# An Empirical Study on Data-driven Requirements Elicitation: Reflections from Nordic Enterprises

Alejandro Martinez[1], Marcus Melin[2], Georgios Koutsopoulos[2] and Jelena Zdravkovic[2]

[1] *Nexer Group AB, Valhallavägen 117G, 11553, Stockholm, Sweden*
[2] *Department of Computer and Systems Sciences, Stockholm University, Borgarfjordsgatan 12, 16407, Kista, Stockholm, Sweden*

**Abstract**

There is a plethora of digital data sources that may be exploited for collecting requirements for system development and evolution. In contrast to human sources, i.e. stakeholders, digital sources continuously generate data that is often not originally created for the purposes of requirements elicitation, e.g. on forums, microblogs, machine-generated trace logs, and sensor data. Streams of large volumes of data can be exploited to enable automation of a continuous requirements elicitation process using AI techniques that combine natural language or machine data processing, with machine learning. On the other hand, the complex characteristics of big data due to its size, lack of structure, high dynamics, and low predictability, present numerous challenges on the process of extracting requirements-related information that would be of a clear value for companies. The purpose of this interview study was to, from the practitioners' perspective, elicit their overall expectations and needs for a method for the elicitation of system requirements from digital data sources. Semi-structured interviews were conducted with several industrial experts from different business domains and the collected empirical data has been analyzed using thematic analysis. The results lead to the identification of a set of high-level requirements related to the method for the elicitation from digital data sources.

**Keywords**

Data-driven Requirements Engineering, Big Data, Requirements Elicitation, Agile Requirements Engineering, Enterprise Modeling

## 1. Introduction

Requirements engineering is a fundamental activity in the system development process of systems, as well as it is seen as a critical part of a project's success as many software deficiencies have their origin and can be traced back to the requirements [1]. Some of these deficiencies are due to requirements not contributing to creating the system that users need or request, or because of incomplete requirements, created or interpreted in a subjective way by the development team. The goal of requirements elicitation is to collect requirements of all relevant stakeholders. Furthermore, the requirements are documented, developed, and then frequently changed in the evolution part of the system's development life-cycle.

Owing to the digital transformation of the business sector, industries, and society in general — and the subsequent emergence of big data — the interest to consider digital sources of information for requirements has emerged, in addition to the traditional stakeholder-driven elicitation. As a result, there are ongoing efforts to support and enrich requirements elicitation activities by automatically processing and mining digital data for information about requirements [2, 3]. The main motivation behind Data-driven Requirements Engineering (DdRE) is to take advantage of large amounts of digital data related to a system in order to guide and support requirements engineers in their decisions about which requirements to include in subsequent system releases [4].

PoEM'21 Forum: 14th IFIP WG 8.1 Working Conference on the Practice of Enterprise Modelling, November 24–26, 2021, Latvia
EMAIL: alejandro.martinez@sigma.se (A. Martinez); marcus.melin1991@gmail.com (M. Melin); georgios@dsv.su.se (G. Koutsopoulos); jelenaz@dsv.su.se (J. Zdravkovic)
ORCID: 0000-0003-2511-9086 (G. Koutsopoulos); 0000-0002-0870-0330 (J. Zdravkovic)

Yet, the characteristics of big data in terms of volume, velocity, variety, and veracity present many challenges for effectively collecting and understanding requirements-related information, requiring different data processing techniques, as well as the identification of appropriate analytical algorithms. These decisions are highly dependent on the types of digital source that are targeted. Furthermore, as the data is not created explicitly for requirements elicitation, it tends to be limited in terms of completeness and correctness in respect to what is understandable as a system requirement and feasible to develop and implement. Therefore, the data-driven elicitation may require a substantial manual effort [5], or could be of low performance and therefore low practical use [6].

