites, Telegram channels, Instagram, and Twitter in 2022 were collected, including over 88,000 articles and approximately 3.8 million words.

By applying the Latent Dirichlet Allocation (LDA) method, researchers were able to extract latent topics from this dataset. LDA is an unsupervised machine learning technique that employs statistical methods to model the thematic structure of a discrete text corpora (Blei, et.al., 2003; Ostrowski, 2015). After that, a thematic analysis on all of the Twitter and Instagram posts was conducted in research dataset.

In continue, researchers examined the results of social media content analysis by interviewing experts. Based on this, a requirement list of Iran public sector big data system was constructed, which can serve as an input for formulating national policies in this field.

2. Literature review

In this section scientific papers and researches about big data and its applications and concerns in the public sector was reviewed. After that, design thinking as a practical research method to answer wicked problems was investigated; Problems like designing a national big data system.

2.1. *Big data and its management in the public sector*

Currently, the significance of data and the impact of big data on human life are widely discussed. However, it is essential to remember that just two decades ago, the concept of big data was virtually non-existent. Over the past few years, numerous studies have focused on the importance of big data for governments, the public sector, businesses, and citizens (Diebold, 2012).

2.1.1. *Definition and background of big data*

The term "data" originates from the Greek word "datomai," meaning "things given." In the past, data were primarily non-digital; however, the adoption of information technology by governments led to the generation of vast amounts of digital data. Big data is often defined using the three-Vs model, considering volume, variety, and velocity (Ivanov, et. al, 2012). Some have extended this model to include value, accuracy, and other factors (Chen, et. al., 2014; Wamba, et. al., 2015; White, 2012).

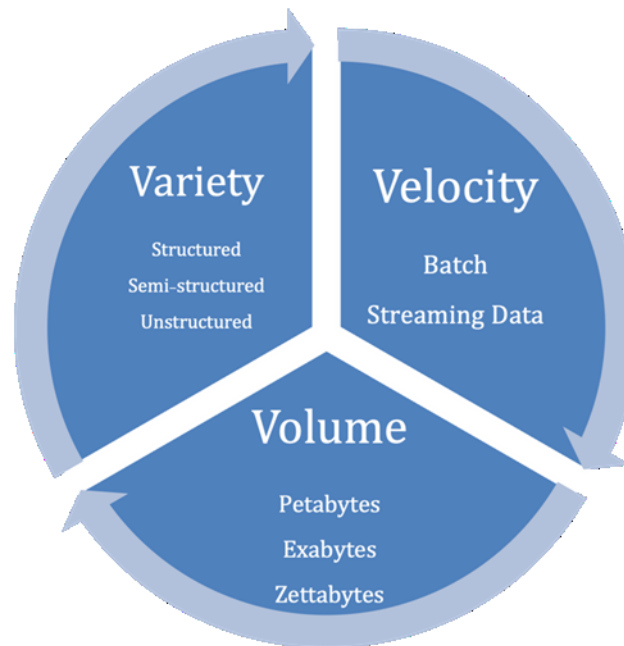


Figure 1. Big data 3Vs (Ivanov, et. al, 2012)

2.1.2. *Big data and the public sector*

Historically, governments were pioneers in using and storing data. They have dealt with data storage and encoding for various purposes, including administration, law, and governance (Henninger, 2013). The shift to digital technology in the 1990s led to the accumulation of vast amounts of digital data and governments were one of the main owners of large-scale data (Chen, et al., 2014). So, nowadays big data became particularly relevant for governments and its importance reflected in initiatives such as Barack Obama's executive order to assess big data capabilities for decision-making (Desouza & Jacob, 2017).

Fredriksson, et al (2017), conducted a systematic literature review on scientific papers about big data in the public sector. They reviewed 156 scientific articles in 5 main databases between 2006 and 2015. They have categorized the main challenges of Big Data in the public sector into three areas: data management, ensuring data quality, and ethical and privacy concerns.

