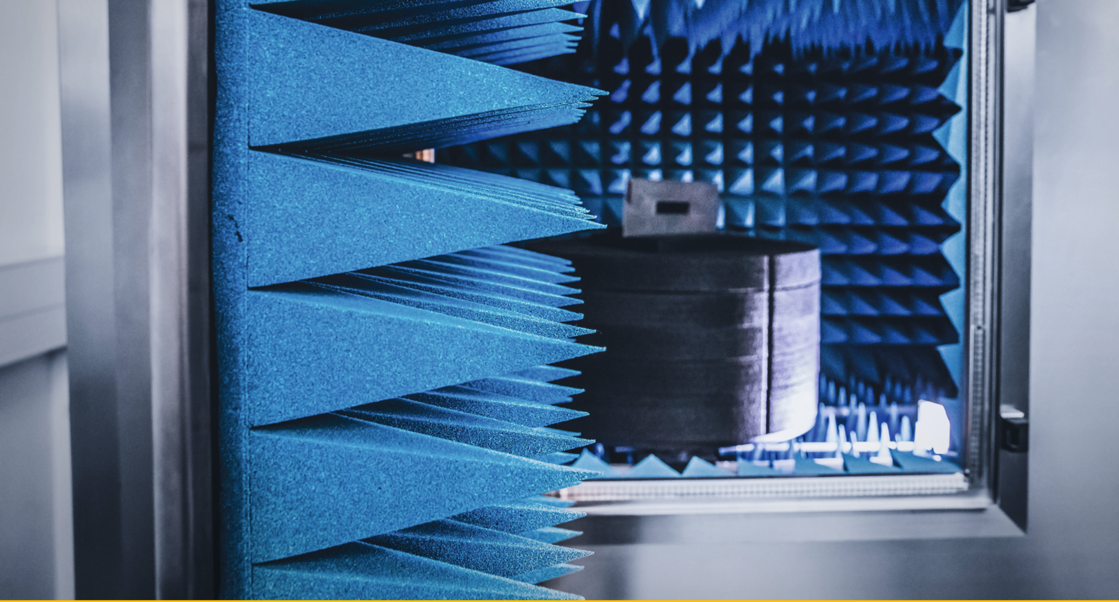

Uckelmann | Romagnoli | Baalsrud Hauge | Kammerlohr [Eds.]

Online-Labs in Education

Proceedings of the 1st International Conference
on Online-Labs in Education,
10 – 12 March 2022, Stuttgart, Germany

Nomos

Edition
Rainer
Hampp

Dieter Uckelmann | Giovanni Romagnoli
Jannicke Baalsrud Hauge | Valentin Kammerlohr [Eds.]

Online-Labs in Education

Proceedings of the 1st International Conference
on Online-Labs in Education,
10 – 12 March 2022, Stuttgart, Germany

Nomos

Edition
Rainer
Hampp

The DigiLab4U project, the conference and these conference proceedings were funded by the Federal Ministry of Education and Research (BMBF), Germany under the funding code 16DHB2112. The responsibility for the content of this publication lies with the authors.

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.d-nb.de

ISBN 978-3-98542-036-0 (Print)
978-3-95710-410-6 (ePDF)

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

ISBN 978-3-98542-036-0 (Print)
978-3-95710-410-6 (ePDF)

Library of Congress Cataloging-in-Publication Data
Uckelmann, Dieter | Romagnoli, Giovanni | Baalsrud Hauge, Jannicke
Kammerlohr, Valentin
Online-Labs in Education
Proceedings of the 1st International Conference on Online-Labs in Education, 10 – 12 March 2022, Stuttgart, Germany
Dieter Uckelmann | Giovanni Romagnoli | Jannicke Baalsrud Hauge
Valentin Kammerlohr (Eds.)
501 pp.
Includes bibliographic references.

ISBN 978-3-98542-036-0 (Print)
978-3-95710-410-6 (ePDF)

Edition Rainer Hampp in der Nomos Verlagsgesellschaft

1st Edition 2022

Published by
Nomos Verlagsgesellschaft mbH & Co. KG
Waldseestraße 3–5 | 76530 Baden-Baden
www.nomos.de

Production of the printed version:
Nomos Verlagsgesellschaft mbH & Co. KG
Waldseestraße 3–5 | 76530 Baden-Baden

ISBN 978-3-98542-036-0 (Print)
ISBN 978-3-95710-410-6 (ePDF)
DOI https://doi.org/10.5771/9783957104106

Onlineversion
Nomos eLibrary

Preface

The *Online-Labs in Education* conference, held on March 10-11, 2022 at HFT Stuttgart, Germany, marked the final phase of the DigiLab4U project (www.digilab4u.com). The goal of the conference was to disseminate the results of the project and to bring together a vibrant research community that is continuously working to enable the use of online labs in research and education without institutional boundaries.

Similar to digitalization in administration and industry, the digitalization or virtualization of lab equipment promises numerous benefits for involved stakeholders. The economic benefits of shared lab infrastructures, remote access to labs anytime and anywhere, the convenience of use, shared lab courses, fully online-labs, and experiments via online platforms are just a few of the benefits that come to mind when thinking of a federated lab network infrastructure. However, the effort required to digitalize 'physical things' should not be underestimated. The DigiLab4U project has investigated technical, organizational, and didactic issues related to online labs and lab-sharing networks. The scientific results have been published in numerous publications (see https://digilab4u.com/publications/). In our perception, publishing and discussing lab-based lecture content among peers is not common practice at scientific conferences, which leads to low visibility of what exists and hinders the uptake and re-use of existing online and remote labs.

