Types of Analytics in Requirements Engineering

Marina Pincuka and Marite Kirikova[0000-0002-1678-9523]

Institute of Applied Computer Systems, Faculty of Computer Science and Information Technology, Riga Technical University, Kaļķu 1, LV-1658, Riga, Latvia
marina.pincuka@rtu.lv, marite.kirikova@rtu.lv

Abstract. Different methods of analytics are popular in various areas of human activities. This also refers to requirements engineering. However, the number of research works on the usage of analytics in requirements engineering is still limited; and the issues addressed in these works are rarely surveyed, structured and organized so that the knowledge of the use of analytics in requirements engineering could be reused and utilized effectively. This paper contains a preliminary survey of types of analytics used in requirements engineering and a mapping of the types of analytics onto the continuous requirements engineering framework. This work is a step towards the effective use of analytics in requirements engineering. The survey of the types of analytics points to the sources of existing knowledge of the use of analytics in requirements engineering. The mapping of the types of analytics onto the requirements engineering framework depicts the sources of data that can be used in different types of analytics.

Keywords: Requirements Engineering, Analytics, Data Analysis.

1 Introduction

Analytics is the application of computer systems to the analysis of large data sets for the support of decisions in a particular domain [1]. The use of analytics gives an opportunity to utilize information which is hard or impossible to handle manually. Today different types of analytics are applied in almost all areas of human and machine activity, including requirements engineering.

The purpose of this paper is to reflect the preliminary results of the research in progress concerning the use of analytics in requirements engineering. It reports on two research questions: (i) "What is the state of art of the use of analytics in requirements engineering?" and (ii) "What are the main sources of data to be used in different types of analytics in requirements engineering?",

A survey of related works in requirements engineering analytics was conducted to answer the first research question. Further, in order to have a structured view on the sources of data used by different types of analytics, a mapping of the types of analytics onto the requirements engineering framework was performed. Thus, the paper contributes a survey of analytics methods used in requirements engineering and provides a structured view on the sources of data that can be utilized by different types of analytics in requirements engineering. It must be noted, that the results refer to currently available sources of literature. Due to popularity of analytics applications and fast development of the field, it is important to take into account that the survey will have to be updated at least once per year until the time when the area of analytics application in requirements engineering will reach a higher level of maturity.

The paper is structured as follows. The survey of types of analytics used in requirements engineering is presented and discussed in Section 2. The mapping of the types of analytics (identified in Section 2) onto a requirements engineering framework is demonstrated in Section 3. Brief conclusions and directions of further research are stated in Section 4.

2 A Survey of Types of Analytics used in Requirements Engineering

To get a preliminary view on the state of art in the use of analytics in requirements engineering, a literature search was conducted using terms "analytics" AND "requirements engineering". Nine relevant sources were selected using IEEE, Springer, ACM, and Science Direct resources. The gathered sources were analysed as follows. First, it was identified which kinds of analytics are used in each source for what purposes. Second, the identified usages were grouped by identified kinds of analytics. The gathered kinds of analytics and their brief definitions are described below:

- Advanced analytics – focuses on forecasting future events and behaviours, allowing businesses to conduct what-if analysis to predict the effects of potential changes in business strategies [2]. This kind of analytics uses historical data to analyse alternative actions.
- Big Data analytics – Big data analytics examines large amounts of data to uncover hidden patterns, correlations and other insights [3], [4].
- Descriptive analytics – the field of study were a (typically large) dataset is described quantitatively on its main features with the aim to reduce the amount of data into 'human consumable information' [5], [6]. Descriptive analytics provides insights about the past and current business performance [7].
- News analytics – the measurement of the various qualitative and quantitative attributes of textual (unstructured data) news stories [8].
- Predictive analytics – the field of study where a prediction is made about the future based on information from the past and current situations [5].
- Prescriptive analytics – the field of study in which the actions are determined that are required to achieve the goal. Prescriptive analytics only determines what will happen if we continue the current trend of activities [5]. It usually works with structured data [5].
- Text analytics – involves information retrieval, lexical analysis to study word frequency distributions, pattern recognition, tagging/annotation, information extraction, data mining techniques including link and association analysis,

visualization and predictive analytics [8]. Here usually unstructured data is used for the analysis.

- Visual analytics - creates a path (from data to decisions) that enables the decision makers to extract insights by interacting with the relevant information [9]. The use of unstructured data to represent it in understandable visualizations is discussed in [10].

