

Association for Information Systems

AIS Electronic Library (AISeL)

AMCIS 2022 Proceedings

SIG DSA - Data Science and Analytics for
Decision Support

Aug 10th, 12:00 AM

Transparency as a Potential Factor for Implementation of Machine Learning-based Systems

Olga Levina

Brandenburg University of Applied Sciences, levina@th-brandenburg.de

Follow this and additional works at: <https://aisel.aisnet.org/amcis2022>

Recommended Citation

Levina, Olga, "Transparency as a Potential Factor for Implementation of Machine Learning-based Systems" (2022). *AMCIS 2022 Proceedings*. 3.

https://aisel.aisnet.org/amcis2022/sig_dsa/sig_dsa/3

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in AMCIS 2022 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Transparency As A Potential Factor For Implementation of Machine Learning-Based Systems

Emergent Research Forum (ERF)

Olga Levina

Brandenburg University of Applied Sciences

levina@th-brandenburg.de

Abstract

Legal regulations such as European General Data Protection Regulation have put transparency as an essential characteristic of data-based information systems firmly into the focus of the system developers. Furthermore, enterprises increasingly consider challenging machine learning-based systems (MLS) to support particular business functions and hence require informed decisions based on explainable results of the MLS. This research in progress explores the information needs for the adaptation of predictive analytics MLS in a commercial setting. Using scenario-based survey and based on the technology acceptance model, insights into the transparency requirements from the users' perspective are derived and future research directions are outlined. This empirical investigation provides suggestions for potential factors influencing the adaptation of MLS in business context and supports design decisions for MLS. Future directions for the exploration of the role of transparency requirements for the acceptance of machine learning-based information systems are outlined.

Keywords

Machine learning based systems, transparency, technology acceptance model, empirical study, ethical implications.

Introduction

Transparency as a requirement for a business information system (BIS) is typically linked to positive ethical attributes such as trust (Mahmood and Moghaddam, 2016). Especially in the context of machine-learning based systems (MLS), increasingly known under the term of artificial intelligence (AI), transparency has received a lot of attention following the definition of trust-worthy AI by the European Commission (European Commission, 2020). Transparency is defined here based on (Mahmood and Moghaddam, 2016) as the "open flow of high quality information in a meaningful and useful manner among stakeholders" of a BIS. Also, opacity, i.e. the lack of explanation, about the origin of the results of an MLS is a point of critique for researchers (Milano *et al.* 2019), users and software developers (Baxter and Sommerville, 2011) alike. On the other hand, introducing this information flow among the BIS stakeholders can also lead to aversive effects such as information overload (Felzmann *et al.*, 2019).

This research in progress explores this position of transparency between information requirements and potential information overload. The guiding research question here is: What influence does transparency have on the intention to use an MLS in a business context? To approach this question, the notion of transparency is decomposed into process, system and economic information needs. As a research method, a scenario-based online survey was launched to allow participants to act in the business context according to their transparency needs and legal obligations based on the GDPR. The scenario outlined an MLS from the predictive analytics domain as a BIS. The scenario was built using the concept of socio-technical information systems design and based on the propositions of the technology acceptance model (TAM) (Davis, 1989; Venkatesh and Davis, 2000). Transparency as information on the MLS design was introduced as an influencing factor on the perceived usefulness, while other influencing factors such as

job relevance, result demonstrability as well as subjective norm were addressed in the scenario description.

The first results of the scenario-based survey presented here provide several starting points for further exploration of the transparency aspects and their effect on adoption of MLS in business context. Furthermore, they deliver some points of contact for developers of MLS on how to implement transparency requirements into the system design, for BIS managers looking to implement MLS into the business processes as well as for information systems researchers exploring the acceptance and usage of the upcoming technologies. The paper is organized as follows: First a short overview of the research on transparency as a factor in the BIS domain is provided. The research method used here and the results of the scenario-based survey are then described. Discussion of the results and further research directions finish the paper.