For this reason, the *Online-Labs in Education* conference called for not only scientific contributions, but also lab-based lecture chapters. We would like to express our thanks to all authors for their contributions. The conference and the proceedings presented here followed the same structure and were complemented by interactive demonstrations. The conference proceedings are structured into the following main sections:

- General topics and organizational issues
- Technical topics
- Didactical considerations
- Educational learning chapters
  (educational considerations – learning chapters are available online)
- Interactive demos (abstracts)

There are institutions and people we would like to thank for their support of the conference. First, we would like to thank the *Federal Ministry for Education and Research (BMBF) Germany* for funding the project and the

conference. Second, we would like to thank our local institution, *Verein Freunde der HFT Stuttgart e.V.*, for their support. Third, we would like to thank the members of the advisory board and the program committee:
Advisory Board:

- Dr. Peter Ferdinand, Institut für Wissensmedien, Universität Koblenz-Landau
- Prof. Massimo Bertolini, Ph.D., Industrial Systems Engineering, University of Modena and Reggio Emilia
- Prof. Dr. Angel A. Juan, Data Science & Artificial Intelligence, Full Professor of Operations Research at the Computer Science Dept., Open University of Catalonia
- Prof. Dr. Abdelmajid Khelil, HAW Landshut
- Prof. David Romero, Ph.D., School of Engineering and Sciences, Tecnológico de Monterrey
- Prof. Dr. Gottfried Zimmermann, HDM Stuttgart
- Prof. Dr. Peter Rödler, University of Koblenz

Program Committee:

- Hadi Adineh, HFT Stuttgart
- Martin Burghardt, IWM, Universität Koblenz-Landau
- Jens Doveren, RWTH Aachen
- Giovanni Esposito, University of Parma
- Matteo Galli, University of Parma
- Birte Heinemann, RWTH Aachen
- Valentin Kammerlohr, HFT Stuttgart
- Nils Höhner, IWM, Universität Koblenz-Landau
- Davide Mezzogori, University of Parma
- Mattia Neroni, University of Parma
- Anke Pfeiffer, HFT Stuttgart
- Davide Reverberi, University of Parma
- David Schepkowski, IWM, Universität Koblenz-Landau

Last, but not least, we want to thank the organizing committee, including Anja Ernst, Anke Pfeiffer, Andreas Jäkel, Elisabeth Kraxner, Kevin Kutzner, and Marc Philipp Jensen for spending countless hours enabling a hybrid conference format.

The discussion at the end of the conference showed that further research and activities related to online-labs in education are expected. Some of the mentioned topics included:

- a demand for more universities collaborating and sharing labs;
- a need for a marketing platform to promote the usage of online labs;
- a need for more events to share experiences on using online labs in engineering education;
- interest in sharing experience on using Learning Analytics (LA) – not only in relation to online labs;
- the need for sustainable financing of labs and lab networks.

It will be interesting in the future to see the sharing of online labs by universities and eventually, cooperation between industries and lab teams will also be well accepted. However, this would require a fundamental shift in existing mindsets, budgets and funding programs.

Dieter Uckelmann
Giovanni Romagnoli
Jannicke Baalsrud Hauge
Valentin Kammerlohr

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

Table of Contents

1 Scientific Contributions on Online Labs

1.1 General Topics and Organizational Issues

Valentin Kammerlohr and David Paradice

Fundamental Organizational Aspects of Shared Lab-Networks: Trust, Business- and Maturity-Model Considerations in DigiLab4U

Abstract

Online labs form the basis of digital exchange in networks and are thus candidates for the use of shared knowledge, shared infrastructure, and shared facilities through the application of ICT technology. In addition to technical and didactic considerations, the importance of organizational considerations in this respect is increasing due to shared use. In this paper, the organizational foundations of digital sharing are highlighted, providing a long-term perspective on lab networks. To this end, three organizational aspects are addressed: (1) a platform business model for activating online lab sharing, (2) considerations on building initial and long-term trust between actors as a critical challenge of a lab sharing platform, and (3) a maturity model for capturing the organizational transformation of online labs for platform actors. Using the case study DigiLab4U, a time-limited, funded research project on online lab sharing, this paper shows how the three organizational considerations can contribute to sustainability over the funding period. The reader is thereby shown which success criteria and functional requirements are necessary for the sustainability of a lab-sharing network.

Keywords

Sharing economy, Online labs, Platform business model, Trust, Transformation maturity model

1 Introduction

Science, technology, engineering, and mathematics (STEM) education requires applied tasks and problems to promote conceptual understanding, practical knowledge, and experience (Feisel & Rosa, 2005). Laboratories (labs) provide students with a special hands-on engineering experience and allow them to explore systems and their real-world behavior in a protected environment (Zutin et al., 2010). However, for universities, these specialized labs involve high investment and operating costs, their utilization is often low, access is limited to local user groups (students and researchers), and the labs are subject to rapid loss of innovation (Heradio et al., 2016). In addition, funding for the labs is solely dependent on budgetary resources and grants and is therefore subject to corresponding funding fluctuations.

Digitalization technologies can be used to transform traditional labs, making them available online, allowing access to labs across locations, eliminating the need for in-person lab attendance, and thus enabling the delivery of lab experiences via distance learning (Mani & Patvardhan, 2006). These online labs are experiments supported by information and communication technologies (ICT) in which manual efforts are eliminated and can be accessed via the Internet (Zutin et al., 2010). According to Zutin et al. (2010), online labs are divided into software simulations (or virtual labs) and labs with real hardware equipment (or remote labs), through which they achieve advantages in availability, observability, accessibility, and security (Heradio et al., 2016). Expanded availability, where users can access online labs from anywhere and at any time, offers universities wide-ranging opportunities to increase usage through new business areas created by the transformed labs, as they can be used outside class hours (Gardel et al., 2012).

By making the labs available online, they can be shared with other facilities and users, opening up a new area of business and thus a potential constant source of revenue. The sharing economy describes behavior that promotes the sharing of resources to benefit from increased resource utilization, cost advantages, and access to new knowledge (Goudin, 2016). The digital transformation of labs to online labs makes them good candidates for the sharing economy, which means additional users can be reached. Several didactic and technological studies have already been conducted to measure the transformation from real to online labs, such as in Brinson (2015) and García-Zubía (2021). Research by Uckelmann (2012) has shown that in addition to didactics and technology, an organizational element is required for online lab sharing, so this paper explores (**RQ**): What are the fundamental organizational aspects of shared lab networks?

