Enabling Digital Transformation Using Secure Decisions as a Service

Hans-Joachim Hof[1], Rainer Schmidt[2(✉)], and Lars Brehm[3]

[1] MuSe - Munich IT Security Research Group, Munich University
of Applied Sciences, Lothstrasse 64, 80335 Munich, Germany
`hof@hm.edu`
[2] Munich University of Applied Sciences,
Lothstrasse 64, 80335 Munich, Germany
`Rainer.Schmidt@hm.edu`
[3] Munich University of Applied Sciences,
Am Stadtpark 20, 81243 Munich, Germany
`Lars.Brehm@hm.edu`

Abstract. Digital Transformation is of crucial importance for many enterprises and creates new challenges both on the conceptual and architectural level. Therefore, a new, secure, and layered architecture is introduced that separates the process from the decision model in order to quickly react to changed requirements. It provides flexibility by separating three aspects of decision-making: foundations, methods, and data. Security and reliability is achieved by using a reputation system to judge data sources and data contributors. The reputation system offers a score for trustworthiness of a data source that could be used for the selection of appropriate data sources for decisions as well as a basis to calculate a confidence score of a decision.

Keywords: Digital transformation · Business process · Innovation · Decision as a Service · Security

1 Introduction

The Digital Transformation of Enterprise Architectures [1] and business processes is an issue of paramount importance for many enterprises [2]. Enormous benefits by digital transformation are predicted [3, 4]. Digital transformation creates new opportunities for enterprises [5], breaks down industry barriers but also destroys existing business models [5]. Serious disruptions take place, new companies emerge [6] and established companies disappear [7]. A number of strategies have been developed in order to drive digitization [4, 8].

In order to improve the capability of enterprise to react quickly to customer and market requirements, a cornerstone of digital transformation is the improvement of decision support within enterprise. Means both include the speed-up of decisions by automation and the increase of decision precision. To speed up decisions, enterprises replace human decisions by automated decision systems. The precision of decisions is increased by taking into account more data and data from more sources. An example is

the analysis of data from social software [9] such as social media. It allows measuring the customer sentiment faster and with more precision than traditional means, e.g. surveys.

Big Data receives a great deal of attention in industry and research for improving decision making [10]. The largest benefits of Big Data can be leveraged by companies engaged in the business areas of IT and technology development and marketing [11]. Big Data is not a specific technology or technology platform such as Hadoop [12], but embraces a series of technological advances creating a significant expansion of the analytical capabilities.

Today Big Data technologies are often integrated into Decision as a Service [13] offerings. Decision as a Service relieves companies of duties as the data collection and the continuous improvement of decision models. Additional, companies have the option to select the best offers from an emerging market. A further benefit of Decision as a Service arises from the fact that the service provider who provides the services has access to more data, as the company due to his aggregator role. In this way, the decision quality can be increased and the time to improve decision-making processes can be shortened.

To enable faster and more precise decisions, many enterprises use Big Data [10] to exploit also external data sources. In this approach, data from different sources are combined during decision-making. Unfortunately, this creates a new security threats. Malicious contributors could distort data in order to influence the final decision in their interest.

The contribution of this paper is to introduce a new architecture for decision as a service that includes a reputation system that could be used to protect against these attacks. The reputation system imposes a goodness score on each contributor, judging on the trustworthiness of the contributor.

The paper proceeds as follows, first decision support technologies in digitized enterprise are introduced. Then an architecture for secure decision as a service is introduced. The reputation system that allows evaluating contributors is defined in the following section. The next chapter shows up the use of decision as a service for digital transformation. After discussing related, a conclusion and outlook is given.

2 Basic Architecture for Secure Decision as a Service

The proposed architecture supports decisions using three essential elements. Foundations such as Laws define the basic rules of decisions. Methods define the way to create these decisions. Data are used to make concrete decisions. Foundations, Methods and Data evolve at different speeds. Foundations are the slowest, data are the quickest element. There our architecture differentiates these elements and thus enables their independent evolution (Fig. 1).

