Measuring the Influence of Characteristics on Decision-Making Scenarios: A Prototype

Sebastian Reiners[1](✉), Lea Sophie Müller[2], Jörg Becker[1], and Guido Hertel[2]

[1] University of Münster, ERCIS, 48143 Münster, Germany
{sebastian.reiners,joerg.becker}@uni-muenster.de
[2] Organizational and Business Psychology, University of Münster, 48143 Münster, Germany
{lea.mueller,guido.hertel}@uni-muenster.de

Abstract. Nowadays, the world is digitalized and data-driven, transforming decision-making as information is available to essentially any person at practically any time. A naive approach would at first propose that with a greater amount of information, i.e., a greater amount of data, a more educated choice can be reached. Rather, many decision scenarios reach a point today where individuals can no longer acquire and comprehend the vast volume of data on their own. As a corollary, people seek the assistance of algorithms and machines to make reasoned and informed decisions when confronted with difficult issues or problems. However, current research lacks a perspective that can recognize and empirically explore the complex multidimensional nature of the relationships between the defining characteristics of such scenarios. We bridge this void by investigating how the influence of these relationships affects human decision making and perception of an artificial intelligence/system. Specifically, we propose a web-based prototype to simulate decision-making scenarios in which it is possible to change certain characteristics (e.g., transparency, control) and analyze how the change affects the person's decisions and perceptions (e.g., trust) toward the system.

Keywords: Decision making · Characteristics · Artificial intelligence · Prototype

1 Introduction

From a analytic, i.e., theoretical, standpoint, the decision-making process can be summarized in a four-stage model: 1) sensory processing, 2) perception, 3) decision-making, and 4) response selection [1]. These decisions can range from the nugatory ("Do I brush my teeth with my right or left hand?"), over the routine ("How many cups of coffee should I drink today?"), to the strenuous ("Which of departments are to be eliminated and thus their employees laid off?").

Today's digitized and data-driven world is transforming our decision-making, with information available to virtually anyone at any time. A naive view would initially suggest that a better justified and facilitated decision is vacant with more information being available, i.e., a more lavish amount of data. Instead, many decision scenarios today

C. Stephanidis et al. (Eds.): HCII 2022, CCIS 1580, pp. 387–392, 2022.
https://doi.org/10.1007/978-3-031-06417-3_52

reach a point where humans can no longer comprehend and process the vast quantity of data on their own [2].

Consequently, when faced with complex questions or assignments, individuals seek the help of algorithms and machines to make informed decisions. However, the often deterministic nature of algorithms requires users to have an impeccable judgment of which propositions to trust [3, 4]. The decision-making process is further revolutionized with artificial intelligence leaving the human entirely out of the final decision in some scenarios. New emerging challenges revolve around evaluating the role and perception of humans in digitized decision-making scenarios [5]. One of the unique (technical) challenges of such systems is the tangibility of the algorithms. Characteristics such as transparency or the ability to explain the decisions made are an increasingly important progression in overcoming information asymmetries [6].

However, current research lacks a perspective that recognizes and empirically examines the complex multidimensional relationships among the defining traits of such scenarios. We seek to bridge this gap by exploring how tampering with those relationships impact human decision-making and the perception of an artificial intelligence/system.

More specifically, we propose a web-based prototype to simulate decision-making scenarios in which it is possible to tweak certain characteristics (e.g., transparency, control) and analyze how the alteration shapes the person's decisions and perceptions (e.g., trust) toward the system [7]. This prototype makes a novel contribution to investigating the complex multidimensional structure of characteristics in human-machine collaboration scenarios by providing a tool capable of empirically testing these scenarios.

2 Artifact Design

2.1 Technology Stack and Database Design

For the prototype's architecture, we adhered to proven designs by other researchers (e.g., [8]) and experience from previously realized software projects. In the implementation process, we use well-established practices of agile software development. Following the agile software development values, we would like to emphasize that potential modifications to licensing models or technical innovations during the implementation might change the intended architecture (cf. Fig. 1). In general, the architecture has been constructed so that long-term compatibility of the used components is probable. Concurrently, there is the potential for enhancements with supplementary features.

The prototype's core, i.e., the developed application, is implemented using the Python web framework Django. The Django framework has a large community on both the user and developer ends. Simultaneously, the comprehensive documentation and modular architecture also meet our required standards of a web framework. As a web server gateway interface responsible for running the prototype's application, we selected Gunicorn. Advantages are platform independence, integration with Django, and scalability with possible extensions. Nginx, an open-source software under BSD license with modular structure, is used as a web server. It enables encrypted connections and is widely used in conjunction with Django applications. For the database, we selected PostgreSQL as it integrates with the Django application by providing an interface to Python: It is also runnable on all conventional operating systems and easy to adapt to future versions of the application. The application is presented to the client via a secure HTTPS request on any web app independent of the client's operating system.