Using a case study methodology, this study focuses on three key organizational aspects: (1) a platform business model for activating online lab sharing, (2) considerations for building initial and long-term trust between actors as a critical challenge of a lab sharing platform, and (3) a maturity model for capturing the organizational transformation of online labs for platform actors. The case study is described in Section 2 and introduces the DigiLab4U research project, which explores lab sharing as part of the research mission and, like most such research projects, faces the challenge of being sustained through the adoption of a business model after the funding period (Esposito et al., 2021). Section 3 then describes the three fundamental organizational aspects of shared lab-networks, further possible approaches, and their interplay for sustainable online lab sharing. For the business model as the first aspect, the following problem is addressed. Labs involve high investment costs, utilization is often low, access is limited to local user groups, labs are subject to rapid loss of innovation, and funding is solely dependent on budget and grant funding and thus subject to corresponding fluctuations. While classic business models do not seem to work, the question of what success criteria and functional requirements should be placed on digital labs is outlined. The second aspect describes the success criterion of trust for the business model in more detail. Trust should be a core element in sharing digital goods such as online labs (Gossen et al., 2019). We, therefore, show how technology-based initial and long-term trust development is approached. Building on this, we show why this is a core element for the sustainability of the sharing business model and how initial and long-term trust can be leveraged for sharing labs. Finally, as a third aspect, it is clear that the effectiveness of the digital transformation of the lab should be made transparent to the user and the platform operator, as has been shown many times for both didactic and technical transformation (Heradio et al., 2016). This could be used by lab operators for design, implementation, or improvement and by users such as students for comparability to build trust. Section 4 then discusses the results, interplay, further approaches, open challenges, and a possible way forward for shared laboratory networks before Section 5 concludes the paper.

The authors point out that individual aspects of this publication have been published before, but the interaction of the aspects is new.

2 DigiLab4U as a Case Study for Shared Online Labs

The mission of the DigiLab4U research project is to make real labs accessible and shared online. Participating institutions work across international borders to achieve common goals for teaching, learning, and research. Collaboration among universities and research institutions allows resources to be pooled so that faculties, learners, and researchers have access to a greater variety of digital courses based on different labs. Currently, the project relies solely on budget and grant funding and is therefore subject to corresponding fluctuations. The inclusion of potential user fees is intended to open up a third pillar of funding and thus create a viable business model that enables sustainability beyond grant funding.

Many research projects, such as the DigiLab4U research project, can be seen as a virtual organization whose typical customer is the funding organization. Their common goal and vision are described in the project proposal and in the statement of work (Seifert, 2009). These are temporary and end when the funding ends. When the temporary lab network is transformed into a sustained, long-term form of collaboration, the goals of the participants may change from jointly meeting the needs of the funder to goals that fit into the long-term business strategy. In some cases, the goals will change only slightly. In other cases, the partners will have such different goals that they will end the collaboration. The corresponding business model will change, however, as the revenue streams will change because the interested funder will have to be replaced by a different type of customer.

Various technical and didactic measures have been taken to best meet the needs and preferences of DigiLab4U stakeholders. The success of operating shared resources in a collaborative network depends not only on business considerations that take into account the needs of all stakeholders, but also on trust between stakeholders and the maturity of the online lab transformation. The introduction of a user fee as a business model changes, among other things, the stakeholders and their goals and relationships.

3 Trust, Business and Maturity Models as Organizational Aspects

Students, researchers, professors, universities, and institutions need organizational measures that go beyond technical and didactic measures to organize and sustainably map the sharing between the stakeholders. These organizational aspects are outlined below and thus form the necessary framework for the successful introduction of a sharing business model. To this end, we first outline a business model that addresses stakeholder needs, initial and long-term technology-based trust to leverage the business model, and a maturity model that maps the effectiveness of transforming an online lab for users.

3.1 A Multi-Sided Platform to Activate the Sharing of Online Labs

Sharing is originally a private matter, but new concepts for sharing goods and services between individuals and companies are emerging worldwide (Beutin, 2018). Sharing is now taking on far-reaching new forms, such as car sharing, code sharing, file sharing or food sharing, and is conquering new business areas with innovative business models. The underlying concept of the sharing economy describes behavior in which either individuals or organizations seek to share existing resources, such as human, tangible, and intangible resources (Goudin, 2016). At first glance, the sharing economy for digital labs offers benefits to providers through increased utilization, and the customer side gains access to a greater supply of labs.

According to Eikaas et al. (2003), a major obstacle to a sustainable business model is "*the willingness of customers to buy access to laboratory resources*". The benefits of sharing must be demonstrated over direct access, and the real benefits are the selling point. Customers expect valuable content, ease of use, affordable services, access to otherwise inaccessible materials and equipment, and customer support (Kammerlohr et al., 2021). In addition, sharing must be trustworthy. This point relates to both the functional and success criteria of sharing. Compared to physical markets, where trust is built through relationships, the digital environment currently uses transparent rating systems that consider the quality and reliability of the actors (Schallmo et al., 2017). Nevertheless, this does not achieve the interpersonal trust that comes into play in social contacts, as discussed in more detail in Section 3.2. However, a closer look at possible user groups, such as industry and students, also shows that they have different requirements (Kammerlohr et al., 2021). For industry, for example, integration into the corporate structure, data security, and the protection of intellectual property are of great importance. For students, on the other hand, the added value must be recognizable in comparison to or in addition to their regular lectures.

Thus, a business model for online labs is needed that is tailored to the needs of customers and providers while fulfilling the trust in network organizations. Following the business models of the leading providers in the sharing economy, a multi-sided platform would be suitable for the activation of the concept. A multi-sided platform is an intermediary for exchanging value between interested parties and providers from two or more markets (Zhao et al., 2019); for example, it is used by Airbnb (landlords and renters), eBay (buyers and sellers), and Facebook (users, advertisers, and content developers). In the DigiLab4U concept, the universities would offer various online labs, and students could meet their needs via a corresponding platform. Here, the interested parties are the students, and the providers are the universities. The marketplace is the DigiLab4U platform, where the joint exchange and coordination service takes place and supports matching providers with buyers (European Commission, 2013). The main difference with a traditional business model is that the DigiLab4U marketplace does not acquire ownership of the resource traded and therefore has no influence on the way it is presented or the price. The terms of sharing are therefore directly controlled by the provider and the buyer. The online lab provider must therefore keep its offer and prices attractive to attract and retain buyers. More users on both sides (supply and demand) increase the benefits of the Digilab4U marketplace, the so-called network effect (Abdelkafi et al., 2019).