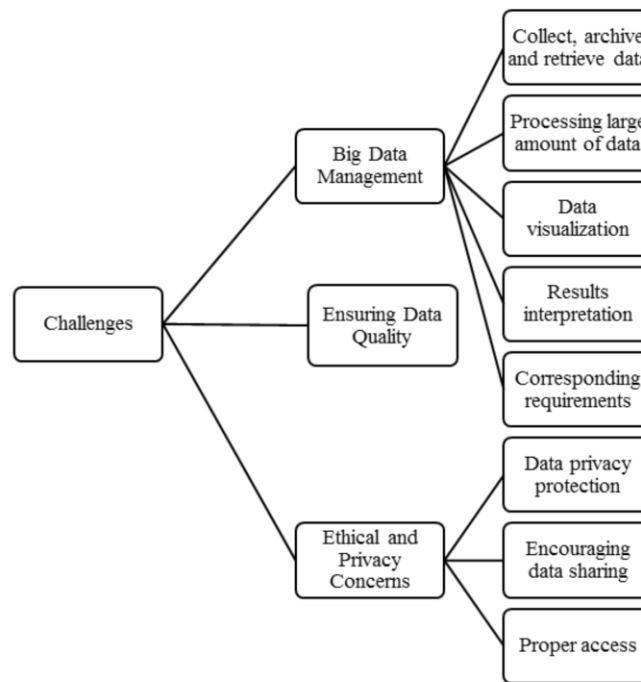


Figure 2- Challenges of big data in the public sector (Fredriksson, et al, 2017)

Daniell, et al (2016) focussed on the applications of big data in policy making. Big data can support politicians' decisions and help them to analyse their policies in fields like energy planning, urban transportation planning, medical emergency planning, healthcare, social services, national security, defence, government finance allocation, understanding public opinion, and fire and police services.

Using big data for policy making needs various sources and analyses. These sources and analyses can be categorized into three analyse types: social data analysis, historical data analysis, and predictive data analysis (Rahmanto, et. al., 2021).

Hong, et al (2019) discussed one of the applications of big data in public sector and analyzed role of big data in smart city policies and service design. They analyzed "night owl bus" service in the Seoul, South Korea as a successful public service design and illustrated role of "Big Data-based policymaking" in its success.

In this section the concerns and applications of big data in the public sector were reviewed. Goal of this paper is to find concerns and applications of big data for Iran public sector and identify requirements of a system that address these concerns and applications.

2.2. Design thinking

Tim Brown (2008) defines design thinking as "a discipline that uses the designer's sensibility and methods to match people's needs with what is technologically feasible and what a viable business strategy can convert into customer value and market opportunity".

[Lockwood \(2010\)](#) also describes design thinking as “a human-centred innovation process that emphasises observation, collaboration, fast learning, visualisation of ideas, rapid concept prototyping, and concurrent business analysis”.

Martin, among other design thinkers, has significantly influenced the field of design thinking with his definition. He characterizes design thinking as the "integration of effective analytical thinking and intuitive thinking" and combining these two distinct models of thinking ([Martin & Euchner, 2012](#)).

[Micheli, et al. \(2019\)](#) through a systematic review on all scientific articles published on design thinking found three main conclusions:

1. Design thinking is often used for solving problems, especially wicked problems ([Buchanan, 1992](#)).
2. Design thinking is an iterative process with a user-centric focus ([Brown, 2008](#)).
3. Some features of design thinking, such as creativity and innovation, are relatively general, while others, like abduction reasoning, are more specific ([Micheli, et al., 2019](#)).

Also, they classified conventional tools and methods of design thinking into eight categories:

1. Cognitive methods
2. Personas
3. Customer journey mapping
4. Brainstorming
5. Mind mapping
6. Visualization
7. Prototyping
8. Field experiments

2.2.1. design thinking and its relationship with systems thinking

A concept that has gained attention in various design thinking research studies is systems thinking, which has been referred to by different terms such as system thinking, holistic thinking, and gestalt approach ([Micheli et al., 2019](#)).

