All Good Things Come in Threes: A Digital Platform for Data-Driven, Interdisciplinary, and Replicable Research

Barbara Dinter[0000-0002-4512-7393] , Sarah Hönigsberg[0000-0003-4508-4650], and Oliver Specht[0000-0003-4441-6588]

Chemnitz University of Technology, Chemnitz, Germany
[barbara.dinter; sarah.hoenigsberg; oliver.specht]@wirtschaft.tu-chemnitz.de

Abstract. Digitalization affects research in such a way that new research subjects, methods, and infrastructures arise. Data, interdisciplinarity, and replicability are becoming increasingly important in research. The question arises how to best support research that combines and reflects these trends in the digital age. In our study, a digital research platform constitutes the envisioned solution artifact and a suitable approach to address this field of tension. Therefore, we ask: How should a digital research platform be designed for replicability in interdisciplinary research projects? We derive three tentative design principles (DP) for a digital research platform in preparation for an upcoming interdisciplinary research project with representatives from the mechanical engineering, mathematics, computer science, and information systems disciplines. Drawing on existing literature and case requirements, we identified the principles of (DP1) Collaborative Engagement, (DP2) Archiving, and (DP3) Virtualized Research Setting. We contribute with our prescriptive knowledge to the design knowledge of digital research platforms.

Keywords: Replicability, Interdisciplinarity, Data-Driven Research, Digital Research Platform, Design Principles.

1 Introduction

In the age of digitalization and global networks, all areas of the economy and society are undergoing a profound change. In this context, research with its primary goal of gaining knowledge is not only investigating many aspects of this transformation, it is changing, too. Consequently, the phenomenon of digitalization impacts research in multiple ways: (1) New research subjects in the field of digitalization; each digital innovation, such as blockchain infrastructures, artificial intelligence (AI), or the Internet of Things, expands the solution space for information systems (IS) design and influences potential impacts on society, both studied by our discipline. (2) New data- and technology-driven research methods that focus on large data sets and apply computerized analysis; (3) New types of research infrastructure and research organization supported by digital tools; as a result, research activities are more and more shifting to digital platforms that support location- and time-independent collaboration as well as the exchange of research data within and between various scientific disciplines.

In addition to such multi-faceted digitalization phenomena in research we can observe three major trends affecting the way of how researchers gain insights and collaborate. On the one hand, data-driven research meanwhile complements theory-driven research in IS as well as in other disciplines [1]. At the same time, globalized science, which requires mutually complementary competences to solve complex problems results in digital and interdisciplinary collaboration among researchers [2]. Not only, but also with regard to interdisciplinarity the third trend – replicability of research – gains relevance. Research setups by researchers from different disciplines, like experiments or data analyses, must be repeatable and comprehensible in detail for the collaboration partners in order to prevent misinterpretations and misunderstandings and to enable interdisciplinary exchange. In this tension, the IS discipline can contribute to increase the rigor of research through replication mechanisms and it can, especially with its design orientation, act as an interface discipline in interdisciplinary research teams [3].

We are currently preparing a multi-year research project on interdisciplinary collaboration with a focus on the manufacturing context with simulation and AI-based prediction. Various representatives from the mechanical engineering, computer science, mathematics, and IS disciplines will participate. In order to benefit from value propositions claimed by the three aforementioned trends we formulate our research question as follows: How should a digital research platform be designed for replicability in interdisciplinary research projects?

The remainder of the paper is structured as follows: In Section 2 we characterize the three perspective we are aiming for and derive requirements for the platform design from literature. After the research approach in Section 3 we present in Section 4 further design requirements originating from our project, platform design features and tentative design principles. In Section 5 we discuss the contributions and future work.

2 Foundations

In the following we briefly characterize the three aspects of data-driven, interdisciplinary, and replicable research and present initial requirements from literature (LR).