Decisions are based on a number of *data sources* in the proposed architecture. Each data source provides data from one or more data *contributors*. For example, a hotel rating website may be a single data source, users writing reviews of hotels are the contributors associated with this data source.
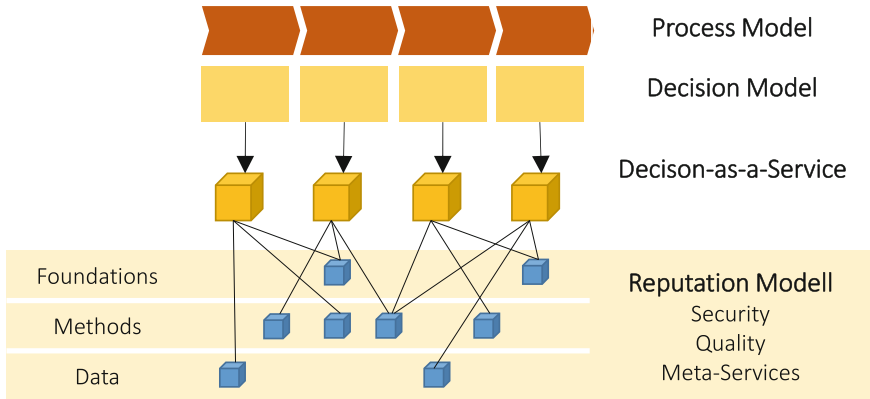
Fig. 1. Layered Decision as a Service

With Decision-as-a-Service models, data from different sources are mixed together to come to a decision. Hence, each data source and each contributor of a data source has the possibility to influence the final decision. In addition, it is expected that the proposed architecture is open and dynamic, i.e. data sources can be added, and contributors are not known in advance. Sources are very likely dynamic, meaning that new content is constantly added to the data source and those new contributors appear and more mature contributors do not contribute anymore to a data source. A Reputation Modell allows to evaluate the Security, Quality and the Meta-Services [14] of services providing foundation, methods or data.

3   Reputation System

Reputation systems are quite common in dynamic and open systems that allow interaction between before unknown parties. A reputation score helps humans to decide whom they will trust. Such a score also allows for automated decisions by machines. One popular example of a reputation system for human interaction is the eBay reputation system (see [15] for a thorough discussion). In [16], a current overview of reputation systems is given. One can notice in [16], that current reputation systems usually only focus on one single domain, hence are limited in their use. The Decision-as-a-Service architecture proposed in this paper however, uses data sources from multiple domains, hence calls for a reputation system that is not limited to one single domain.

For a Decision-as-a-Service provider, it is of crucial importance that decisions are reliable for the customers of the provider. The reliability of a decision can be negatively affected by the following actions:

- Attack 1: A contributor of a data source provides fake data.
- Attack 2: Contributors of a data source cooperate to provide faked data and to discredit other contributors (e.g. a large number of paid reviews at a hotel review site).

- Attack 3: An adversary influences the selection of data sources a Decision-as-a-Service service is based on.
- Attack 4: An adversary manipulates data of a data source, either during storage or during transfer of data.

To avoid the last attack of this list, authenticity and integrity of all data must be guaranteed during transfer and storage. Several methods are available to achieve authenticity and integrity, e.g. secure communication protocols like HTTP over TLS (https, see [17] for details). A correct implementation (including certificate validity checking) assumed, these solution works well and have the potential to avoid the last attack of the list.

Avoiding the first three attacks of the list above is harder to achieve. This paper presents a reputation system that could be used to protect against these attacks. The reputation system imposes a goodness score g(contributor) on each contributor, judging on the trustworthiness of the contributor. For all contributors, it is defined that

$$0 \le g(contributor) \le 1$$

where a goodness store of 0 denotes a totally untrustable contributor whereas a goodness score of 1 means that a contributor is fully trusted.

Using the goodness scores of all contributors of a single data source (one domain), a goodness score for the data source itself can be calculated, e.g. by using a weighted sum of the goodness sources of all n contributors of this data source:

$$g(source) = \sum_{i=1}^{n} (w_i * contributor_i)$$

The weight w_i of a $contributor_i$ expresses the degree a contributor contributes to the data of a data source. For example, the weight could be the number of reviews a contributor has written on a hotel review site. The weights wi should be selected such that

$$\sum_{i=1}^{m} w_i = 1$$

In the hotel review site example, m is the total number of reviews and the w_i for a $contributor_i$ is calculated from the number of reviews (nor) this contributor has written:

$$w_i = \frac{nor}{m}$$

With these definitions, it is ensured that

$$0 \le g(source) \le 1$$

The goodness score of a data source may be enriched using other information about a data source. For example, the identifiability of contributors plays an important role in our model, as it is otherwise not possible to use individual goodness scores to calculate the overall score of a data source. Another aspect is the authentication method used by the data source to authenticate contributors. If the authentication method is weak, it may be easy for an attacker to impersonate other contributors and insert fake data. Yet another aspect is the ease of creating new contributors in a data source. If it is easy for an attacker to create a huge number of new contributors and then only once contribute data to the data source for each created contributor, an attacker can spoil the data of the data source. A quality factor is introduced to address the issues described in this paragraph:

$$g(source) = quality * \sum_{i=1}^{n}(w_i * g(contributor_i))$$

Where

$$quality = \sum_{i=1}^{n}(u_i * qa_i).$$

In this equation, qa_i is one quality aspect that applies for this type of data source and u_i is a weight factor that reflects the importance of the associated quality aspect.

With

$$qa_i \in [0, 1]$$

$$u_i \in [0, 1]$$

$$\sum_{i=1}^{n} u_i = 1$$

it is ensured that

$$0 \leq quality \leq 1,$$

hence

$$0 \leq g(source) \leq 1$$

The knowledge of goodness scores of data sources helps Decision-as-a-Service services to select appropriate data sources for the decision process. Rules may define a lower boundary for the goodness score of data sources to be used. In addition, the goodness score could be used as primary input for a confidence score of a decision. The calculation of the goodness score for a contributor is the heart of the reputation system. All other scores rely on this score. First, it is necessary to have a way to identify

reliably the contributors of a single data source. IDs for identification may for example be published (verified) email addresses. If no IDs are available, the only way to use this data source is to assume it is a data source with only one contributor. However, in this case it is very likely that the associated goodness score will be only average or lower.

The goodness score of a contributor should take into consideration that a contributor may appear in more than one data source (domain). For example, a contributor rating hotels on a hotel review website may also rate restaurants on a restaurant review website. A reputation system that works on a potentially unlimited number of domains (data sources) is a significant advance to the state of the art of reputation systems.

To address the problem of multiple data sources, it is necessary, to identify the appearance of one single contributor in multiple data sources. Each data source may use a different ID for this single contributor, for example, one site may use the email address, another site may use a self-selected username (distinct from email addresses), yet another side may use the full name of users. To address this issue, the proposed reputation system associates a randomly chosen GUID (Global Unique ID) with one or more IDs of contributors. For each ID, the confidence that this ID really belongs to the GUID is also stored. This is important in the case of ambiguities. For example, there may be more than one person named "Hans Maier" as the first name Hans and the last name Maier are very common in Germany. In addition, depending of the authentication method and the validation method at registration, it may be possible to impersonate a contributor. The confidence about an association may change over time, hence the goodness score of a contributor may be very dynamic over time. Finding a good and secure way of associating different IDs will be an important part of the proposed reputation system.

Criteria for judgment of the goodness of a contributor are specific for each data source. These criteria may include

- Existing reputation mechanisms from data sources, e.g. user voting systems, user rating systems, …
- Past behavior of a user.
- Characteristics that are appropriate to detect fake content (e.g. typical wording of paid advertisments).
- …

The goodness score of a contributor is calculated using the n data sources a user appears in with the best confidence score

$$g(contributor) = \frac{\sum_{i=1}^{n}(confidence_i * g(contributor@source_i))}{n}.$$

The goodness score of a contributor in one single data source is

$$g(contributor@source_i) = \sum_{i=1}^{n}(v_i * gc_i),$$

where gc_i is one criteria for goodness, and v_i reflects the importance of this goodness criteria.

With

$$qc_i \in [0, 1]$$

$$v_i \in [0, 1]$$

$$\sum_{i=1}^{n} u_i = 1$$

it is ensured that

$$0 \leq g(contributor@source_i) \leq 1,$$

hence

$$0 \leq g(contributor) \leq 1.$$

The same semantic (0 = not trustworthy, 1 = fully trustworthy) as above applies.