System thinking methods and tools played a significant role in the software system developments in the past decades, but due to the high dynamism and complexity of this field, traditional system tools are no longer widely applicable ([Paetsch et al., 2003](#); [Nerur et al., 2005](#)). Therefore, Investigating the roots of systems thinking and considering innovative system approaches to update these tools becomes essential.

By investigating system methods and approaches observers can find some that aims to solve wicked problems. soft systems methodology, strategic options development and analysis, drama analysis, viable system model, and critical systems heuristics are examples of these aims

(Checkland, 2000; Eden and Ackermann, 2001; Bennett and Howard, 1996; Beer, 1984; Ulrich and Reynolds, 2010).

Design thinking is another method that wants to handle wicked problems, so it seems plausible to recognising design thinking as a system approach in system of system methodologies.

3. Materials and methods

To identify the requirements of the Iranian public sector big data system, the following framework was used, based on design thinking approach:

1. Collecting data sets for user's opinions
2. Conducting LDA on data sets
3. Labelling each topic and designing requirements architecture version 1
4. Theme analysis on all collected articles of Twitter and Instagram
5. Analysing problem persona and finding a representative for each persona
6. Interviewing with each expert persona representative
7. Reviewing labels and requirements architecture based on current interview

This model is an iterative model, meaning that researchers repeat last two steps to considering all personas. It is design thinking way to ensure theoretical saturation is achieved.

To better implement the above method, also some design thinking tools were leveraged. Population-based methods such as interviews, and persona analysis and design are tools directly employed in this research. By examining the overall audience of this research, four personas can be identified from a specialized perspective:

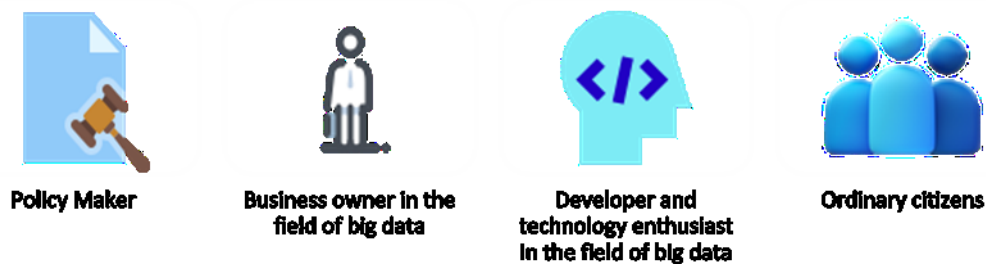


Figure 3- Research persona

In the topic modelling section, the ordinary citizens opinions were investigated. To ensure that experts' opinions were considered and refining topic labels, researchers interviewed with six expert persona representatives. Each of these individuals may adopt either a liberal or conservative approach to the big data field. Therefore, researchers made an effort to consider all persona of this research in interviews, including:

1. Two active policymakers in the Iranian big data field
2. Two business activists in the big data field
3. Two technical figures in the big data field

Each interviewee would be introduced in the next section.

To examine user opinions on big data in the context of social networks, an appropriate search query was designed and conducted searches on three popular social platforms, Twitter, Instagram, and Telegram and Persian news websites, over a one-year period. These media were chosen because of their popularity in Iran.

A search query comprising the term “big data” in Persian Language and in various formats was employed. The search was conducted from December 31, 2021, to December 31, 2022.

The search query was not restricted to obtaining the maximum volume of relevant content. Therefore, during the analysis process, all collected content was thoroughly examined, and those with irrelevant content were removed as noise.

The search resulted 300 tweets and 296 Instagram posts. It is worth noting that owing to the use of the official Twitter programming interface, the retrieved tweets represent all Persian tweets during the specified time period. However, because of various limitations on data collection imposed by Instagram, the search query was applied to users in the Dataak System, encompassing a majority of active Iranian Instagram users.