Basing on the analysis of the use of different types of analytics that were gathered during systemic literature review and are listed above, a grouping of data used in analytics is proposed. This grouping helps to understand which types of analytics can be used and applied in which situations of requirements engineering. The following four groups of data where revealed:

- Unstructured data – analytics works with unstructured data.
- Large datasets – analytics works with datasets that usually are large.
- Historical data – analytics works with past information.
- Structured data – analytics works with structured data.

The groups of data are not mutually exclusive. They just point to the main data variations addressed in current researches that report on the use of analytics in requirements engineering. Based on the groups of data, the analyzed analytics methods were organized in 8 overlapping types, which together with the examples of their application are reflected in Table 1. These types of analytics are further used for the mapping of the analytics methods onto the requirements engineering framework in Section 3.

Table 1. Analytics in requirements engineering

Types of analytics	Application examples in the related work
Advanced historical data analytics	Future prediction [2]
Big Data large datasets analytics	Understanding customers needs, Understanding and optimization of business processes, Performance optimization [3]
Descriptive large datasets analytics	Requirements analysis, History analysis [11]
News unstructured data analytics	Measure qualitative and quantitative datasets [8]
Predictive historical data analytics	Risk mitigation, Future prediction [11]
Prescriptive structured data analytics	Future prediction [11]
Text unstructured data analytics	Pattern recognition, Annotation, Information extraction, Future prediction [8]
Visual unstructured data analytics	Requirements analysis, Requirements visualization, Displaying Data [9]

The related works discussed in this section reveal that the use of analytics in requirements engineering can help to identify the requirements which otherwise might be overlooked or un-recognized. For instance, in large software systems (system size, functionality breadth, component maturity, supplier heterogeneity), it is advisable to apply software repository mining for understanding, evaluating, and predicting the development, management, and economics of such systems [12]. The use of analytics can also help to correct wrong requirements. For instance, by the use of visual analytics, with the dashboard and sketching systems interface [10] we can capture issues like a field with incorrect input data or button with wrong functionality. Analytics can also help to improve the quality of collected requirements, for instance, by analysing large datasets from similar (e.g. open source) software systems [3]; all common functions can be collected and, based on them, the missing system functionalities retrieved.

At the current stage of research, the overlapping types of analytics and overlapping groups of data sources; and continuous emergence of new analytics approaches are a challenge for designing a method for the choice of particular types of analytics for particular requirements engineering cases. To overcome this challenge, it was decided to design a method that helps to gradually amalgamate the acquired knowledge so that it might be extended and refined; and, simultaneously, already now, could be put in the practical use to arrive at generic recommendations and methods with respect to the choice of the types of analytics in particular cases of requirements engineering. The proposed approach is discussed in the next section.

3 Data Sources Based Mapping of Types of Analytics onto Continuous Requirements Engineering Framework

To understand how analytics can be used in requirements engineering, we have designed a simple method of mapping the known types of analytics onto the requirements engineering framework. The method consists of the following five steps (the first three steps refer to the representation of the existing knowledge about the use of analytics in requirements engineering; the last two steps refer to the use of this knowledge in actual requirements engineering cases):

1. Choose the requirements engineering framework, such that its components would allow distinguishing between analytics data sources.
2. Map visually each type of the analytics onto the framework so that it would be visible which component of the framework provides data for which type of analytics.
3. Repeat the second step whenever the new type of analytics or the new data source for the already mapped types of analytics is detected in the related work.
4. For a specific case of requirements engineering, identify possible sources of data for particular types of analytics.
5. Apply those types of analytics where data are available, if these types of analytics can provide useful knowledge in the requirements engineering process.

The choice of a particular framework onto which to map the types of analytics depends on the preferences of users. Each framework that satisfies the conditions mentioned in the first step of the method is applicable. We have chosen the FREEDOM framework that was developed for the purposes of continuous requirements engineering and is reflected in Fig. 1 [13], [14]. The framework is suitable for the method because of the following reasons:

- It gives an opportunity to distinguish between different data sources.
- It gives an opportunity to distinguish between internal and external data sources with respect to the enterprise or project where the requirements engineering is applied.
- It has a fractal nature [14], i.e. requirements engineering, on a smaller scale, is a constituent of each component of the FREEDOM framework.
- The framework itself prescribes the usage of monitoring, auditing and analytics (maa links in Fig. 1).

The FREEDOM framework has the following constituents-functions (see Fig. 1): F– Future representation, R – Reality representation, E1 – requirements Engineering, E2 – fulfillment Engineering, D – Design and implementation, O – Operations, and M– Management.