While many companies are increasingly interested in exploiting the potential value and opportunities that these data sources provide for requirements engineering, they face the challenge of understanding how to manage big data efficiently, as well as to how to integrate it with the current development methods in place. Previous research points that neither traditional nor agile development methods explicitly define how the collection and documentation of requirements should be done when the requirements are determined on the basis of diverse big data sources [2]. Furthermore, [7, 8, 9] advocate that there is a great need to define a complementary process for guiding how to use big data for requirements elicitation, where agility is considered relevant in this case as big data is created continuously over time, and under different business conditions. Therefore, the goal of this paper is to present an interview study devoted to industrial requirements for a data-driven RE method. The study is a part of an ongoing design science research project on this topic [7, 9, 10].

The rest of the paper is organized as follows. Section 2 presents a background on the dominant RE methods and the recent research in data-driven RE. Section 3 presents the method applied for conducting this semi-structured interview study. In section 4, results of the thematic analysis are reported together with the elicited requirements for data-driven requirements elicitation. Section 5 provides a discussion on the findings, while section 6 concludes the paper and outlines future research.

## 2. Background

Requirements engineering methods exist for decades. The best known is the plan-driven method, a sequential approach to system development that was introduced in the 70s, also known as “waterfall”. Two decades later, an incremental development approach emerged. In this approach the requirements are elicited in iterations where the core system functionality is defined in the first increments, and other features and quality requirements are left for later iterations. Based on this approach, the agile development method was proposed during the 2000s [11,12]. The method advocated continuous communication with the user, late and limited documentation, frequent system releases, and quick reactions to changes. In all methods, the elicitation phase implies obtaining intended statements related to a system under development that are transformed to a canonical format, which is understandable to software teams and feasible to develop and implement, such as to the semi-formal requirement structure in the plan-based approach [1] or to the user story template in agile approaches [12].

Data-driven RE has emerged during the last few years due to a wide range of online digital sources that may be exploited for requirements elicitation [2, 4]. Increasing attention has been given to sources that are more dynamic, i.e. continuously generating large amounts of data through various mediums and platforms, such as online discussion forums, app reviews, and microblogs, as well as to sensor data and machine-generated logs. Most of the efforts have been focused on identifying and classifying requirements-related information from few sources such as user feedback and machine-generated data. A systematic literature review showed that there is still lack of integration with the existing system development methods [3]. Recently, more holistic views on data-driven RE have been presented, which integrate information from different sources, based on domain ontologies [2], metamodeling [10], or contextual factors [9]. Although these studies address relevant research challenges, they pay less attention to the practitioners’ views and challenges with data-driven RE, which is still scarce and ad-hoc applied. Therefore, the aim of this study has been to investigate how practitioners see the challenge of eliciting data-driven requirements, and should it be methodologically supported.

When the digital data is in the form of free text, Natural Language Processing (NLP) is applied to extract requirements information. Identification of information in free text is commonly performed by the Named Entity Recognition (NER) task that concerns classifying word sequences to identify

requirements-related information, such as feature mentions. Machine Learning (ML) is applied to learn from historical data to perform the classification of requirements-related data. This learning process is typically supervised and relies on manual annotation of examples. This requires effort, but once an ML model has been created, it can be applied repeatedly to facilitate automation of requirements elicitation.

## 3. Method

This study is part of a design science research (DSR) project which aims to develop a method for data-driven requirements elicitation and management. According to the DSR methodology [13], exploratory semi-structured interview studies may be used for iterative requirements analysis of the design artefact. The exploratory semi-structured interviews of this study followed an interview guideline with a set of predetermined open-ended questions. The questions were structured into three themes – *current elicitation methods in use, the relevance of big data*, and the *viewpoints on the data's potential for improving requirements elicitation*. Using thus a deductive approach, a set of interview questions were derived for each of these three themes. A thematic analysis was used to analyze the answers on the given questions - to annotate and collate the interview transcriptions and summarize the findings using the qualitative data analysis. The categories and codes were created based on the interview data, i.e. citations, further giving rise to a set of high-level requirements related to the approach to the elicitation of data-driven requirements. Purposive sampling was conducted in order to select study participants among industry experts, where the respondents with different professional roles have been chosen in order to ensure a comprehensive analysis basis; selection criteria were: a) the respondent has a long work experience in IT; the respondent has >3 years' work experience in big data; the respondent has a long experience and expertise with requirements management. Each interviewee received one ~2 hour-long interview session (Table 1) starting with a brief introduction of data-driven RE research.