Table 1- Number of contents by platform

Platform	Number of articles	Number of words
News	85639	3,192,501
Telegram	1692	534,541
Instagram	296	60,504
Twitter	300	13,756

After collecting the content, the text of each article was normalized using the Hazm Persian language normalization algorithm. This normalization enables us to use various quantitative methods to process text.

In the next stage of this research, researchers focused on topic modelling collected dataset. Topic modelling, one of the most powerful techniques in natural language processing, has gained significant attention in recent years because of its ability to uncover latent thematic structures within large text corpora.

Latent Dirichlet Allocation (LDA) has emerged as a powerful and widely adopted probabilistic model for topic modelling in various domains (Jelodar, et al, 2019). Introduced by Blei, et al., (2003), LDA provides a generative probabilistic framework for representing documents as mixtures of topics, with each topic being a distribution over words. This influential work by Blei, et al., (2003) laid the foundation for numerous applications in natural language processing, text mining, and information retrieval. Ostrowski (2015) contributed to

the literature on LDA by exploring its applications and potential enhancements. Ostrowski's work likely extended the understanding of LDA in terms of its performance, limitations, and areas for improvement.

To conduct LDA topic modelling on our dataset, after data cleaning, stop-words were identified and a model was trained by extracting five distinct topics from each dataset. In the following table, parameters of our model that used with Gensim ([Řehůřek and Sojka, 2010](#)) python library can be reviewed.

Table 2- LDA model parameters

Parameter	Value
Number of topics	5
Number of words per each topic	10
Number of documents in each training chunks	100
Randomness	Used a number as a seed for random function (in this research 42)
Optimizing α	Gensim can use an algorithm to optimize α for best fit the corpora. This algorithm was used with setting the alpha parameter to "auto"

As topic modelling is an unsupervised approach, the discovered topics do not have labels. Subsequently, the identified topics from the media dataset were associated, and labels and descriptions assigned to each topic and after each interview and iteration, all of these were reviewed.

Following this, the theme of all the content collected from Twitter and Instagram were qualitatively analysed. Subsequently, through interviews with six experts representing the three specified expert persona, a deeper analysis was conducted, particularly focusing on the findings from the LDA section. These interviews helped us to gain a better understanding of users, based on the design thinking empathy concept, so in each iteration, the system requirements model was improved and validated.

In conclusion, by summarizing the above points, a comprehensive requirement list of Iran public sector big data system was developed. These requirements, as input in the technical design of Iran's big data architecture, empowers policy makers and engineers to consider various aspects of Iranians' concerns in the design of a new big data architecture.

4. Analysing data

After examining each of the selected platforms using the Latent Dirichlet Allocation (LDA) topic modeling method, five topics in each platform were identified and each topic includes ten keywords. So, the process of labelling each topic started, based on its keywords and researchers knowledge from literature review and some insights from context of our corpora. Also, after each iteration of research, the labels were refined.

The labels assigned to the identified topics on each platform and its keywords can be observed in the following tables.

Table 3- Instagram topics

Topic labels	Topic keywords
Security (At the Macro and international level)	Technology- Artificial Intelligence- Intelligence- International- Big-security- International relations- Peace- Capital- Organization
Business	Business- Company- Big- Technology- Digital- Plan- Presentation- Area-Financial- Development
Future technology	Internet- Intelligence- Artificial- Future- Information- Things- Technology-Smart- Production- Technology
Trending topic	Process- Registration- Period- Revolution- Management- Holding-Industrial- Hour- Big- Conference
The “everybody lies” book and role of big data in social analysis	Book- Lie- Big- #Big_Data- #Artificial_Intelligence- Social- Google- Data- #Technology- #Data_Science

The first topic concentrates on keywords about security. Role of big data in national and international security and concerns like peace, are obvious in this topic. The second and third topics are transparent but the last two ones need more explanation.

The fourth topic is about conferences and events in the field of big data. This topic was labelled as ‘trending topic’ because it’s a sign of attention to big data. The fifth topic is about a book. This book is about big data but its occurrence in our corpora related to a popular TV show and it can be considered as a noise.