The challenges of a multi-sided platform, according to Henseling et al. (2018), are: (1) building user trust, (2) evolving what the marketplace offers, and (3) attracting new user groups. In addition, the life cycle of a marketplace is described as consisting of three phases (Abdelkafi et al., 2019 ; Otto & Jarke, 2019): (1) design: technological architecture and innovation of the platform (software and hardware), (2) dynamics: evolution of the platform and ecosystem by attracting users and adding new functionalities, and (3) performance: scaling, growth, and competitive success. The challenges for a multi-sided platform, such as developing trust, need to be addressed first, and depending on the current phase of the marketplace, the other goals and associated challenges afterward. Trust is extremely important for the development of the online lab, as a loss of trust could lead to the collapse of the network effect.

In the example of DigiLab4U, the user must trust that the online lab will be available at the right time and in the expected condition, while the operator, e.g. the university, trusts that the lab will be used correctly and under the conditions agreed upon and that no damage will occur. Particularly with online resources such as online labs, the parties involved may not know each other and must have confidence that each other's requirements will be met. Independent information from third parties can provide clues to this and help to build initial trust before one's own experience can be gained (Ba et al., 2003). Two further organizational aspects for the DigiLab4U shared lab network can be derived from the business model: first, a model for the marketplace that initially, but also in the long term, builds trust between the provider and customer to avoid disruptions in the network; second, a maturity model that allows providers to pre-evaluate the effectiveness of the online lab's transformation, and that provides a kind of trust reference for the user of a third actor that evaluates the effectiveness of the online lab's organizational transformation.

3.2 Trust to Leverage the Business Model and Increase Organizational Effectiveness

In general, trust arises from and in relationships, and therefore it can be created and destroyed (Flores & Solomon, 1998). A trust relationship involves two parties, there is uncertainty and risk, and the trust giver relies on the honesty and goodwill of the trust taker (Siau & Shen, 2003). A distinction is made between weak and strong trust relationships, with a strong relationship characterized by feeling secure and trusting that our partner can rely on us and will respond to our needs (Rempel et al., 2001). Trust occurs in various social contexts and can arise both between individuals and between individuals and organizations as a hybrid form (Zaheer et al., 1998). A further distinction is made between trust that already arises on the basis of an existing trust relationship and trust that must first be established. In addition to initial trust, there are trust models that reflect the development of trust during the interaction of the parties, such as that of Lewicki and Bunker (2010). The transitional stages of their trust development model describe how two parties form and develop a new relationship and explain how trust and relationships change, develop, or decline over time and how trust can be restored. In doing so, the transitional stages of the trust model maps different benefits and different costs for each stage (sequential iteration) (Lewicki & Bunker, 2010).

In contrast to trust in a social context, online trust or technology-based trust is increasingly being studied, e.g., e-commerce (Gefen, 2000), trust in smart personal assistants (Zierau et al., 2020), in blockchain platforms (Zavolokina et al., 2020), or for entire research disciplines such as information systems (Söllner et al., 2016). The difference with online trust is that in an online situation, it is more difficult to reasonably assess the potential harm and goodwill intended by others (Friedman et al., 2000). To this end, new methods have been developed, such as a user-centered rating system, trusted third-party certifications, or trusted third-party recommendations. Long-term trust conditions have not yet been used in technology practice because the focus to date has been on initial trust rather than the impact of long-term development. Similar to Lewicki and Bunker's (2010) model, Williamson's (1993) transaction cost theory follows a parallel idea when parties begin to validate activities in terms of trust to build a knowledge base about their needs, preferences, and priorities. According to this theory, a transaction can be processed and organized more or less efficiently, which describes the transaction costs. An adaptation of this transaction concept for trust in a technology context could map trust interactions into a trust-level model and promote benevolent behavior through lower (transaction) costs. For example, a user with a higher level of trust will be more willing to accept a trusted online lab offer because they will expect the transaction costs to be lower. One way to provide information to users to build long-term trust is to develop technology maturity models, as described in the next section, which are enabled by a business model by influencing costs through trust levels. The combination of a business model and a maturity model, both of which promote long-term trust, should ensure that after the initial trust, there is stakeholder interest in building a long-term trust relationship, thus underpinning the sharing.

3.3 Maturity Model for the Effectiveness of Digital Lab Transformation

Digital transformation is defined by Pousttchi et al. (2019) as a change process that companies undergo due to the emergence of new technologies and their social and economic impact. Digital transformation of labs is therefore defined as a continuous development process that goes beyond the emergence of new technologies and their social and economic impact to include the construction of a new business ecosystem. Various studies measure this transformation from a didactic and technical perspective in order to make its effectiveness transparent to stakeholders and to build trust. The pedagogical effectiveness of online labs at different stages of digital transformation as an indicator of the usefulness of an experiment to achieve the desired goal has been studied by various authors, e.g., Brinson (2015). Similarly, studies on technological effectiveness, such as the design, development, and implementation of different digital lab transformations have been pursued, as by Prada et al. (2015). Corresponding maturity models for both areas can also be found in research literature (Abbas, 2019), but a model that takes into account organizational change towards sharing between institutions and thus the needs of users and operators is currently lacking. The organizational effectiveness of digital lab transformation has not been further researched since then but has gained importance over time due to the changing requirements of lab sharing, such as building initial trust between different actors and organizations. Numerous international research projects involving online labs have failed to continue the environments developed after the project funding phase (Esposito et al., 2021), not least due to lack of effectiveness.