State of the art on transparency and MLS adoption

From the legal perspective, the Art. 5 of GDPR lays down the universal principles of the data processing: “Personal data shall be processed lawfully, fairly and in a transparent manner in relation to the data subject” (Art. 5.1. GDPR). The main scope of the transparency requirements is defined by the information obligations in Art. 12 and Art. 13 GDPR. Art. 12 GDPR initially states that all information and communications must be provided in a precise, transparent, comprehensible and easily accessible form in clear and simple language. However, *transparency* also requires explainability of the decision process in algorithmic decision procedures. It is essential from the GDPR perspective that users of algorithmic systems are as a minimum able to understand, explain and control their operation, and that those affected receive sufficient information to enable them to exercise their rights in Art 12 – 23 GDPR properly. (Levina and Mattern, 2022). Additionally, transparency is an often discussed factor in the context of recommender or advice giving systems (AGS). In the research scenario in this paper an AGS is selected as an MLS example. AGS’s transparency is defined by (Zhao, Benbasat and Cavusoglu, 2019) as “users’ understanding of systems inner logic”. (Mahmood and Moghaddam, 2016) provide a broader definition of transparency as the “open flow of high quality information in a meaningful and useful manner among stakeholders” that is extended here for the BIS domain. This research in progress focuses on the objective transparency in the input stage of the model calculations as defined by (Zhao, Benbasat and Cavusoglu, 2019). It describes the extent to which information is provided by the system designers about what they do and what data they use or used to train their model. Technology acceptance model (TAM) and TAM2 by (Davis, 1989) and (Venkatesh and Davis, 2000) respectively provide the theoretical fundament for the research design. TAM and TAM2 are frequently used and also modified to explore the adoption and acceptance of new consumer technologies (Salloum *et al.*, 2021), and machine learning (Go, Kang and Suh, 2020; Alwabel and Zeng, 2021). Here, transparency, i.e. its operationalization as information requirements, was included as an influencing factor for the perceived usefulness and thus was considered as potentially influencing the intention to use the BIS. Transparency as part of explainability of machine learning algorithms and MLS is also increasingly gaining attention from research, e.g. (Barredo Arrieta *et al.*, 2020; Meske *et al.*, 2020), governmental institutions (Ridley, 2019), standardization organizations (IEEE Computer Society, 2021), as well as industries (Chander *et al.*, 2018; Nabbosa and Kaar, 2020).

Research Method

To answer the research question on the effect of objective transparency on collected data for the acceptance of MLS, an online scenario-based survey was launched. The scenario was built around the use of a fictive BIS that involves predictive analytics techniques (MLS).

Scenario description

The survey participant was put into a role of the owner of a newly founded café “Green Lawn”. The café owner in the scenario has changed her career to open the café in a small town, including accompanying substantial financial expenses. Nevertheless, the local municipality would grant a free access to a MLS-based predictive analytics software, called GuestInTime, for business users for a trial period of two weeks. The developers of GuestInTime, a fictional company named RealGuest, claim the software to be able to make correct predictions about the amount of guests expected during the day and on the following day.

No further description of GuestInTime is given in this stage of the scenario. After this introduction, the participants were offered different options of how to proceed on behalf of the owner of the “Green Lawn”, choosing between possible scenario paths and outcomes. In the first selection step, participants, in the role of the owner of “Green Lawn”, were asked whether they would like to try GuestInTime in their business for two weeks without charge.

Decision 1: Participation in the trial period of GuestInTime. No further description of the MLS was given beside the introduction to the scenario.

- Option 1.1: Yes. Participation offer is accepted.
- Option 1.2: No. The café owner is not using GuestInTime.

For the participants choosing option 1.1., decision 2 needed to be made after the trial period ended with good performance from the GuestInTime software. The decision was concerned future engagement with GuestInTime based on its good prediction performance. Following options for the scenario paths were given:

- Option 2.1: Buy the software, because it delivered good results.
- Option 2.2: Get more information about how the results are calculated, as the results may be due to chance, and decide about the acquisition later.
- Option 2.3: Not buying the software, since the owner could have predicted the number of guests herself.
- Option 2.4: Not buying the software, because its structure is unclear and its understanding would take up time that could be spent by building a customer base.

Decision 3 was now to be taken accordingly to the selected options. If the option 2.2 was selected, the scenario led to a meeting between the café owner and the developers of GuestInTime to address the information needs of the owner. In this meeting, the developers explained that they used publicly available weather data of the region in their prediction model, and identified potential guests of “Green Lawn” via their location data on the SuperMaps, another fictional application invented for the scenario. The Social Media profiles of the users that were located in the vicinity of “Green Lawn” were then screened to calculate the probability of them to visit the coffee shop. These metrics were used as inputs for the final prediction by the GuestInTime model. The user tracking option is also described in the scenario to be part of the terms of use of SuperMaps, as required by the GDPR. After this explanation, RealGuest company makes a licensing offer to “Green Lawn” for the next two years that requests for 1.5% of “Green Lawn”’s sales.