Table 4- News websites topics

Topic labels	Topic keywords
Future technology and country development	Intelligence- Artificial- Technology- Area- Development- Country- Iran-Agreement- Year- Social
Legislation	Article- Approved- Technology- National- Country- Subsection-Development- Clause- Production- Ministry
Fifth generation mobile network	Iran-cell- Digital- Exhibition- Iran- International- 1400- Sentence-Generation- Transformation- Dey
Big data policy-making	Virtual- Space- Commission- Law- Council- Plan- Regulation- Approval-Parliament- Governance
Industrial development	Development- Technology- Information- Mobile- Company- Country-Industry- Big- Smart- Organization

Results of news corpora is more relevant to policy making. One of its reasons is that usually formal news media reflect experts and elites’ concerns, but these topics usually suffering lack of contents that reflect ordinary people views. The combination of social media and formal news media would help this research to consider all of problem persona point of view.

Table 5- Telegram topics

Topic labels	Topic keywords
Policy-making (Critique of passive, conservative, and directive policies)	Arena- Settlers- Society- Government- Life- Class- Cave-dwellers- Economy- Three- Monthly
Trending topic (Technical education in the field)	Education- Program- Programming- Learn- Intelligence- Artificial- Want to- Computer- 100- Engineering
Social analysis (Analysis of opinions in social media)	Report- Production- Analysis- Transition- Content- Media- Percentage- Opinion- Revolutionary- Weakness
Business development and large companies	Technology- Big- Digital- Area- Company- Information- Irancell- Social- Iran- Development
Media war and virtual space	Virtual- Space- War- Revolution- Enemy- Honor- Plan- Religious- Approval- Governance

Telegram channels are popular in Iran and many of other media owners, experts and organizations publish their thoughts in these channels. Therefore, its results are really relevant to our concerns and its topics are reflecting concerns of experts, elites, researchers and ordinary people.

The first topic is related to the ‘cave metaphor’. The cave metaphor indicates on open policy making versus passive, conservative, and directive policies. In addition, one of interesting points of Telegram topics are uniqueness of third and fourth topics. It shows higher level of contents of this media in comparison with other sources and it can be rooted on its variety and popularity in Iran.

Table 6- Twitter topics

Topic labels	Topic keywords
Security	Big (Persian Version)- Data- Big- Intelligence- Security- Artificial Intelligence- Free- Network- Virtual- Instagram
Artificial intelligence	Big- Data- Big (Persian version)- Intelligence- Artificial- Area- Analysis- Hand- Year- Account
Social control (China model)	Data- Big- Big (Persian version)- Project- Iran- China- Analysis- Information- Work- Government
Social media big data analysis	Data- Big- Big (Persian version)- Social- Network- Analysis- Government- Engineering- Can't- Publish
Machine learning	Data- Big- Book- Learning- Big (Persian version)- Intelligence- Work- Artificial- Machine- Analysis

Twitter is the last media that its topic was labelled. Political nature of this platform is reflected on its topics. Twitter usually shows thoughts of political citizens and their concerns about social control, like China programs, and ordinary people security and privacy issues can be seen in the table above.

All the topics of each media were presented in the following table:

Table 7- Labels assigned to the identified topics on each platform

Platform	Assigned label
Instagram	Security (at the Macro and International Level) Business Future technology Trending topic The “everybody lies” book and role of big data in social analysis
News websites	Future technology and country development Legislation Fifth generation mobile phones Big data policy-making Industrial development
Telegram	Policy-making (Critique of passive, Conservative, and Directive policies) Trending topic (Technical education in the field) Social analysis (Analysis of opinions in social media) Business development and large companies Media war and virtual space
Twitter	Security (User) Artificial intelligence Social control (China model) Social media big data analysis Machine Learning

Following the above examination, all Twitter and Instagram posts were read, and theme analysis was performed on them. In result, themes of all posts were identified and categorized into five concepts. The table below displays the results of this thematic analysis.