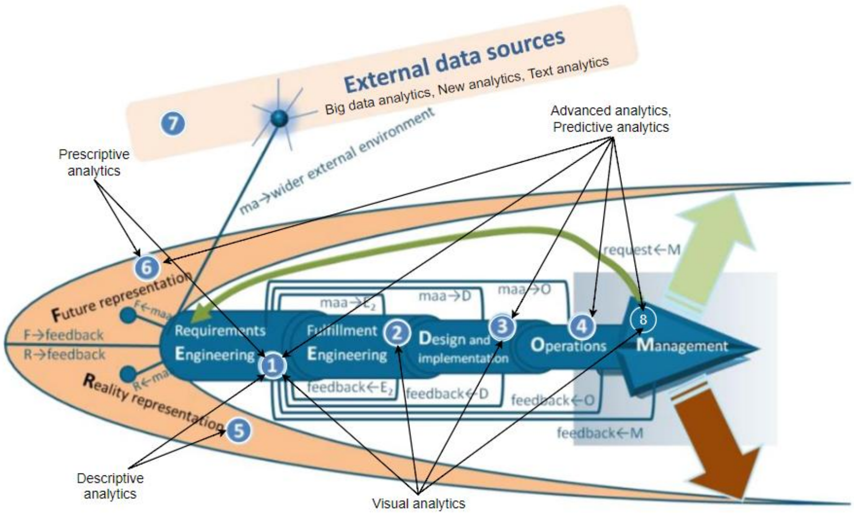

Fig. 1. An example of mapping between the types of analytics and the requirements engineering framework.

F – Future representation is the constituent of the framework that is responsible for representation of the To-Be situation. In Fig. 1 the sources of data in this function are represented with (6).

R – Reality representation is responsible for all artifacts that represent the present (As-Is) situation. The sources of data in this function are represented with (5)

E1 – requirements Engineering is the function dedicated to the model and tool based acquisition and management of high quality requirements. The sources of data in this function are represented with (1).

E2 – fulfillment Engineering is the function that takes care of handling project portfolios that would lead to the fulfillment of stated requirements. The sources of data in this function are represented with (2).

D – Design and implementation is the function that produces the design and handles implementation of the target system. The sources of data in this function are represented with (3).

O – Operations regard the actual operation of the implemented system, including its maintenance. The sources of data in this function are represented with (4).

M – Management refers to all levels of management under which the target system operates [13]. There is number (8) that refers to data in management function, however, it has to be taken into account that in Management function information linkages differ from those of other functions of the framework.

Number (7) is assigned to data sources external to the enterprise. This data can be used for Big Data analytics, news analytics and text analytics that refer to the external data sources (7) in Fig. 1 [3], [15]. Numbers (1) to (6) refer to internal enterprise data.

As it is shown in Fig. 1 and in Table 1, the analytics of one and the same type can use data from several functions (components) of the framework. For instance, prescriptive analytics helps organizations make better decisions by optimizing trade-offs between business goals, such as costs or customer service, while considering predictions, rules, and constraints on available resources, to recommend the best course of action; whether decisions are made on a configuration, design, plan, or a schedule [5]. Thus, the prescriptive analytics can be used not only for the data (6) of Future representation (analyzing data collected in past to predict the future changes), but also for data (1) in requirements Engineering component of the FREEDOM framework (applying analytics to the data of previous requirements gathering cases can help to improve requirements of systems in future).

Predictive analytics has a lot of capabilities - visualization, modeling, data mining management and deployment [5], [7], so predictive analytics can be used in (6) Future representation (analyzing past events to predict changes in future); in (1) requirements Engineering (gathering new requirements while analyzing existing systems); in (3) design and implementation (achieving the most effective development alternatives while modeling each development phase); and in operations and management component (applying present and future analytics based on historical data). Assuming that prescriptive analytics and predictive analytics both are advanced analytics technologies [2], we can make a conclusion that advanced analytics can be used in all functions (components) of the FREEDOM framework that are mapped to predictive and prescriptive analytics. Descriptive analytics can be a preliminary stage of data processing that creates a summary of historical data to yield useful information and possibly prepare the data for further analysis [2], [6]. Descriptive analytics can be used to represent both: the past and the reality, and to use collected information in requirements engineering. The basic idea of visual analytics is to visually represent the data so as to allow the employees to directly interact with the information [10], [16]. It means that visual analytics can be used not only for data sources (1) in the requirements Engineering function/component (using dashboards to draw sketches), but also, for instance, in data sources (3) in the Design and implementation function (e.g. making charts with phase goals, to keep the track of development progress).

The mapping shown in Fig. 1 gives an opportunity to see where from the data can be taken in the process of requirements engineering and what types of analytics can be applied to this data. Such representation helps to make decisions about what data analytics might be useful and possible in particular requirements engineering situations.