**Table 1**
Description and identification codes of the interviewees

| ID | Description of interviewee (respondent) |
|---|---|
| I1 | Business Intelligence Connoisseur in a consulting company in the data-driven business transformation. |
| I2 | Business Intelligence Developer in a Nordic IT company that offers solutions for business operations and management of IT services. |
| I3 | Senior Test Consultant in a consulting company for digital services and enterprise information systems. |
| I4 | Agile Coach in the world's largest music streaming and media services provider. |
| I5 | Chief Architect in a company providing integrated solutions for dairy and farming machinery production. |
| I6 | CTO of the company developing and publishing themed strategy video games. |
| I7 | Requirements Analyst in an online pan-Nordic financial services company. |

## 4. Results

In this section, the categories and codes that emerged during the thematic analysis of the interviews' data and the associated themes are presented. From the three themes - Method, Big Data, and Viewpoint and the corresponding questions, seven categories and 15 essential codes transpired. Based on the analysis, 16 requirements concerning a data-driven RE method were elicited.

All the respondents confirmed the need for using big data when collecting requirements. Furthermore, attention was paid to the consideration of big data sources - most of the companies already have access to a lot of data from several different types of digital sources but barely utilize them for requirements elicitation. At the same time, the respondents confirmed that the digitalization that takes place in the IT industry gives rise to new forms of requirements collection.

## 4.1. Theme: Method

The theme *Method* captures the respondents' knowledge on the use of big data for requirements, as well as their experiences of working with methods for system development, and in particular – requirements elicitation. All respondents confirmed that they do not have any systemized methods for collecting requirements from big data. The theme includes two categories of analyzed data, Digitalization (Table 2) and Approach to requirements elicitation (Table 3).

**Table 2**
Category: Digitalization

| Code | Citation |
|---|---|
| Use of big data for requirements elicitation. *Motivates R1* | *"When you use sensors, you kind of try to capture exactly what's going on, all of them here and get as a digital source and be able to read from…I can imagine just in zoos that it will be as large volumes of data that you want to be able to see what is actually needed based on that and like all these interactions, as well as, what the animals do, what happens to all machines - that there are many as digital footprints that would contribute to new requirements." (I2)* |
| Efficiency in requirements elicitation from big data. *Motivate R1 and R2* | *"Data should give us value and then you draw up a technical solution to get that data and then you still sit and think that you should have something that has power of BI or something else. Then I will analyze data and find these patterns and I think there will be much more computer-analyzed solutions." (I3)* *"I think all such smart tasks, requirements collection etc... come with digitalization so all Big-Data should be used. It creates the conditions for doing all that stuff. That's where performance really becomes important to check" (I1)* *"Digital user data is giving us the ability to monitor and act on the evolution of the requirements of our player base, particularly in order to improve performance, automation, and objectivity for new releases." (I6)* |

The respondents observe that the use of big data has a potential for improved requirements elicitation. They argue that the big data will provide value if it is used in a way that will improve performance in the elicitation – by automation of the collection, as well as by elicitation from a larger and wider user base. Based on the codes collated in this category, requirements R1 and R2 are elicited as follow:

*R1: There shall be a method to enable the collection and analysis of digital data from different digital sources, in an automated manner.*
*R2: The method shall enable repetitive collection and analysis tasks.*