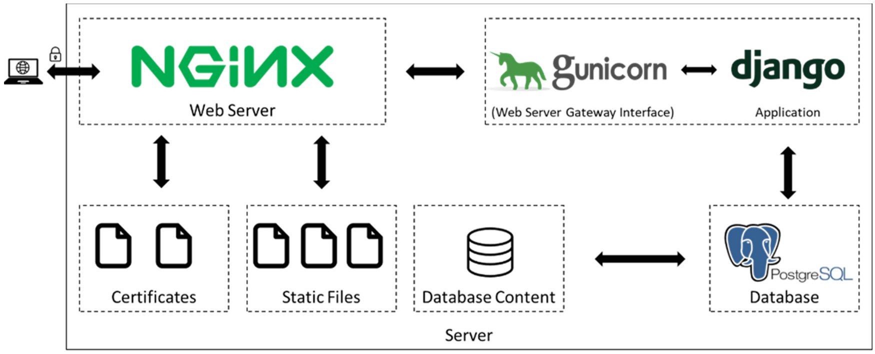

Fig. 1. Architecture

The data structure of the prototype is derived from a data model, which specifies the crucial elements of the application. The holistic database design is done in an Entity Relationship Model (ERM) (cf. Fig. 2) featuring entities as rectangles, relationships as diamonds, and the cardinalities between entities, i.e., the numerical correlation of rows of one with rows of the other table, as pairs of numbers in square brackets.

At the center of the model is the scenario. A scenario can only be created by one manager but altered by other managers. Managers also are responsible for creating characteristics that are used in the scenarios. Each scenario consists of at least one characteristic but can have an arbitrary number of characteristics. A participant can complete a scenario at a given time. A given completion is associated with measurements in which the manipulation of characteristics, performance, and participant perception are saved. In the ERM, this is displayed by re-interpreting the relationship between participant, time, and scenario as a new entity type.

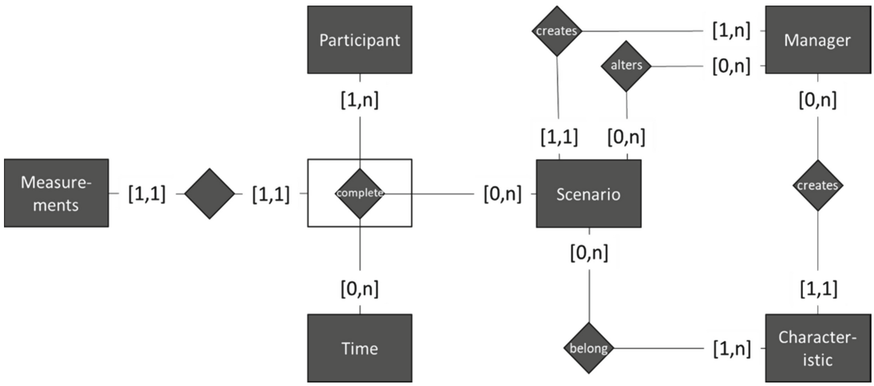

Fig. 2. ERM model

2.2 Features and User Journey

The main features of the prototype are centered around the scenario, characteristics, and the relationship between the aforementioned. In this, a scenario and characteristics creator are generated. We, furthermore, evolve the user interface from two angles. The scientists or the people referred to as managers in ERM, are on one side. They have access to the scenario creator and the characteristics creator. On the other side are the visitors of the web prototype or participants in the ERM.

The user journey (cf. Fig. 3) represents, although on a high level, those steps a participant would encounter when completing a scenario with the web-based prototype. First, the participant is greeted by a welcome screen featuring a button to start the study and two tabs with additional information about the prototype and the project positioned at the top bar. When clicking the start button, general information about the study ahead and the possibility to select specific scenarios are shown. There will also be the possibility to receive a random scenario. In a first questionnaire, "Inquire well-being", we ask for the users' well-being ([9]; 5-point smiley rating scale), and stress level ([10]; three items rated on a 7-point Likert scale; e.g., 'At the moment, I feel burdened'). We do not ask for the user's perception of characteristics in information systems (e.g., transparency) as we do not want to influence the user's judgment in the scenarios and later measurements. After completing the scenario, the user is guided to the end of the study. We them ask for the user's perception of specific characteristics. This includes both manipulated and not manipulated characteristics. We do this to investigate the complex multi-relationships between the characteristics exploratively. For finishing the user journey, we measure the user's performance by either comparing their answer 1) to the correct solutions or 2) to the answers of other participants. This depends on the given scenario and cannot be generalized at this point.