Digital lab transformation effectiveness is defined as the evaluation of the lab's digital transformation efforts with the goal of sharing (Kuntsman & Arenkov, 2019). In this study, effectiveness is specifically defined as the quality of the change process organizations undergo that involves technology and its social and economic impact. Specifically, effectiveness is about four dimensions: (1) universality and accessibility, (2) user management, (3) scalability and extensibility, and (4) learning support. These are aligned with Garcá-Zubía's (2021) structures for the requirements of a remote lab management system and the characteristics of a remote experiment. Universality and accessibility describe if and how a lab is accessible to the user in any technological scenario and refer to the original design of the experiment (García-Zubía, 2021). In this context, Prada et al. (2015) added easier support and efficient management. The second requirement, user management, consists of four subsections that describe how users gain access to online labs, how their data is managed, what user rights they have, and how their experiment data set is stored (García-Zubía, 2021 ; Ying & Zhu, 2004). The third requirement is scalability and extensibility, which consists of five subcategories. Scalability and extensibility describe how easy it is to adapt the labs to new audiences, expand them to include more experiments, extend them to more facilities, certify the results, and ensure sustainability (García-Zubía, 2021). The fourth requirement is learning support or pedagogical effectiveness, i.e., whether and to what extent the online laboratory supports coursework (García-Zubía, 2021). As Kara et al. (2010) stated, "*effective learning in engineering education can only be achieved through approaches that link theoretical courses to the laboratory*." The lab supports not only experimentation, but also social coordination, the lab environment, and individual differences (Nickerson et al., 2007). We distinguish between learning environment, interactivity and realism, technical support and maintenance, and didactic support.

In terms of application in the DigiLab4U project, it should therefore be further investigated whether the effectiveness of digital lab transformation in support of the sharing economy can be mapped in a maturity model. The practical use of DigiLab4U promises comparability of the effectiveness of shared labs, both from the provider's perspective in terms of administrability and financial and personnel effort, and from the perspective of demand in terms of learning success. The theoretical perspective has shown that the effectiveness of digital lab transformation assumes that criteria are subject to multiple truths and that these are determined by the subject matter and the underlying use case. Therefore, a constructivist approach based on multi-stakeholder interaction should be used to gain insights into and build knowledge about the effectiveness of labs' digital transformation within the shared DigiLab4U network.

4 Discussion

Returning to the research questions, we can conclude that this research has provided an overview of the fundamental organizational aspects of shared lab networks through a general understanding of business models, trust, and maturity models for digital labs. Specifically, a business model that promotes sustainability by enabling online lab sharing. A maturity model that can determine the effectiveness of digital lab transformation, and a consideration of trust as a key element of platform business.

As described in the second phase dynamics of the life cycle of the multi-sided marketplace, the platform and the ecosystem should be further developed depending on the current situation (Otto & Jarke, 2019). In addition to the three organizational aspects already described, there are numerous other ways to further develop the business model marketplace and keep it attractive to customers, which we divide into central services, individual services, and community services. By central services, we mean applications that are offered centrally from the marketplace to improve matching, for example. We think of cooperative resource management, from which a common booking and billing process, work properties and resources, and common and standardized terms of use emerge, but also a booking system that is able to cover the different needs of user roles, such as a recurring event for a lecture series. Associated with this should be a flexible billing system on a transaction basis to map the described levels of trust, individual but also standardized, national and international, and billing for companies and universities. Individual services are services provided by the marketplace on an individual basis, such as lab didactic or transformation services, research services, or the sale of processed research data. Community services are actions taken to build and sustain the community as an ecosystem around the marketplace. Simple things like a shared vision, or mission statement can help, but so can conferences, awards, badges linked to learning paths for external visibility of learning success, or individualized advertising for job openings or research contracts.

Generalizations can be applied to related application areas such as vocational training, but also to broader areas such as shared infrastructures and digital transformation, e.g., the business model of shared resources, as in research facilities for industry and research, or the effectiveness of transforming the digital infrastructure of government agencies or universities. Technology-based initial trust and long-term trust could be generalized in order to build trust in new technologies.

Limitations arise from differences in the education system, such as regional differences in the willingness to pay for education and the degree of digitalization. One problematic issue is the willingness to pay; traditionally, education in Europe has been free. Students accept that the cost of a digital lab must be paid, but there is considerable dispute about whether the state, the university, or the students themselves should pay for it. Another limitation is the language and cultural differences that affect the type of education (practical vs. theoretical) and the level of education (BA, MS, PHD). In addition, regional taxes and public sector billing may impose limitations.

5 Conclusion

This paper addresses the problem of sharing online digital education resources for STEM subjects. Lab exercises to gain hands-on experience and practical knowledge play an important role in the education of future engineers and scientists. Online labs can be used to gain this experience online. However, sharing online labs is currently insufficient; in fact, a large percentage of lab providers fail to keep them running (profitably) after research funding. This paper addresses this problem by highlighting the organizational aspects of online lab sharing to provide a long-term perspective for lab networks and to serve as a foundation for online lab sharing between providers and buyer. The analysis includes three different organizational proposals for improving the sharing of these labs to increase sustainability, using DigiLab4U as an example. The research potential, generalizations, and limitations were highlighted.

The next step is to further explore the organizational aspects scientifically and put them into practice using DigiLab4U as an example to gain insights and experience. This will provide a more detailed insight into the community, further elaborate on buyer and provider demand, make different online lab transformations organizationally comparable, clarify dependencies in more detail, and test the initial but also the long-term trust network. In addition, experience can be gained from later life cycle phases of the multi-sided platform business model for sharing online labs.

Acknowledgements

This work has been funded by the Federal Ministry of Education and Research (BMBF), Germany under the funding code 16DHB2112.