4 Conclusions

In this paper we discussed the usage of analytics in requirements engineering. The paper contributes (i) a preliminary survey on the use of different types of analytics in requirements engineering, (ii) a mapping of analytics types onto the requirements engineering framework that helps to visualize and amalgamate the knowledge on the use of analytics in requirements engineering.

The presented research has several limitations (i) the surveyed types of analytics overlap, (ii) also the identified sources of data, used in analytics, overlap, (iii) this is only a preliminary survey – more sources might be identified with more sophisticated search methods, and (iv) the proposed method of mapping the types of analytics onto the requirements engineering framework is manual, and thus has a restricted representation flexibility.

Nevertheless, the contribution of this research in progress is a step towards the effective utilization of different types of analytics in requirements engineering. The further research will include overcoming of the above listed limitations and developing the software tool that supports the effective use of requirements engineering analytics.

References

1. Runkler, T.A.: Data Analytics: Models and Algorithms for Intelligent Data Analysis. 2nd Edition, Springer (2016).
2. Rouse, M.: WhatIs.com. (2014), http://searchbusinessanalytics.techtarget.com/definition/
3. Elgendy, N., Elragal, A.: Big Data Analytics: A Literature Review Paper. In: Proceedings with 14th Industrical Conference – ICDM 2014, St.Peterburg, Russia, July 16–20, 2014, pp. 214–228, Springer (2014).
4. Soltanpoor, R., Sellis, T.: Prescriptive Analytics for Big Data. In: Proceedings with 27th Australasian Database Conference – ADC 2016, Sydney, NSW, September 28–29, 2016, pp. 245–256. Springer (2016).
5. Sappelli, M., et al.: A Vision on Prescriptive Analytics. In: Proceeding of The Third International Conference of Big Data, Small Data, Linked Data and Open Data (includes KESA 2017), pp. 45 – 50 (2017).
6. Olson, D. L.: Descriptive Data Mining. Computional Risk Management, Lincoln, USA, p.114 (2016).
7. Nasiri, A., Nalchigar, S., Yu, E., Ahmed, W., Wrembel, R., Zimányi, E.: From Indicatiors to Predictive Analytics: A Conceptual Modelling Framework. In: Proceedings of 10th

IFIP Working Conference – PoEM 2017, Leuven, Belgium – November 22- 24, 2017 pp. 171–186. Springer (2017).

8. Batrinca, B., Treleaven, P. C.: Social media analytics: a survey of techniques, tools and platforms. AI & Society, pp. 89 – 116 (2015).
9. Reddivari, S.: Visual Analytics for Software Requirements Engineering. In: Proceedings of the RE 2013, Doctoral Symposium, pp. 389 – 392 (2013).
10. Niu, N., Reddivari, S., Chen, Z.: Keeping Requirements on Track via Visual Analytics. In Proceedings of the Requirements Engineering Conference, 2013 21st IEEE International (RE), pp. 205–214 (2013).
11. SAS Institute Inc., SAS, (2013-2018). https://www.sas.com/
12. Kirikova, M., Purmalietis, K.: Analytics in Continuous Requirements Engineering. In: Joint Proceedings of REFSQ-2017 Workshops, Doctoral Symposium, Research Method Track, and Poster Track co-located with the 22nd International Conference on Requirements Engineering: Foundation for Software Quality (REFSQ 2017), Essen, Germany, February 27, 2017, CRE'17, Ceur-ws.org, Vol.1796, pp. 1–9 (2017).
13. Kirikova, M.: Continuous Requirements Engineering in the FREEDOM Framework: a Position Paper. In: Joint Proceedings of REFSQ-2016 Workshops, Doctoral Symposium, Research Method Track, and Poster Track co-located with the 22nd International Conference on Requirements Engineering: Foundation for Software Quality (REFSQ 2016), March 14-17, 2016, Gothenburg, Sweden. CEUR-WS.org, Vol. 1564, pp. 1–6 (2016).
14. Kirikova, M.: Continuous Requirements Engineering. In: Proceedings of the 18th International Conference on Computer Systems and Technologies – CompSysTech'17, Ruse, Bulgaria, June 23–24, 2017, pp. 1–10. ACM Press (2017).
15. Tsai, C. W., Lai, C. F.: Big data analysis: a survey. Journal of Big Data. p. 32 (2015).
16. Steed, C., Swan, J. E., Fitzpatrick, P. J., Jankun- Kelly, T. J.: Innovative Approaches of Data Visualization and Visual Analytics: A Visual Analytics Approach for Correlation. Classification, and Regression Analysis, Huang, M. L., Huang, W., Sydney, Australia, CSIRO ICT Centre, 464p. (2013).