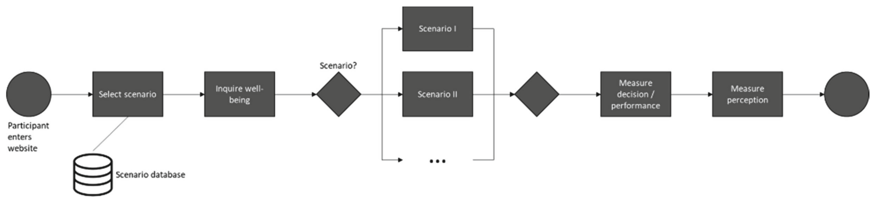

Fig. 3. User journey from a participant perspective

3 Significance to Practice and Research

The successful understanding of human-machine collaborations is governed by interpreting the relationships of characteristics in those scenarios. In the real world, it isn't easy to decipher the influence of characteristics, a person's mood, or their disposition to technology in general. Hence, we use a experimental setting combined with a quantitative, empirical approach to produce results that can explain the complex multidimensional structure of characteristics in human-machine collaboration scenarios.

Against this background, the main contribution of the proposed prototype is the capability to scientifically evaluate designed scenarios involving notions of artificial intelligence and semi-automated decision-making in human-machine collaboration scenarios. To be more precise, the prototype enables the identification, evaluation, and selection of characteristics grouped in various scenarios. The results of completed scenarios can be interpreted and investigated using multiple means of descriptive analysis (e.g., intercorrelations between investigated variables). The prototype is designed to be extended continuously while also adapting to the learnings of researchers in the course of participants completing the scenarios. We also plan to release the source code publicly to allow others to verify and adapt our work in their research.

4 Evaluation and Conclusion

The development of the prototype is driven by an iterative evaluation in both theoretical and implementation aspects. In assessing characteristics and their relationships, we build on previous empirical research to verify their relevance in the first place. Afterward, we re-evaluate the previous insights on attributes continuously in interpreting the results of the scenarios.

To evaluate scenarios and general user experience, we plan to collect individual feedback from potential end-users of the application. The application assessment will be done in both moderated, i.e., performing specific scenarios, and explorative, i.e., "click on whatever you fancy", testing.

Despite the focus on the actual use and implementation, the prototype also touches on diverse research topics. First, we combine aspects from information systems (e.g., human-machine collaboration) with psychological constructs (e.g., decision-making, trust, well-being). Traditionally, the intersections of different domains and topics foster collaborative idea identification. Hence, we generate new insights to understand better human-machine collaboration on both the psychological and the technical sides. As a result, we expect our results to provide initial guidance to management audiences on how intelligent information systems should be designed depending on the desired role of the system.

Acknowledgment. This work has been written in the context of the research project "Getrost Vergessen". The project was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft) under Grant HE 2745/16-2 and BE 1422/21-2.

References

1. Parasuraman, R., Sheridan, T.B., Wickens, C.D.: A model for types and levels of human interaction with automation. IEEE Trans. Syst. Man Cybern. A Syst. Hum. **30**(3), 286–297 (2000)
2. Duan, Y., Edwards, J.S., Dwivedi, Y.K.: Artificial intelligence for decision making in the era of big data – evolution, challenges and research agenda. Int. J. Inf. Manag. **48**, 63–71 (2019). https://www.sciencedirect.com/science/article/pii/s0268401219300581
3. Bansal, G., Nushi, B., Kamar, E., Weld, D.S., Lasecki, W.S., Horvitz, E.: Updates in human-ai teams: understanding and addressing the performance/compatibility tradeoff. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, no. 01, pp. 2429–2437 (2019)

4. Zhang, Y., Liao, Q.V., Bellamy, R.K.E.: Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, pp. 295–305 (2020)
5. McAfee, A., Brynjolfsson, E., Davenport, T.H., Patil, D.J., Barton, D.: Big data: the management revolution. Harv. Bus. Rev. **90**(10), 60–68 (2012)
6. Forssbaeck, J., Oxelheim, L.: The Oxford Handbook of Economic and Institutional Transparency. Oxford Handbooks (2014)
7. Meeßen, S.M., Thielsch, M.T., Riehle, D.M., Hertel, G.: Trust is essential: positive effects of information systems on users' memory require trust in the system. Ergonomics **63**(7), 909–926 (2020)
8. Hermann, A., Gollhardt, T., Cordes, A.-K., Kruse, P.: PlanDigital: a software tool supporting the digital transformation. In: Chandra Kruse, L., Seidel, S., Hausvik, G.I. (eds.) DESRIST 2021. LNCS, vol. 12807, pp. 356–361. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-82405-1_34
9. Jäger, R.: Konstruktion einer Ratingskala mit Smilies als symbolische Marken. Diagnostica **50**(1), 31–38 (2004)
10. Stanton, J.M., Balzer, W.K., Smith, P.C., Parra, L.F., Ironson, G.: A general measure of work stress: the stress in general scale. Educ. Psychol. Meas. **61**(5), 866–888 (2001)