References

Abbas, A. (2019). *The Learning Labs Maturity Model: From Sandbox to Guided Learning*. https://trainingindustry.com/articles/it-and-technical-training/the-learning-labs-maturity-model-from-sandbox-to-guided-learning/

Abdelkafi, N., Raasch, C., Roth, A. & Srinivasan, R. (2019). Multi-sided platforms. *Electronic Markets*, *29*(4), 553–559. https://doi.org/10.1007/s12525-019-00385-4

Ba, S., Whinston, A. B. & Zhang, H. (2003). Building trust in online auction markets through an economic incentive mechanism. *Decision Support Systems*, *35*(3), 273–286. https://doi.org/10.1016/S0167-9236(02)00074-X

Beutin, N. (2018). *Share Economy 2017: The New Business Model*. www.pwc.de/share-economy

Brinson, J. R. (2015). Learning outcome achievement in non-traditional (virtual and remote) versus traditional (hands-on) laboratories: A review of the empirical research. *Computers & Education*, *87*, 218–237. https://doi.org/10.1016/j.compedu.2015.07.003

Eikaas, T. I., Schmid, C., Foss, B. A. & Gillet, D. (2003). A Global Remote Laboratory Experimentation Network and the Experiment Service Provider Business Model and Plans. *Modeling, Identification and Control: A Norwegian Research Bulletin*, *24*(3), 159–168. https://doi.org/10.4173/mic.2003.3.2

Esposito, G., Mezzogori, D., Reverberi, D., Romagnoli, G., Ustenko, M. & Zammori, F. (2021). Non-Traditional Labs and Lab Network Initiatives: A Review. *International Journal of Online Engineering (iJOE)*, *17*(05), 4. https://doi.org/10.3991/ijoe.v17i05.20991

European Commission. (2013). *The Sharing Economy*. https://www.academia.edu/9481688/Business_Innovation_Observatory_The_Sharing_Economy_Accessibility_Based_Business_Models_for_Peer-to-Peer_Markets

Feisel, L. D. & Rosa, A. J. (2005). The Role of the Laboratory in Undergraduate Engineering Education: Journal of Engineering Education.

Flores, F. & Solomon, R. C. (1998). Creating Trust. *Business Ethics Quarterly*, *8*(2), 205–232. https://doi.org/10.2307/3857326

Friedman, B., Khan, P. H. & Howe, D. C. (2000). Trust online. *Communications of the ACM*, *43*(12), 34–40. https://doi.org/10.1145/355112.355120

García-Zubía, J. (2021). *Empowering STEM education with technology: Remote laboratories*. World Scientific Publishing.

Gardel, A., Bravo, I., Revenga, P. A., Lázaro, J. L. & García, J. (2012). Implementation of Industrial Automation Laboratories for E-learning. *The International Journal of Electrical Engineering & Education*, *49*(4), 402–418. https://doi.org/10.7227/IJEEE.49.4.4

Gefen, D. (2000). E-commerce: the role of familiarity and trust. *Omega*, *28*(6), 725–737. https://doi.org/10.1016/S0305-0483(00)00021-9

Gossen, M., Pentzien, J. & Peuckert, J. (2019). What use is it really for sustainability? Potentials and impacts of peer-to-peer sharing in the domains of accommodation and mobility. *Nachhaltigkeits Management Forum*, *27*(2), 125–138. https://doi.org/10.1007/s00550-019-00488-8

Goudin, P. (Hrsg.). (2016). *The Cost of Non-Europe in the Sharing Economy: Economic, Social and Legal Challenges and Opportunities*. European Parliamentary Research Service.

Henseling, C., Hobelsberger, C., Flick, C. & Behrendt, S. (2018). *Nachhaltige Entwicklungsperspektiven für Geschäftsmodelle des Peer-to-Peer Sharing: PeerSharing Arbeitsbericht 6. PeerSharing.*

Heradio, R., La Torre, L. de, Galan, D., Cabrerizo, F. J., Herrera-Viedma, E. & Dormido, S. (2016). Virtual and remote labs in education: A bibliometric analysis. *Computers & Education*, *98*, 14–38. https://doi.org/10.1016/j.compedu.2016.03.010

Kammerlohr, V., Uckelmann, D. & Baalsrud Hauge, J. (2021). A Multi-Sided Platform to Activate the Sharing of Digital Labs. *International Journal of Online Engineering (iJOE)*, *17*(11), 4–33. https://doi.org/10.3991/ijoe.v17i11.25183

Kara, M., Aydin, E., Ozbek, M. E. & Cagiltay, N. (2010). Design and development of a remote and virtual environment for experimental training in Electrical and Electronics Engineering. In *2010 9th International Conference on Information Technology Based Higher Education and Training (ITHET)* (194–200). IEEE. https://doi.org/10.1109/ITHET.2010.5480040

Kuntsman, A. & Arenkov, I. A. (2019). Method for Assessing Effectiveness of Company Digital Transformation: Integrated approach. *IBIMA Business Review*, 1–16. https://doi.org/10.5171/2019.334457

Lewicki, R. J. & Bunker, B. B. (2010). Developing and Maintaining Trust in Work Relationships. In R. M. Kramer & T. R. Tyler (ed.), *Trust in organizations: Frontiers of theory and research* (114–139). Sage Publications. https://doi.org/10.4135/9781452243610.n7

Mani, A. & Patvardhan, C. (2006). A Study of ICT Enabled Laboratories. In *2006 Annual IEEE India Conference* (1–6). IEEE. https://doi.org/10.1109/INDCON.2006.302765

Nickerson, J. V., Corter, J. E., Esche, S. K. & Chassapis, C. (2007). A model for evaluating the effectiveness of remote engineering laboratories and simulations in education. *Computers & Education*, *49*(3), 708–725. https://doi.org/10.1016/j.compedu.2005.11.019

Otto, B. & Jarke, M. (2019). Designing a multi-sided data platform: findings from the International Data Spaces case. *Electronic Markets*, *29*(4), 561–580. https://doi.org/10.1007/s12525-019-00362-x

Pousttchi, K., Gleiss, A., Buzzi, B. & Kohlhagen, M. (2019). Technology Impact Types for Digital Transformation. In *2019 IEEE 21st Conference on Business Informatics (CBI)* (S. 487–494). IEEE. https://doi.org/10.1109/CBI.2019.00063

Prada, M. A., Fuertes, J. J., Alonso, S., García, S. & Domínguez, M. (2015). Challenges and solutions in remote laboratories. Application to a remote laboratory of an electro-pneumatic classification cell. *Computers & Education*, *85*, 180–190. https://doi.org/10.1016/j.compedu.2015.03.004