Now, the participants that met with the developers were given the following options for the subsequential scenario path, i.e. decision 3.

- Option 3.1 Buy the software.
- Option 3.2 Buy the software, despite an ethically questionable, but legal data collection.
- Option 3.3: Do not buy the software, because more information is needed.
- Option 3.4: Do not buy the software and rely on proper expertise.

This was the final decision on the proceeding of the scenario. Depending on the chosen paths and their combination, participants were granted an assertion of a successful business. Selection of alternative paths led to a scenario outcome where a public backlash that was based on the newspaper article about the data used for GuestInTime prediction calculation and that had negative business consequences for “Green Lawn” was depicted.

Operationalizing transparency as an influencing factor for perceived usefulness

In the scenario the usefulness of the BIS is outlined by describing the potential business dependencies such as: responsibilities towards the employees, business success as well as information about the MLS design. Hence, transparency, was operationalized as an influencing factor for the perceived usefulness and as such was considered as potentially influencing the intention to use the software according to TAM. The perceived usefulness of this factor was measured via the selection of the proceeding option in the scenario. Options 2.2, 2.4. and 3.3 in the scenario description directly inquire about the informational input requirements concerning the MLS as the factor in the decision making. Hence, the aspect of transparency as information need was operationalized on three levels in the scenario description:

Economic level: Description of the need for support of business processes, since the café was started as a new career choice; the economic responsibility of the café owner for the employees, as well as the positive results of the MLS on café's economic performance indicators, such as exact planning based on good predictability. Financial involvement after the end of the free trial period was outlined as a license plan for the use of the MLS. Described in the scenario preceding decision 1 and measured by the number of participants opting for the participating option, i.e. options 1.1 and 2.1 described in section 3.1.

System level: It was outlined in the scenario that the predictive MLS would run on any mobile device and would not require any additional communication architecture or input data. No description of potential for feedback or explanation of the MLS results was provided. Described in the scenario preceding decision 2 and 3 and measured by the number of participants opting for the buying option, i.e. options 3.1 and 3.2.

Process level: Scenario described the potential use of the MLS in business processes and value creation, i.e. more precise planning of staff, planning of provisions as well as food waste reduction. Described in the scenarios preceding decision 1 and 2 and measured by the number of participants opting for the buying option or the option to gather more information on the algorithm behind the system results, i.e. options 1.1, 2.1 and 2.2.

The research method used here is a scenario-based survey that was distributed online among bachelor and master students of business administration, information systems management and law. The survey started with a description of a scenario of a business start-up, a café, allowing the participants to act in the business context according to their transparency and business needs. 74 people participated in this online survey. Since the fictive MLS was implemented in a scenario-environment that might have required personal data collection, legislation of the European Union was used as the legal framework for the scenario.

Results overview of the scenario-based survey

With N=74 participants only a descriptive statistical analysis of the data is performed. Nevertheless, the results allow some insights that can be used in further research directions and experimental setups.

Given the decision 1 on whether to try the MLS in their business workflow or not 63 of 74 participants, i.e. almost 85%, decided to participate in the free trial. 11 participants decided against the usage of MLS in their business context. 9 participants, i.e. almost 15%, decided to buy GuestInTime right after the trial period without any further information on its algorithms or collected data, choosing option 2.1. 50 participants, i.e. close to 81%, decided to gather more information on the workings of the MLS and thus meeting with the developer team of GuestInTime. One participant declined buying the software based on option 2.4, i.e. not wanting to spend time exploring the data related questions. After the meeting with the developer team 28 participants, i.e. 56%, decided to buy the software (option 3.1 and 3.2), while 15, i.e. 30% percent, were still hesitant and requesting more information (option 3.3). 7 participants, i.e. 14% of those participants who requested additional information, decided not to buy the software after the meeting with the developer team (option 3.4). As a final observation, 37 participants, 50%, would have bought and implemented GuestInTime in the described scenario, with the majority of them (24 participants) deciding on the purchase after gaining additional information on the data used for the prediction confirmation. As suggested by (Levina, 2021), this purchasing decision might have been supported by the fact, that the data used for the prediction model was conformant with the legislation.