4 Using Decision as a Service in Digital Transformation

Decision as a service in general and the reputation system in particular are critical elements for the digital transformation of enterprises. Digital transformation is defined according to [4] as "the use of technology to radically improve performance or reach of enterprises" and affects the areas customer experience, operational processes and business models. (see also [18, 19]).

Four digital capabilities build the foundation layer for the improvements in these areas and encompass unified processes and data, analytics capability, business and IT integration and solution delivery [4]. While the last two capabilities are more concerned with the realization of effective, solutions for the digital transformation are the first two capabilities directly linked to decision as a service and the reputation system. Unified processes and data are especially difficult to archive in an age of massive volatile processes and unstructured data as described earlier and additionally, it is required to include these processes and unstructured data as sources into the business analytics. The recommended layer architecture for Decision as a service helps to overcome these difficulties.

In the following some potential application of Decision as a service and the reputation system in the areas customer experience, operational processes, and business models will be described.

The area of customer experience provides several opportunities for application.

- Decision as a service can support the customer segmentation by providing relevant information, e.g. from even divers social media contents, to define better customer segments
- In the digital selling process decisions are constantly required to provide customers with individualized marketing information and specific offers with better quality due

to the usage of the reputation system. Additionally, Decision as a service can also support a much leaner, streamlined customer process by eliminating non-relevant alternatives, like in service add-ons, payment options. This also applies to service processes with the customers or retailers.

- Cross-channel coherence can also profit from Decision as a service by ensuring that the marketing content and offers are consistent across the channels.

The area of operational processes is a natural use-case for Decision as a service.

- Process improvements require often decisions within certain activities, but also the decision, which process variant is applicable to a specific case (e.g. case based routing). Decision as a service helps here to include not just company internal, but also external sources with high data quality.
- As performance management needs higher operational transparency and the decision-making process will more and more on data-drive the same reasoning also applies to this building block of digital transformation

Also in the area of business model for Decision as a service can help to support the digital transformation.

- Digital modified business in the form of product/service augmentation or digital wrapper to existing products rely heavily on larger and more divers data sources and high data quality about products, services, customer and other partners to provide the right solution in time. Decision as a service help to fulfill this needs
- New digital business is by definition linked to digital services and with that Decision as a service can also become an import part on new digital products. On example might be trustworthy recommendation in share economy applications.
- With the digital globalization Decision as a service helps to foster a global network and build transparency even for quite distant partners.

These applications show how the recommended layer architecture for Decision as a service helps to successfully digital transform an existing enterprise.

5   Related Work

A more strategic view on Decision as a Service as means for Transformation is discussed in [13]. The positive effects of data-driven decision making on firm performance are analyzed in [20]. Also in [21] the positive effects of analytics are discussed. In [22] the basic concept of Decision as a Service is introduced.

A general introduction into reputation systems is given in [16]. The role of trust in the internet is analyzed in [15]. The first use of reputation as metric for data quality has been defined in [23]. The use of reputation in peer to peer networks is suggested in [24]. The positive effects on separating different aspects of business processes are discussed in [25].

6   Conclusion

Many enterprises and organizations improve their decisions making processes by automation and the integration of more data sources. Although this approach increases the velocity and precision of decision-making, it also creates new risks, because it implies to integrate many more external data sources as before. This creates new challenges to information system security. External data could be forged and leading to false decision. Therefore, a reputation system for Decision-as-a-Service is presented. The reputation system provides a score for the trustworthiness of a data source used for Decision-as-Service. The reputation system is not limited to one domain but can be used on a potentially unlimited number of domains, hence offers a high degree of flexibility for the Decision-as-a-Service service. The output of the reputation system may be used for selection of appropriate data sources as well as for the calculation of a confidence score for the decision of a Decision-as-a-Service service. Future work will have to further detail our reputation model and system architecture.