The megatrend of digitalization affects research because traditional research methods are increasingly complemented by new technologies and data-driven approaches [4]. In this context, we define data-driven research as the creation and exploratory analysis of data to gain insights about interesting phenomena, e.g. by identifying patterns within in the data [1]. Algorithms and machine learning are allowing the analysis of large amounts of data and the prediction (e.g. by simulation) of results [5]. Therefore, we derive the requirements related to data-driven research **(LR1.1):** The design of a data-driven research platform should enable users to exchange research data efficiently. Furthermore **(LR1.2):** The design of a data-driven research platform should allow virtualization and simulation of research corresponding to reality.

Solving a complex multidimensional problem often requires the involvement of several disciplines as well as heterogeneous and interdisciplinary collaboration in order to foster radically new and merged concepts [3]. We understand interdisciplinarity as the collaboration of two or more disciplines with their theories and concepts to make

(integrated) contributions to one's own and other disciplines or in their common intersections [3]. The success of interdisciplinary projects depends on various factors. A mutual understanding of research methods and results, as well as a consistent vocabulary among researchers from different disciplines, are seen as key success factors [2]. Consequently, we formulate **(LR2):** The design of a digital interdisciplinary research platform should reduce the cognitive discrepancy between the heterogeneous users.

Our third perspective refers to the replicability of research. Replicability constitutes a fundamental requirement in physical sciences, and its relevance in social sciences continuously grows [6, 7]. In general, replicability in IS refers to the capability to solve a problem by reproducing or verifying the same artifact (model, theory, or principle) with different data drawn from a different sample (or sampling) of the same population (or socio-economic environment) in the same spatio-temporal context [8]. Thus, the main focus lies on the actual repeatability of studies and calculations rather than on the sole traceability of results [6, 8]. The availability and completeness of documents and data, especially on theory, method, and context of the research to be replicated, must be ensured to enable such recalculations [7]. Technical procedures to support replicability include sandboxing for isolation and (data) snapshotting for archiving [9]. Therefore, we finally derive **(LR3.1):** The design of a digital research platform focusing on replicability should isolate and archive the digital representation of the research setup. Furthermore **(LR3.2):** The design of a digital research platform focusing on replicability should support transparency and comprehensibility of research.

3 Research Approach

The goal is to develop a digital research platform for replicability in an interdisciplinary research project. In contrast to network platforms such as ResearchGate or data repositories, we aim for an engagement and interaction platform on which virtualized research of interdisciplinary teams can occur. We systematically derived requirements, features, and tentative design principles (DPs) for this platform, referring to Kuechler and Vaishnavi's design science research framework [10].

Awareness of Problem: Initially, we collected data in focus groups and workshops for the problem description and suggested solution in the interdisciplinary project [11]. We conducted four focus group interviews (ca. 1 h each) and eight meeting rounds with workshop components (2 to 2,5 h each) with all involved researchers from the different disciplines during the setup phase of the project. The goal was to understand the problem, develop a joint solution approach, and gather case requirements (CR).

Suggestion: Our suggestion is to develop a digital research platform for replicability in interdisciplinary research projects. We base our proposal on a combination of gathered requirements from literature (LR) and our case (CR). From those requirements we derive design features (DF), which represent planned components of the digital platform for our case. The platform aims to enable interdisciplinary research teams, e.g., to conduct experiments virtually and to make all phases of this research process replicable.

Development & Evaluation: We derived tentative design principles following [12] to provide a starting point for the platform development. Throughout the project, the

requirements engineering will continue, and an agile development of the platform is planned. The evaluation will follow the human risk and effectiveness strategy, starting with an artificial evaluation (e.g., testing against our defined requirements) and quickly moving to a naturalistic evaluation (e.g., interventions in our project setting) [13].

4 An Interdisciplinary Research Platform for Replicability

Awareness of Problem: Our project's problem setting is that tests on physical objects and experimental setups in mechanical engineering are time and material consuming, but not sustainable and often unnecessary considering today's digital technologies, e.g. for simulating material behavior. Thus, the mechanical engineers told us they would like to store generated sensor data from their experiments "just somewhere" (CR1). The data should be used to train an AI system that can predict future experiment results (CR2). The process should be automated (CR3), and, if possible, the predictions should be used to derive suggestions for adjustments in the experimental setups (CR4).