Rempel, J. K., Ross, M. & Holmes, J. G. (2001). Trust and communicated attributions in close relationships. *Journal of Personality and Social Psychology*, *81*(1), 57–64. https://doi.org/10.1037/0022-3514.81.1.57

Schallmo, D., Rusnjak, A., Anzengruber, J., Werani, T. & Jünger, M. (eds.). (2017). *Digitale Transformation von Geschäftsmodellen: Grundlagen, Instrumente und Best Practices*. Springer Fachmedien. https://doi.org/10.1007/978-3-658-12388-8

Seifert, M. (2009). *Collaboration formation in virtual organisations by applying prospective performance measurement*. *Bremer Schriften zur integrierten Produkt- und Prozessentwicklung: vol. 63.*

Siau, K. & Shen, Z. (2003). Building customer trust in mobile commerce. *Communications of the ACM*, *46*(4), 91–94. https://doi.org/10.1145/641205.641211

Söllner, M., Hoffmann, A. & Leimeister, J. M. (2016). Why different trust relationships matter for information systems users. *European Journal of Information Systems*, *25*(3), 274–287. https://doi.org/10.1057/ejis.2015.17

Uckelmann, D. (2012). The Role of Logistics Labs in Research and Higher Education. In *Communications in computer and information science: vol. 282. The impact of virtual, remote and real logistics labs: First international conference, ImViReLL 2012, Bremen, Germany, February 28–March 1, 2012; proceedings* (1–12). Springer.

Williamson, O. E. (1993). Calculativeness, Trust, and Economic Organization. *The Journal of Law and Economics*, *36*(1, Part 2), 453–486. https://doi.org/10.1086/467284

Ying, S. & Zhu, S. (2004). Remote laboratory based on client-server-controller architecture. In *ICARCV 2004 8th Control, Automation, Robotics and Vision Conference, 2004* (S. 2194–2198). IEEE. https://doi.org/10.1109/ICARCV.2004.1469506

Zaheer, A., McEvily, B. & Perrone, V. (1998). Does Trust Matter? Exploring the Effects of Interorganizational and Interpersonal Trust on Performance. *Organization Science*, *9*(2), 141–159. https://doi.org/10.1287/orsc.9.2.141

Zavolokina, L., Zani, N. & Schwabe, G. (2020). Designing for Trust in Blockchain Platforms. *IEEE Transactions on Engineering Management*, 1–15. https://doi.org/10.1109/TEM.2020.301 5359

Zhao, Y., Delft, S. von, Morgan-Thomas, A. & Buck, T. (2019). The evolution of platform business models: Exploring competitive battles in the world of platforms. *Long Range Planning, vol. 53 No. 4 (2019)*, 101892. https://doi.org/10.1016/j.lrp.2019.101892

Zierau, N., Engel, C., Söllner, M. & Leimeister, J. M. (2020). Trust in Smart Personal Assistants: A Systematic Literature Review and Development of a Research Agenda. In N. Gronau, M. Heine, K. Poustcchi & H. Krasnova (eds.), *WI2020 Zentrale Tracks* (99–114). GITO Verlag. https://doi.org/10.30844/wi_2020_a7-zierau

Zutin, D. G., Auer, M. E., Maier, C. & Niederstatter, M. (2010). Lab2go – A repository to locate educational online laboratories. In *IEEE EDUCON 2010 Conference* (1741–1746). IEEE. https://doi.org/10.1109/EDUCON.2010.5492412

Authors

Valentin Kammerlohr, PhD
HFT Stuttgart
70174 Stuttgart, Germany
Auburn University,
AL 36849 Auburn, USA
https://orcid.org/0000-0003-1795-3759
Valentin.Kammerlohr@hft-stuttgart.de

David B. Paradice
Auburn University
AL 36849 Auburn, USA
https://orcid.org/0000-0002-0287-3249

Giovanni Esposito, Davide Reverberi,
Giovanni Romagnoli and Riccardo Ghinzelli

Research Data Management for Laboratory Services: the DigiLab4U Use Case of Dataverse

Abstract

The ongoing digitalization of academia and research institutes has led to an increasing need for suitable processes of data management and dissemination. As a result, research is increasingly asking for standardized data management processes. Research data management (RDM) has emerged as an important concern for the whole scientific community, and several platforms to support data deposits have been designed and released. Actually, (i) rules of management and (ii) the curation of advanced data catalogues seem to be generally lacking. Moreover, one of the major challenges is encouraging consultation and 'buy-in' from researchers and senior managers. This paper presents the implementation of Dataverse by the DigiLab4U consortium from this standpoint. The benchmarking process among research data management (RDM) platforms available on the open-source market, and the hierarchical structure for storing and managing data are introduced and discussed.

Keywords

Research Data Management (RDM), Dataverse, Servitization

1 Introduction

The ongoing digitalization of academia and research institutes has led to an increasing need for suitable processes of data management and dissemination. As a result, important scientific journals, academia, but also third-party funding institutions are increasingly asking for standardized data management processes (Wilms et al., 2018). Research data management (RDM) has emerged as an important concern for the whole scientific community, and several platforms to support data deposits have been designed and released (Amorim et al., 2015). Technological progress, especially under the advent of the Internet of Things (IoT) era, created a state of the art involving high-level performances, with respect to the possibility of storing huge amounts of data, and accessing them everywhere by means of cloud services, for instance. This allowed providers to supply advanced data management services. Although these important results were achieved, the management and curation of advanced data catalogues seem to be lacking (Cox et al., 2017). As a consequence, institutions that eventually share the data must establish their own rules for the suitable management and promotion of research data (Wilms et al., 2018), otherwise, data catalogues could result in a mess. In addition, major challenges include (i) resourcing, (ii) adaptive capacities and communicability with other services, and especially (iii) encouraging consultation and 'buy in' from researchers and senior managers (Cox et al., 2017). The 'buy in' formula for the provisioning of data especially is a new and important topic in research, and it is gaining attention, especially from business entities that are interested in buying research data and scientific knowledge from academia and research institutes (Esposito et al., 2021). Therefore, this paper presents the implementation of Dataverse by the DigiLab4U consortium. The DigiLab4U consortium, under the Open Digilab4U project funded by the German Federal Ministry of Education and Research (BMBF), aims at creating a network of digitalized labs via the Internet of Things, towards hybrid education, cross-institutional research, and cross industrial cooperation. Dataverse is a type of open-source software for the management of files in the academic field. It has been selected within a benchmarking assessment, with respect to some requirements required by the DigiLab4U, and identified according to Amorim et al., (2015). This paper provides a standardized hierarchical structure for organizing the research material in a RDM system and hence answers the research question about a possible solution to managing the curation of content in research data catalogues. The remainder of the paper is as follows: section 2 provides an overview of the literature on digital online labs and RDM in this field. Section 3 briefly introduces the DigiLab4U case and provides the benchmarking assessments of the RDM platforms and software. In section 4, the hierarchical structure is presented and tested, and then its validation is discussed. Finally, section 5 addresses conclusions and outlooks for future works.