**Table 3**
Category: Approach to requirements elicitation

| Code | Citation |
|---|---|
| Fit to existing approaches to development. *Motivates R3* | *"It's a mix of methods. Yes, we are quite too big changes, we are still quite project-driven." (I2)* *"We try follow a product centered approach with inspiration from DevOps.." (I4)* *"We have product owner that works a lot according to the waterfall method. So many waterfalls… sequential" (I3)* *"Large solutions need a deep understanding of successful requirements management" (I5)* *"We try to work agile but I would probably say that it is 50% agile" (I3)* |

| | |
|---|---|
| Highly iterative requirements elicitation. *Motivates R4* | *"We work towards the smallest feasible product development, frequent deliveries" (I1)*<br>*"So yes, we want to move more towards becoming more agile with new requirements but we are not there today." (I7)*<br>*"There should be continuous cycles of iterations" (I5)*<br>*"We need all time to know how our players think, how they play, to fix problems in the release…and to constantly please them by adding cool things..." (I7)* |

The respondents say that they need to combine different methods to be able to achieve some arbitrary results. At the same time, the respondents want to move towards agile methods, but this requires commitment from the entire business (organization). Constantly changing big data is the main reason why several respondents want to move towards more agile alternatives, with the possibility to work in short sprints for requirements collection and continuously analyze the needs. Based on the codes collated in this category, requirements R3 and R4 are elicited as follow:

*R3: The method shall be compatible with the organization's approach to system development.*
*R4: The method shall provide adequate support for short and continuous iterations for system development and evolution.*

## 4.2. Theme: Big Data

The theme *Big Data* captures the respondents' different types of big data sources that are available, the maintenance of data quality and the forms of big data analysis for identifying new needs and developing requirements. During the interviews, the respondents stated that they have access to different types of big data sources. The theme includes three categories of analyzed data, Data sources (Table 4), Data quality (Table 5), and Data analysis (Table 6).

**Table 4**
Category: Data sources

| Code | Citation |
|---|---|
| Consideration of any available big data source. *Motivates R5* | *"It's one of the shortcomings that we may still have, but that I think people have started to realize, is that we have too little customer involvement in our development." (I2)*<br>*"We should use most big data sources…large scope that identifies mass gaps that need to be addressed" (I1)*<br>*"Any customer transaction is a relevant data source" (I4)*<br>*"You should use sensors, to kind of try to capture exactly what is happening, and get as a digital source" (I2)*<br>*"Processing of NL sources is prioritized in our company due to their dominant quantity. However, we also utilize an eye-tracker sensor service for obtaining various statistics on players' behavior, i.e. concerning eye gazes and moves. Even the documents concerning ethics and privacy need to be watched in relation to the requirements constraints for game features" (I6)*<br>*"We have high interest to increase objectivity of management of players' feedback by collecting data from even other highly relevant sources, primarily from forums and Twitter. "* |

The respondents commented on several alternative sources that could be used to understand the users' needs related to the software product. This new form of data collection would result in a better and more accurate system development. The respondents believe that any digital source is worth

considering, or having in mind for the future. Based on the codes collated in this category, requirement R5 is elicited as follow:

*R5: The method shall enable the collection of data from both user- and machine-based digital data sources.*

**Table 5**
Category: Data quality

| Code | Citation |
|---|---|
| Identification of relevant data. *Motivates R6* | *"The main idea is to understand what you really need to understand and prioritize, that is, what is actually your user base" (I5)* *"Yes, the data quality really varies…everyone understands that it is important, but sometimes it costs much. Data management is a matter of priority." (I1)* |
| Use of methods for ensuring data quality. *Motivates R7* | *"Data needs to be sorted and filtered" (I2)* *"We need to understand data, make it usable" (I6)* *"The use of bad methods can lead to major data quality problems." (I3)* *"It is difficult to draw conclusions on big data, it needs to be structured and accurate" (I4)* |

Data quality is a factor that all the respondents consider important to systematically address. It is considered that one must sort collected data for the best possible use, which is something that is lacking for many respondents' companies. If the business does not have well-defined processes and techniques for handling data, the data quality will be lacking. Based on the codes created in this category, requirements R6 and R7 are elicited as follow:

*R6: The method shall prioritize data and the respective digital sources based on their relevance for the organization.*
*R7: The method shall include method components specialized for improving the quality of data without altering their meaning.*

**Table 6**
Category: Data analysis

| Code | Citation |
|---|---|
| Documentation and structure of data analysis means. *Motivates R8* | *"You need to so see if the strategy is successful - if big data shows good results."* *"Yes, it's much about the processes that exist for all data sources, if you have good processes and routines in place. That is, if things are well applied and documented, then you usually have fewer problems with data quality. While if you have poor control of your data and do not really understand how you got what, then you are in trouble" (I3)* |
| Learn patterns in data analysis. *Motivates R9* | *"Big data analysis is to identify behavior" (I4)* *"When requests for changes emerge different disclaimers need to be checked and added, so they need to be in the regulatory and applied." (I7)* *"To be effective we need to identify patterns in big data with AI" (I1)* |
| Predict new features. *Motivates R10* | *"We should be able to anticipate user needs and desires from current user behavior and comments" (I2)* *"We should be able to analyze user data to find strategies to increase software usage" (I5)* |

The majority of the respondents highlighted how big data could contribute with substantial advantages in system development and requirements collection. There are large amounts of data that are constantly generated and as such should be kept as the basis for broad and deep analyses. This data

could, for example, be used to develop repeatable analysis techniques. One important aspect is to identify and find patterns in big data in order to later be able to reap substantial benefits. Finally, the respondents believe that big data can be used to provide users with a need that they themselves were not aware of based on their behavior when using a software product. Based on the codes created in this category, requirements R8, R9, and R10 are elicited as follow:

*R8: The method shall be able to store and assess the methods and models for data analysis.*
*R9: The method shall be able to discover patterns during data analysis.*
*R10: The method shall include data analysis components for predicting users' future needs.*

## 4.3. Theme: Viewpoint

The theme *Viewpoint* captures the respondents' different opinions, experiences, and position on the requirements collection from big data. Many respondents discuss the future of requirements management and how it can possibly be affected by alternative sources than just traditional stakeholders. The theme includes two categories of analyzed data, Opportunities (Table 7) and Challenges (Table 8).

**Table 7**
Category: Opportunities

| Code | Citation |
|---|---|
| Continuous requirements collection. *Motivates R11 and R12* | *"It just increases and increases with the amount of data. It becomes as natural that it becomes more and more. Not just more and more sources" (I3)*<br>*"But yes, it feels like a sequel, if you can call it the 'Big Data World'" (I2)*<br>*"One needs to act on big data whole time…" (I1)*<br>*"Big data is relevant for continuation of requirements management, for software improvement" (I2)*<br>*"We should be able to control a behavior subconsciously…." (I5)* |
| Data availability. *Motivates R13* | *"Big data represents population level well." (I4)*<br>*"Then you had a completely different opportunity thanks to all this data storage" (I1)* |
| Identification of new behavior. *Motivates R14* | *"We want to know what our users are missing in our products! (I6)*<br>*"We need to know how exactly we could improve our online services" (I7)*<br>*"We should be able to analyze user data to find strategies to increase software usage" (I5)*<br>*"Users' sentiments are critical, for example negative one to automatically detect and have fast reaction when the number or impact of the players who express it is significant; this is important!" (I6)* |

All the respondents have identified a need to use big data sources when collecting requirements. With the advance of digitalization, the amounts of data increase and lead to new forms of data sources that can be used. You can discover new information that you did not have access to before. This is summed up as a continuation of requirements collection with big data, a future situation that should be used. Based on the codes created in this category, requirements R8, R9, and R10 are elicited as follow:

*R11: The method shall be able to continuously acquire relevant data from the respective sources.*
*R12: The method shall be able to combine data associated to similar behavior.*
*R13: The method should store collected data into a storage to enable analysis and learning.*
*R14: The method shall be able to recognize behavioral variations.*