Table 3. Themes and concepts of Instagram and Twitter thematic analysis

Concepts	Themes
Ethics and privacy	Privacy and technology ethics Industrial security National security Social control Social consequences
Creating value from big data	Development of industries and businesses using big data Revenue generation through big data
Applications	Social analysis Prediction Decision-Making Policy making and strategy development Optimization Recommendation Fair allocation in society Identification and alerting of deviation (Error and fraud) Increasing responsiveness to change
Science and technology	Economic transformation with technological revolution and integration with other technologies Application in research across various branches of sciences
Policy and management	Policy-making for big data development Data governance and the open data approach Education, advocacy, and promotion Human resource development and migration of skilled human resources

After theme analysis, the process of interviewing six experts based on three expert persona and two approaches (liberal and conservative) identified in the previous section, was started. In the next table these persona were introduced:

Table 8- Interviewees

Expert code	Biography
Expert 1	Ph.D. in political science and bachelor of computer engendering. Expert of e-governance. Former manager of IT department in Iran free zones organization. Participant in many national IT projects. Worked with reformist and liberal politicians.
Expert 2	Participant in national big data projects. Expert in the field of social analysis based on big data. Partner and consultant of conservative intuitions.
Expert 3	Founder of pioneer big data company in Iran. Worked with many public and private organizations (Liberal point of view).
Expert 4	Founder of one of the well-known social listening and big data based think-tanks. Researcher of the big data field (Conservative point of view).
Expert 5	Chief technology officer of one of the main Iranian big data companies which designed and implemented remarkable big data systems in Iran (Liberal point of view).
Expert 6	Main developer of one of the most important national data platforms of Iran (Conservative point of view).

As expected, liberal interviewees concentrated on applications and benefits of big data and conservative interviewees concentrated on issues like national security; But besides these, some similarities are obvious between them. Interviewees opinion can be summarized as following:

- Lack of a working framework for contribution and cooperation of public and private sector players

When someone speak about public sector big data, he/ she is speaking about data that create and store by many public and private sector players. We have no working framework that these players can share their data together by considering security and privacy issues. This issue prevents forming a working big data system in many cases.

- Service oriented view against security oriented one

Security without service is meaningless. We can store all of our data in a closed system and prevent other players to use it, but there is no difference between this situation and the situation that we have no data. Our experts think that in many cases security concerns are just an excuse to not working.

When we prioritize service versus security, security concerns would be answered by the requirements of each service. For example, a standard access control systems can handle many of security concerns in this field. In addition, many of public sector data sets are open in democratic countries and transparency regulations force the government to publish them.

- Data ID document

One of the main problems of Iran data ecosystem is lack of data ID. Data ID is a standard document that indicates standard model of data presentation, access of these data, privacy statement and procedures to validate the data. Data ID would be a core document in each big data systems and prevent many further problems.

- Standardizing access controls

Designing a coherence access control system is the other main concern of our experts. In the absence of this crucial system. Everything is arbitrary and every organization and data owner would play a random game based on his/her aims and believes. For example, in current situation, chief data officer of an organization can give a full access to a data source or completely block any access to his organization data by his aims.

- High level reports and dashboards

Governments use big data to support their decisions and designing better services for their citizens. They cannot achieve these goals without using proper reports and dashboards that represents insights of their society big data.

- Data integration

Data conflicts are one of the main issues of the public sector big data systems. There are many cases that a data has multiple data sources and these sources doesn't represent similar result. This can be caused by data gathering errors, different criteria or deliberated manipulation of data. A working big data system will identify these conflicts and offer procedures for data integrity.

- Preventing corruption and data manipulation

Corruption is one of the most important issues of Iran governance system. National big data systems can play a crucial role to identifying and preventing corruption. Saving summary of each organization data in third party or decentralized databases can help us to identify and eliminate many types of data manipulation events.

By summarizing and discussing all of the above concerns, it illustrates the crucial role of national data organization. This organization can facilitate data standardization in the form of data ID, ensures that access controls are secure, fair and operational, design national flagship projects like big data framework, national big data dashboards, ..., and facilitate cooperation between all national big data players, including private and public sector players.