Discussion and Outlook

This research in progress used a scenario-based survey to simulate the inclusion of transparency on input data of the MLS as an influencing factor for TAM's perceived usefulness. Transparency was operationalized as information needs on the levels of: economic, system and process. First insights that were gained using descriptive analysis of the survey results indicated multiple further research directions. As such, the majority of the survey participants was willing to try the MLS to help fulfill their job in a foreign realm without requesting any additional information on its mode of operation. This might be interpreted as a positive influence of technical curiosity on the technology acceptance. An aspect that will be included in further analysis. On the other hand, the information on the workings of the MLS (the operationalization of transparency) resulted in over 40% of participants (22) rejecting the commercial use

of the supporting MLS (options 3.3. and 3.4), indicating potential disagreement with the legal but ethically questionable data collecting practices that were involved in the calculation of the MLS prediction. Based on this indication, further empirical analysis of potential ethical questions related to information needs and acceptance of MLS will be performed. These insights provide several starting points for further research and also some indicators for the developers and designers of MLS about the information and transparency needs of its users. Addressing these needs directly in MLS design and description may help to avoid information overload or deceptive expectations towards results and business benefits of the MLS.

REFERENCES

- Alwabel, A. S. A. and Zeng, X. J. (2021) 'Data-driven modeling of technology acceptance: A machine learning perspective', *Expert Systems with Applications*, 185, p. 1155.
- Barredo Arrieta, A. *et al.* (2020) 'Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI', *Information Fusion*, 58, pp.82–115.
- Baxter, G. and Sommerville, I. (2011) 'Socio-technical systems: From design methods to systems engineering', *Interacting with Computers*, 23(1), pp. 4–17.
- Chander, A. *et al.* (2018) 'Working with beliefs: AI transparency in the enterprise', in *CEUR Workshop Proceedings*.
- Davis, F. D. (1989) 'Perceived usefulness, perceived ease of use, and user acceptance of information technology', *MIS Quarterly: Management Information Systems*, 13(3), pp. 319–339.
- European Commission (2020) *Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment | Shaping Europe's digital future*. Available at: <https://ec.europa.eu/digital-single-market/en/news/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment> (Accessed: 7 February 2021).
- Felzmann, H. *et al.* (2019) 'Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns', *Big Data & Society*, 6(1).
- Go, H., Kang, M. and Suh, S. B. C. (2020) 'Machine learning of robots in tourism and hospitality: interactive technology acceptance model (iTAM) – cutting edge', *Tourism Review*, 75(4)
- IEEE Computer Society (2021) 'IEEE 7000-2021: IEEE approved draft model process for addressing ethical concerns during system design'. Available at: <https://standards.ieee.org/standard/7000-2021.html>
- Karpati, D., Najjar, A. and Agustin Ambrossio, D. (2020) 'Ethics of Food Recommender Applications'
- Levina, O. (2021) 'Transparency of Data Origin and Business Value as Factors for the Assessment of Opportunities and Risks of Machine Learning- based Systems', in *AI4EU Workshop: The Culture of Trustworthy AI. Public debate, education, practical learning*. Venice.
- Levina, O. and Mattern, S. (2022) 'Ethical and legal analysis of machine learning based systems: a scenario analysis of a recommender system', in *Symposium of Recommender Systes*. Köln, p. in print.
- Mahmood, S. and Moghaddam, H. (2016) 'Engineering of Transparency Requirements in Business Information Systems', *Computer Science*.
- Meske, C. *et al.* (2020) 'Explainable Artificial Intelligence: Objectives, Stakeholders, and Future Research Opportunities', *Information Systems Management*.
- Milano, S., Taddeo, M. and Floridi, L. (2019) 'Recommender Systems and their Ethical Challenges', *Minds and Machines*, 2, pp. 187–191.
- Nabbosa, V. and Kaar, C. (2020) 'Societal and Ethical Issues of Digitalization', *ACM International Conference Proceeding Series*, pp. 118–124.
- Ridley, M. (2019) 'Explainable Artificial Intelligence', *Research Library Issues*, (299), pp. 28–46.
- Salloum, S. A. *et al.* (2021) 'Predicting the Intention to Use Social Media Sites: A Hybrid SEM - Machine Learning Approach', *Advances in Intelligent Systems and Computing*, 1339, pp. 324–334.
- Venkatesh, V. and Davis, F. D. (2000) 'A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies', *Management Science*, 46(2), pp. 186–204.
- Zhao, R., Benbasat, I. and Cavusoglu, H. (2019) 'Do users always want to know more? Investigating the relationship between system transparency and user's trust in advise-giving systems', *Proceedings of the 27th European Conference on Information Systems*, pp. 1–13.