**Table 8**
Category: Challenges

| Code | Citation |
|---|---|
| Difficulty to process big data. *Motivates R15* | *"Data-driven is new…it is difficult to handle large amounts of unstructured data" (I2)*<br>*"It is simpler with traditional stakeholders…here there are no processes and the type of different data you have to take in that way will be more difficult." (I1)*<br>*"Fixed routines are needed for the requirements management for big data" (I3)*<br>*"So what you constantly dig in is some kind of historical rubbish heap and you have to be aware of that, there are absolutely many things that are good to find there." (I4)* |
| Data is many, unstructured, immature. *Motivates R16* | *"…unstructured data cannot be used…" (I5)*<br>*"…it is difficult to understand big data requirements…" (I1)*<br>*"Yes, it is well with this as then if it is AI in that it is difficult to as well as for the important thing when you have a requirement is also to understand what is the expected result then. What do we think this will give us?" (I3)* |

The respondents suggest that collecting requirements from big data will be a challenge. It is a difficult process to handle large amounts of data and it is important for businesses to have clear methods and processes for handling the data. This is because there is a lot of data that is generated and does not always mean something, neither is relevant, nor adds any value. Two respondents believe that the degree of maturity of the business defines how easy or difficult such work with requirements collection will be. Finally, a respondent also emphasizes the importance of keeping in mind that one only analyzes historical data, which can be seen as a "historical rubbish heap". One cannot always predict the need for the future as it is old data that risks giving a misleading picture of a situation.

*R15: The method should be supported be a systematic process, with defined steps and guidelines.*
*R16: The method should enable creating meaningful data by the use of AI techniques.*

Figure 1 summarizes the requirements collected from the thematic analysis, and which codes, categories, and themes, they concern:

,

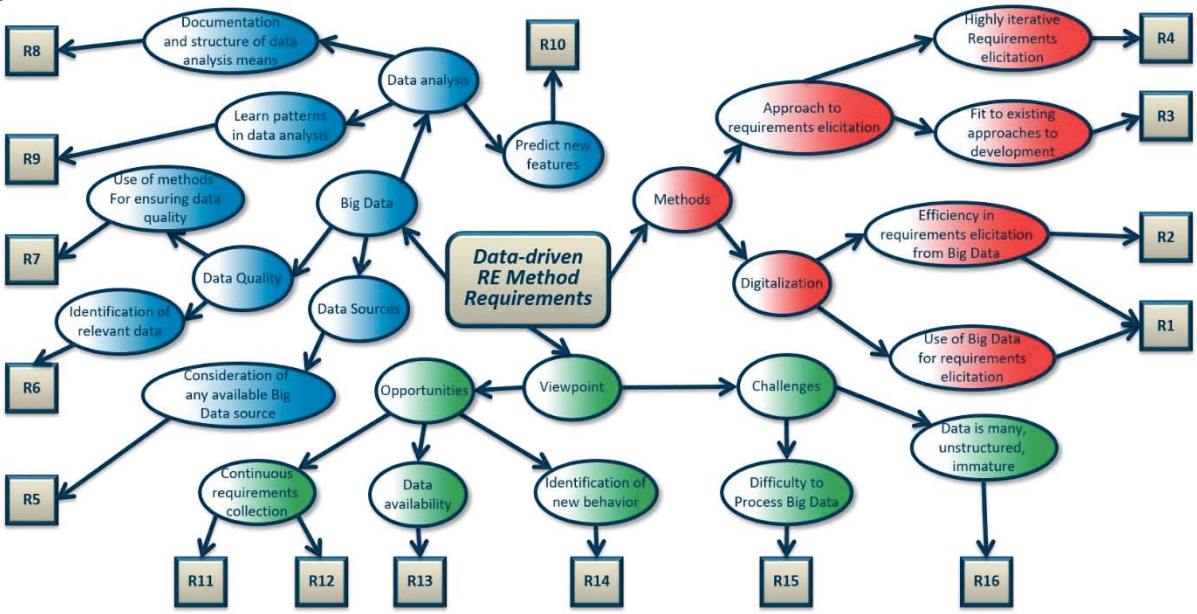


**Figure 1**: A map including the themes, categories and codes, along with the elicited requirements.

## 5. Discussion of Results

The themes that have been identified in this exploratory study reflect the state of the field. The experts expressed their requirements for development of the field in aspects related to methods, which relate to "how" questions and the procedure of eliciting requirements from big data, in aspects related to the big data itself, which reflect "what" questions that are associated to the nature of big data, and in specific aspects of viewpoints that reflect "why" and "why not" questions, in terms of opportunities and challenges involved in the process of transitioning towards a big data-driven state. Regarding the derived requirements in relation to the themes, there seems to be a balance between big data and viewpoints, since six requirements have originated from each theme and slightly less, in particular, four, have been derived from the method theme. This is a possible indication on which aspects of DdRE require more research attention and are of greater concern to the users, but this is a mere indication that should be further researched in future steps of this project for verification.

The experts that participated in this study have provided important insight towards the development of a specification for a method that aims to utilize big data to elicit requirements. However, apart from the direct answers and information, there seems to be a latent theme expressed by the participants. A finding that should not be neglected is the fact that, among the interviewees, a consensus exists regarding the complete absence of established methods for requirements collection with big data in their companies. This fact, in conjunction with their stated consensus on attempting to explore the field, seems to indicate a contradiction. All the interviewees are experts in the field, which means that they are aware of the benefits of adjusting to the new situation. They also state that they are willing to adopt data-driven requirement elicitation with big data, but there seem to be factors hindering the adoption that need to be identified and researched. One such potential hindrance may be the infamous "resistance to change" expressed during these exploratory interviews in terms of challenges related to the nature and processability of big data, and also in terms of compatibility with existing development approaches. Combined with tradition, resistance to change can be a strong factor that needs to be taken into consideration in all similar research projects, but it is not necessarily a negative phenomenon.

One way to respond to this situation is to ensure the quality and efficiency that has been emphasized by this study's participants and, in parallel, use the resistance as a driver to provide a method that supports the automatic elicitation of requirements using continuous, relevant and available data, as stated by the interviewees. This indication makes any theoretical argumentation on the benefits of a method seem weak unless it includes an aspect of operationalization. This operationalization can be complemented by a supporting tool. For this reason, the future steps of this project are the provision of method and tool support for the elicitation of requirements using big data. The complementing tool should be able to "mine" requirements using classification and facilitate the work of requirements engineers.

## 6. Conclusions, and Future Work

In this study, we have applied a qualitative approach with semi-structured interviews and thematic analysis to elicit the requirements for a method for the elicitation of system requirements from digital data sources (big data). 16 requirements classified to the themes *Method*, *Big Data*, and *Viewpoint*, have been elicited based on the codes generated during the thematic analysis. The main motivation for the study was to report the insights of industry experts on the topic of data driven requirements elicitation. The overall objective is to provide business organizations with a systematic aid for dealing with the complexity of digital data management in relation to available and prioritized data sources, amount and dynamics of the data in these sources, and different techniques for data processing and analysis. A framework is intended to complement existing approaches to requirements elicitation.

Even a theoretical validity of the lack of a method has been reported in the survey [3], a limitation of this study is that the elicited list of requirements might not be complete due to the given number of interviews; however, the last two interviewees (I6, I7) belonged to companies where the study has been performed over a period of several weeks and months [9] and by also discussing the topic with even other workers. This led to a saturation of insights in relation to the data that was collected in the interviews. Regardless of the possibility of being incomplete, the elicited requirements are considered

valuable and valid since they are derived from interviews with experts and practitioners working in companies that are daily dealing with large amounts of user or machine-driven data related to their software products. It is common that, in a DSR project, elicitation of requirements is done iteratively. Additional requirements or refined requirements, may emerge in the subsequent phases of the project when the initial version of the design artefact will be developed and validated.