After identifying experts concerns, all of this research findings can be summarized by combining similar concerns. The following table is the result of this process. Our persona concerns and needs discovered by all the iterations were summarized into four main category which each category has some subcategory requirement. All of this process was audited by experts and these requirements were discussed with them as proposed in test stage of design thinking.

Table 9- Iran big data system requirements

Main requirements	Subcategory requirements
Access and security	Security- Access control- Privacy- Data integration
Public and private sector cooperation	Data sharing- Data commercializing- National flagship data projects
Smart and data-oriented governance	Service-orientation- Prediction and alarms- Data-driven decision making
National data organization and data ID document	Legitimation and policy making- Data standardization

5. Conclusions

In this study, the goal was determining requirements of Iran public sector big data system and finding them by focusing on stakeholders of this system. Our findings demonstrate any public sector big data system needs two core components: A national data organization and data ID.

National data organization cooperates with other organizations to identify every data type that should be collected and will determine responsible of this data, forms of data representation, the ways other organization can access this data and data validation process. All of these will be stored in the data ID document and will be shared with stakeholders. By publishing data ID, national data organization can audit security and privacy issues.

A determined data Id would help public and private sector players to collaborate. For example, internet taxies have many valid information about traffic that can help municipality for urban planning. By standardizing access to these data, private companies like internet taxies can sell their data to public sector organizations by checking it compliance with security and privacy concerns and vice versa.

By storing and analyzing standardized data, government can transform itself to a smart government that help it to make better decisions, propose customized service to its citizens, predict social crises and conduct programs to solve them, and eliminate many potential opportunities for corruption.

Finally, one of the most issues of this model is governance of this national data organization. This organization should be superior to all of the power branches including executive, legislation and Judiciary branches, military organizations and other branches of power.

Governance of this organization can be discussed in future studies. Also, discussing convergence of big data with other emergent technologies like blockchain and IoT in public sector can be the other future study that researchers propose.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

- Beer, S., 1984. The viable system model: Its provenance, development, methodology and pathology. *Journal of the operational research society*, 35(1), pp.7-25. <https://doi.org/10.1057/jors.1984.2>.
- Bennett, P. and Howard, N., 1996. Rationality, emotion and preference change drama-theoretic models of choice. *European Journal of Operational Research*, 92(3), pp.603-614. [https://doi.org/10.1016/0377-2217\(95\)00141-7](https://doi.org/10.1016/0377-2217(95)00141-7).
- Blei, D.M., Ng, A.Y. and Jordan, M.I., 2003. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan), pp.993-1022. <https://doi.org/10.7551/mitpress/1120.003.0082>.
- Brown, T., 2008. Design thinking. *Harvard business review*, 86(6), p.84.
- Buchanan R., 1992. Wicked problems in design thinking. *Design issues*, 8(2), pp.5-21. <https://doi.org/10.2307/1511637>.
- Checkland, P., 2000. The emergent properties of SSM in use: a symposium by reflective practitioners. *Systemic Practice and Action Research*, 13, pp.799-823. <https://doi.org/10.1023/A:1026431613200>.
- Chen, M., Mao, S. and Liu, Y., 2014. Big data: A survey. *Mobile networks and applications*, 19, pp.171-209. <https://doi.org/10.1007/s11036-013-0489-0>.
- Daniell, K.A., Morton, A. and Ríos Insua, D., 2016. Policy analysis and policy analytics. *Annals of Operations Research*, 236, pp.1-13. <https://doi.org/10.1007/s10479-015-1902-9>.
- Desouza, K.C. and Jacob, B., 2017. Big data in the public sector: Lessons for practitioners and scholars. *Administration & society*, 49(7), pp.1043-1064. <https://doi.org/10.1177/0095399714555751>.
- Diebold, F.X., 2012. On the Origin (s) and Development of the Term'Big Data'. <https://doi.org/10.2139/ssrn.2152421>.
- Eden, C. and Ackermann, F., 2001. SODA—the principles. *Rational analysis for a problematic world revisited*, pp.21-41. John Wiley & Sons Inc., United Kingdom.
- Fredriksson, C., Mubarak, F., Tuohimaa, M. and Zhan, M., 2017. Big data in the public sector: A systematic literature review. *Scandinavian Journal of Public Administration*, 21(3), pp.39-61. <https://doi.org/10.58235/sjpa.v21i3.11563>.
- Henninger, M., 2013. The value and challenges of public sector information. *Cosmopolitan Civil Societies: An Interdisciplinary Journal*, 5(3), pp.75-95. <https://doi.org/10.3316/informit.944373326039286>.
- Hong, S., Hyoung Kim, S., Kim, Y. and Park, J., 2019. Big Data and government: Evidence of the role