References

1. Andersson, H., Tuddenham, P.: Reinventing IT to Support Digitization. McKinsey, New York (2014)
2. Markovitch, S., Willmott, P.: Accelerating the Digitization of Business Processes. McKinsey & Company, New York (2014)
3. Manyika, J., Chui, M., Bisson, P., Woetzel, J., Dobbs, R., Bughin, J., Aharon, D.: The Internet of Things: Mapping the Value Beyond the Hype. McKinsey & Company, San Francisco (2015)
4. Capgemini Consulting, MIT Sloan Managment: Digital Transformation: A Road-Map for Billion-Dollar Organizations (2011). http://www.capgemini.com/resources/digital-transformation-a-roadmap-for-billiondollar-organizations
5. Weill, P., Woerner, S.: Thriving in an increasingly digital ecosystem. MIT Sloan Manage. Rev. **56**(4), 26–34 (2015)
6. Fortune, Inc.: Fortune 500 firms in 1955 vs. 2014; 89 % are gone, and we're all better off because of that dynamic "creative destruction," http://www.aei.org/publication/fortune-500-firms-in-1955-vs-2014-89-are-gone-and-were-all-better-off-because-of-that-dynamic-creative-destruction/
7. Locker, M.: 8 iconic brands that have disappeared – Fortune. http://fortune.com/2014/11/09/defunct-brands/
8. Fitzgerald, M., Kruschwitz, N., Bonnet, D., Welch, M.: Embracing digital technology: a new strategic imperative. MIT Sloan Manage. Rev. **55**, 1–12 (2013)
9. Schmidt, R., Nurcan, S.: BPM and social software. In: Ardagna, D., Mecella, M., Yang, J., Aalst, W., Mylopoulos, J., Rosemann, M., Shaw, M.J., Szyperski, C. (eds.) Business Process Management Workshops, pp. 649–658. Springer, Heidelberg (2009)
10. LaValle, S., Lesser, E., Shockley, R., Hopkins, M.S., Kruschwitz, N.: Big data, analytics and the path from insights to value. MIT Sloan Manage. Rev. **52**, 21–32 (2011)
11. Schmidt, R., Möhring, M., Maier, S., Pietsch, J., Härting, R.-C.: Big data as strategic enabler - insights from central European enterprises. In: Abramowicz, W., Kokkinaki, A. (eds.) BIS 2014. LNBIP, vol. 176, pp. 50–60. Springer, Heidelberg (2014)

12. White, T.: Hadoop: The Definitive Guide. O'Reilly Media, Sebastopol (2012)
13. Schmidt, R., Möhring, M., Koot, C.: Data-centered cloud-environments as enabler for decision as a service in small and medium enterprises. In: 2013 NAPLES Forum on Service, Ischia, Napoli (2013)
14. Schmidt, R.: Meta-services as third dimension of service-oriented enterprise architecture. In: Presented at the 2010 14th IEEE International Enterprise Distributed Object Computing Conference Workshops (EDOCW) (2010)
15. Resnick, P., Zeckhauser, R.: Trust among strangers in internet transactions: empirical analysis of ebay's reputation system. Econ. Internet E-commerce 11, 23–25 (2002)
16. Hendrikx, F., Bubendorfer, K., Chard, R.: Reputation systems: a survey and taxonomy. J. Parallel Distrib. Comput. 75, 184–197 (2015)
17. Rescorla, E.: Http Over TLS (2000)
18. Fichman, R.G., Dos Santos, B.L., Zheng, Z.: (Eric): digital innovation as a fundamental and powerful concept in the information systems curriculum. MIS Q. 38(2), 329–343 (2014)
19. Safrudin, N., Rosemann, M., Recker, J.C., Genrich, M.: A typology of business transformations. 360° Bus. Transform. J. 2014, 24–41 (2014)
20. Brynjolfsson, E., Hitt, L., Kim, H.: Strength in numbers: how does data-driven decisionmaking affect firm performance? (2011)
21. Davenport, T.H., Harris, J.G., Morison, R.: Analytics at Work: Smarter Decisions, Better Results. Harvard Business Press, Cambridge (2010)
22. Delen, D., Demirkan, H.: Data, information and analytics as services. Decis. Support Syst. 55, 359–363 (2013)
23. Strong, D.M., Lee, Y.W., Wang, R.Y.: Data quality in context. Commun. ACM 40, 103–110 (1997)
24. Oliver, H., Daanen, J.: Management of peer-to-peer networks using reputation data. Google Patents (2004)
25. Regev, G., Soffer, P., Schmidt, R.: Taxonomy of flexibility in business processes. In: Proceedings Seventh Workshop on Business Process Modeling, Development, and Support (BPMDS 2006), Requirements for Flexibility and the Ways to Achieve It, p. S. 90–93. Luxemburg (2006)