The main direction for future work concerns elaboration of automation of the elicitation method, as well as classifying the use of processing techniques and algorithms according to the specifics of data sources and the software product to which they concern, and in accordance to that define reusable capabilities for elicitation [14]. Another direction of interest is a further development of the framework for model-driven engineering, to leveraging abstraction and automation in software development [15].

## References


[1] K. Pohl, Requirements engineering: fundamentals, principles, and techniques, Springer, Heidelberg ; New York, 2010.

[2] C. Quer, X. Franch, C. Palomares, A. Falkner, A. Felfernig, D. Fucci, W. Maalej, J. Nerlich, M. Raatikainen, G. Schenner, M. Stettinger, J. Tiihonen, Reconciling Practice and Rigour in Ontology-Based Heterogeneous Information Systems Construction, in: The Practice of Enterprise Modeling, volume 335, Springer, Cham, 2018, pp. 205–220.

[3] S. Lim, A. Henriksson, J. Zdravkovic, Data-Driven Requirements Elicitation: A Systematic Literature Review, SN Computer Science 2 (2021) 16.

[4] W. Maalej, M. Nayebi, T. Johann, G. Ruhe, Toward Data-Driven Requirements Engineering, IEEE Software 33 (2016) 48–54.

[5] M. van Vliet, E. C. Groen, F. Dalpiaz, S. Brinkkemper, Identifying and Classifying User Requirements in Online Feedback via Crowdsourcing, in: Requirements Engineering: Foundation for Software Quality, volume 12045, Springer, Cham, 2020, pp. 143–159.

[6] J. Dąbrowski, E. Letier, A. Perini, A. Susi, Mining User Opinions to Support Requirement Engineering: An Empirical Study, in: Advanced Information Systems Engineering, volume 12127, Springer, Cham, 2020, pp. 401–416.

[7] A. Henriksson, J. Zdravkovic, A Data-Driven Framework for Automated Requirements Elicitation from Heterogeneous Digital Sources, in: The Practice of Enterprise Modeling, volume 400, Springer, Cham, 2020, pp. 351–365.

[8] S. Lou, Y. Feng, H. Zheng, Y. Gao, J. Tan, Data-driven customer requirements discernment in the product lifecycle management via intuitionistic fuzzy sets and electroencephalogram, Journal of Intelligent Manufacturing 31 (2020) 1721–1736.

[9] X. Franch, J. Ralyté, A. Henriksson, J. Zdravkovic, Data-driven Agile Requirements Elicitation through the Lenses of Situational Method Engineering, IEEE Computer Society, 2021, pp. 402–407.

[10] A. Henriksson, J. Zdravkovic, Holistic data-driven requirements elicitation in the big data era, Software and Systems Modeling (2021).

[11] K. Beck, M. Beedle, A. Van Bennekum, A. Cockburn, W. Cunningham, M. Fowler, J. Grenning, J. Highsmith, A. Hunt, R. Jeffries, others, Manifesto for agile software development (2001).

[12] K. S. Rubin, Essential Scrum: a practical guide to the most popular agile process, Addison-Wesley, Upper Saddle River, NJ, 2012.

[13] A. Hevner, S. T. March, J. Park, S. Ram, Design Science in Information Systems Research, MIS Quarterly 28 (2004) 75–105.

[14] J. Zdravkovic, J. Stirna, J.-C. Kuhr, H. Koç, Requirements Engineering for Capability Driven Development, in: The Practice of Enterprise Modeling, volume 197, Springer, Berlin, Heidelberg, 2014, pp. 193–207.

[15] O. Pastor, Model-Driven Development in Practice: From Requirements to Code, in: SOFSEM 2017: Theory and Practice of Computer Science, volume 10139, Springer International Publishing, Cham, 2017, pp. 405–410.