The data are transferred to researchers from other disciplines (mathematics and computer science) to perform modeling and calculations in the studied case. However, the researchers expressed concern in the interviews that the results could be confusing and incomprehensible. Therefore, versioning capabilities of raw and processed data and models were identified as a requirement (CR5). In addition, computer scientists and mathematicians emphasized the integration of other data management infrastructures in a transparent way to users (CR6). In the interviews, it became apparent that the various researchers preferred familiar user interfaces, varying from programming interfaces to search masks or wizards. Therefore, the last requirement we identified was the need for different access options for the different disciplines (CR7).

Suggestion: Fig. 1 shows the components of the future digital research platform as design features derived from the literature and case requirements and how they lead to our tentative DPs. For easy data storage, a data lake (DF1) was defined as the repository of all data. The data navigation (DF2) will provide a semantic layer, e.g., to retrieve research data from the data lake and to facilitate the research platform's use. The view-based interfaces (DF3) ensure that different researchers have different access options. For the automated and virtualized research experiments, data pipelines (DF4) for data processing and a simulation and prediction component (DF5) for calculating the results were identified as essential components. The data pipelines should lead the data from the generation at the sensor up to the input into the simulation and prediction component, which is planned as an AI system based on mathematical models. Finally, a snapshot and isolation component (DF6) will enable the versioning of all research data. In addition to data and metadata storage, the containerization of the simulation and prediction components is planned with the latter.

Derivation of Tentative Design Principles: In preparation for the first design phase in the project, we derived three tentative design principles that capture the core concepts of the envisioned digital research platform. The design principles of (DP1) Collaborative Engagement, (DP2) Archiving, and (DP3) Virtualized Research Setting will guide

us in the initial design of the platform prototype and will be iteratively revised and refined. In detail, our design principles are as follows:

(DP1) Principle of Collaborative Engagement: To allow researchers in an interdisciplinary research project to engage and leverage a shared research platform, provide view-based interfaces, and enable data navigation because interdisciplinary research teams with shared world views will perform higher [3].

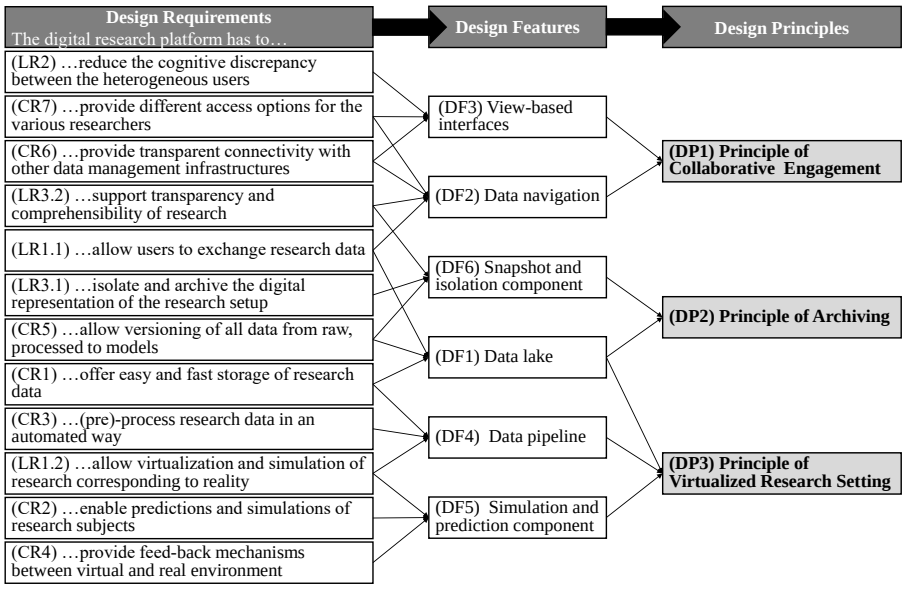

Fig. 1. Derivation of Design Principles.

(DP2) Principle of Archiving: To allow researchers in an interdisciplinary research project to conduct their research replicably, provide storage and archiving mechanisms at all levels of data processing because sandboxing and snapshotting enable the comprehensibility and repeatability of digitally supported research [9].