- of Big Data for smart cities. *Big data & society*, 6(1), p.2053951719842543. <https://doi.org/10.1177/2053951719842543>.
- Ivanov, T., Korfiatis, N. and Zicari, R.V., 2013. On the inequality of the 3V's of Big Data Architectural Paradigms: A case for heterogeneity. *arXiv preprint arXiv:1311.0805*. <https://doi.org/10.48550/arXiv.1311.0805>.
- Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y. and Zhao, L., 2019. Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia tools and applications*, 78, pp.15169-15211. <https://doi.org/10.1007/s11042-018-6894-4>.
- Lockwood, T., 2010. Design thinking in business: An interview with Gianfranco Zaccai. *Design Management Review*, 21(3), pp.16-24. <https://doi.org/10.1111/j.1948-7169.2010.00074.x>
- Martin, R. and Euchner, J., 2012. Design thinking. *Research-Technology Management*, 55(3), pp.10-14. <https://doi.org/10.5437/08956308X5503003>.
- Micheli, P., Wilner, S.J., Bhatti, S.H., Mura, M. and Beverland, M.B., 2019. Doing design thinking: Conceptual review, synthesis, and research agenda. *Journal of Product innovation management*, 36(2), pp.124-148. <https://doi.org/10.1111/jpim.12466>
- Nerur, S., Mahapatra, R. and Mangalaraj, G., 2005. Challenges of migrating to agile methodologies. *Communications of the ACM*, 48(5), pp.72-78. <https://doi.org/10.1145/1060710.1060712>.
- Ostrowski, D.A., 2015, February. Using latent dirichlet allocation for topic modelling in twitter. In *Proceedings of the 2015 IEEE 9th international conference on semantic computing (IEEE ICSC 2015)* (pp. 493-497). IEEE. <https://doi.org/10.1109/ICOSC.2015.7050858>.
- Paetsch, F., Eberlein, A. and Maurer, F., 2003, June. Requirements engineering and agile software development. In *WET ICE 2003. Proceedings. Twelfth IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, 2003*. (pp. 308-313). IEEE. <https://doi.org/10.1109/ENABL.2003.1231428>.
- Rahmanto, F., Pribadi, U. and Priyanto, A., 2021, March. Big data: What are the implications for public sector Policy in society 5.0 era?. In *IOP Conference Series: Earth and Environmental Science* (Vol. 717, No. 1, p. 012009). IOP Publishing. <https://doi.org/10.1088/1755-1315/717/1/012009>.
- Řehůřek, R. and Sojka, P., 2010. Software framework for topic modelling with large corpora.
- Ulrich, W. and Reynolds, M., 2010. Critical systems heuristics. In *Systems approaches to managing change: A practical guide* (pp. 243-292). London: Springer London. https://doi.org/10.1007/978-1-84882-809-4_6.
- Wamba, S.F., Akter, S., Edwards, A., Chopin, G. and Gnanzou, D., 2015. How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International journal of production economics*, 165, pp.234-246. <https://doi.org/10.1016/j.ijpe.2014.12.031>.
- White, M., 2012. Digital workplaces: Vision and reality. *Business information review*, 29(4), pp.205-214. <https://doi.org/10.1177/0266382112470412>.