(DP3) Principle of Virtualized Research Setting: To allow researchers in an interdisciplinary research project to virtually conduct and simulate their research, provide a computerized core for simulations and prediction and a data pipeline to extract from the storage because researchers can leverage data-driven approaches to examine complex phenomena in a resource-efficient manner [5].

We could capture all design requirements and features in our design principles resulting in the first concrete recommendations for the design of the digital research platform.

5 Contribution and Future Work

In this research-in-progress paper, we address how to design a digital platform for data-driven, interdisciplinary, and replicable research. To this end, we derive three tentative design principles abstracted from six design features, which are based on five literature

and seven case requirements. Hereby, we provide cumulated prescriptive knowledge and thus contribute to the knowledge base of digital research platforms.

In subsequent research, we plan to extend our tentative design principles by considering and including further design requirements. Flexibility might constitute another emerging requirement where the platform can be easily adjusted to new situations, such as integrating new researchers. Our interview partners indicated openness to new collaborations on the platform as a potential future requirement. Furthermore, more research needs to be done to confirm our derived design principles – which we plan by an actual instantiation of our principles. The preliminary nature and high abstraction of our design principles also represents a limitation. Accordingly, it should be emphasized that completeness cannot be assumed at this stage. Seven of our requirements arise from our project case, thus, general transferability of all requirements to other contexts might not be given. Nevertheless, our research provides an initial building block for further work on digital research platforms focusing on interdisciplinary and replicability. It can offer relevant insights into the practice of interdisciplinary research projects.

References

1. Maass, W., Parsons, J., Purao, S., Storey, V.C., Woo, C.: Data-driven meets theory-driven research in the era of big data: Opportunities and challenges for information systems research. J. Assoc. Inf. Syst. 19, 1253–1273 (2018).
2. Kroeze, J.H., Van Zyl, I.: Transdisciplinarity in Information Systems: Extended Reflections. In: Twentieth Americas Conference on Information Systems. pp. 1–10. , Savannah (2014).
3. Tarafdar, M., Davison, R.M.: Research in information systems: Intra-disciplinary and inter-disciplinary approaches. J. Assoc. Inf. Syst. 19, 523–551 (2018).
4. Sheng, J., Amankwah-Amoah, J., Wang, X.: A multidisciplinary perspective of big data in management research. Int. J. Prod. Econ. 191, 97–112 (2017).
5. Müller, O., Junglas, I., Brocke, J. Vom, Debortoli, S.: Utilizing big data analytics for information systems research: Challenges, promises and guidelines. Eur. J. Inf. Syst. 25, 289–302 (2016).
6. Dennis, A., Valacich, J.: A Replication Manifesto. AIS Trans. Replication Res. 1, 1–4 (2014).
7. Brendel, A.B., Greulich, R.S., Niederman, F., Trang, S.: Towards a Greater Diversity of Replication Studies. AIS Trans. Replication Res. 6, 1–15 (2020).
8. Olbrich, S., Frank, U., Gregor, S., Niederman, F., Rowe, F.: On the Merits and Limits of Replication and Negation for IS Research. AIS Trans. Replication Res. 3, 1–19 (2017).
9. Meng, H., Kommineni, R., Pham, Q., Gardner, R., Malik, T., Thain, D.: An invariant framework for conducting reproducible computational science. J. Comput. Sci. 9, 137–142 (2015).
10. Kuechler, B., Vaishnavi, V.: On theory development in design science research: anatomy of a research project. Eur. J. Inf. Syst. 17, 489–504 (2008).
11. Morgan, D.L.: The focus group guidebook. SAGE Publications, Inc., Thousand Oaks, CA (1997).
12. Gregor, S., Chandra Kruse, L., Seidel, S.: The Anatomy of a Design Principle. J. Assoc. Inf. Syst. 21, 1622–1652 (2020).
13. Venable, J., Pries-Heje, J., Baskerville, R.: FEDS: A Framework for Evaluation in Design Science Research. Eur. J. Inf. Syst. 25, 77–89 (2016).