2 Literature review

The literature on digital online labs has been increasing even more over the last decade (Heradio et al., 2016). Basically, two research lines have arisen: one about the didactical perspective, and one about technical implementation (Zappatore et al., 2015). What is missing is a deep analysis of financial and organisational aspects for making labs and networks in which they are inserted robust from a life cycle perspective (Esposito, Kammerlohr, et al., 2021). In this regard, Esposito, Mezzogori, et al. (2021) have showed that most digital online experiences last only for the time in which they are funded by institutions and organizations under the programs of national or international projects. From this point of view, Esposito, Kammerlohr, et al. (2021) analyzed the possibility of using research data for business partnerships upon payment, with generally positive results from several Italian companies. As a result, platforms for data curation and sharing are needed of course. Although advisory and consultancy services have been recently stressed, technology and data deposit assistance seem still to be lacking and are only forecast in the near future (Corrall et al., 2013). As a counterproof, a single result is obtained by querying Scopus with the following search string: (TITLE-ABS-KEY ("Research Data Management" OR RDM)) AND ((("data curation")) AND ("data curation")) AND ("data catalogue"), and the only work by Cox et al. (2017) attests to the lack of works and research on data catalogues and the active curation of data.

3 Dataverse within the DigiLab4U environment

In this chapter, firstly the DigiLab4U case is described. The technical system is just referenced here using the work by Galli et al. (2020) and Kammerlohr et al. (2021) since it is not of interest in this paper and is not discussed further. A deeper overview of the services provided is discussed, instead. Secondly, the requirements of the DigiLab4U for the selection of the RDM system are introduced. There are four key aspects to identify the DigiLab4U requirements, according to Amorim et al. (2015): (1) architecture, (2) metadata handling capabilities, (3) interoperability, content dissemination, and search features, and finally (4) community acceptance. Lastly, commercial solutions are analyzed in a benchmarking assessment, resulting in the selection of Dataverse.

The DigiLab4U case and its services

DigiLab4U is the cross-Institutional network of Industry 4.0 lab infrastructure. The consortium is led by the Hochschule für Technik Stuttgart (HFT), and joined by the other four founding members: the Bremen Institute for Production and Logistics (BIBA), the Institut für Wissensmedien

(IWM) of Koblenz-Landau, the Rheinisch-Westfälische Technische Hochschule (RWTH) of Aachen, and the University of Parma. Nowadays, it has another nine partners all around the world (see https://digilab4u.com/consortium/). The network was funded by the German Federal Ministry of Education and Research (BMBF) for developing the project 'Open Digital Lab for You', which created an integrated, hybrid learning and research environment consisting of a large variety of lab technologies offering digital services and reaching all kinds of possible users. The digitized lab environment intends to enable a real IoT learning marketplace, consisting of both digital labs from several suppliers and users who access to the labs. The cooperation between universities, research institutions and industry allows the suppliers to be pooled so that users have access to a larger variety of digital courses based on different IoT labs. As a consequence, one of the main features of such a network is the availability of tracked and traceable data repositories for research data, and suitable systems for retrieving and accessing them. Hence, an RDM was needed.

Requirements of the DigiLab4U for the RDM system

The proposed system must comply with two main functionalities of storing data: their ownership is assigned and they can be retrieved with a suitable reference system. First, from the user's perspective, people have access to old batches of data or new data generated by remote, performing analysis without collecting data physically in a lab. Second, from a content uploader's perspective, while they are uploading a data set, a new DOI must be automatically created, making data referencing easier and faster when using them and the information generated by them. Lastly, an efficient and organized framework to store all the data is required to promote data consultation and referencing. These translate into the following seven requirements, and each one is related to the four key aspects of Amorim et al. (2015): (i) the tool must support REST API, (ii) it possibly needs to be open source, and (iii) data must be hosted within Germany—these three requirements meet the key aspect (1); (iv) the ability to visualize the stored data must be accomplished, (v) it must support a wide range of data formats, based on the recommendations by the German association "Verbund Forschungsdaten Bildung", and (vi) the facility to generate edit and back up data is required—these three requirements meet the key aspects (3) and (4); finally, (vii) extraction of metadata from data must be simple and structured—this requirement meets the key aspects (2) and (3).

Benchmarking commercial solutions

Several RDM systems are commercially available, and in this phase, some have been analyzed with respect to the requirements of the DigiLab4U. Software identified from a market investigation performed on the Web are listed here: CKAN, RADAR, bwScienceToShare, Freidok plus, Dataverse, EdShare, dSPACE, and DKAN. Functionalities and characteristics have been mapped through ten main features, according to the requirements. These are (i) support for file formats, REST API, service support, and community, (ii) suitable data visualization; (iii) consistently building technology; (iv) type of license and price; finally, (v) host and service provider, and main users. If a single requirement is missing from the above features, the software is neglected. As a result, two types of software have been selected for comparison here: Dataverse from the Dataverse project, and dSPACE by dSPACE GmbH. Dataverse is rich in features and has an active service and user community to support users and service providers. Hence, the benchmarking has been concluded selecting Dataverse, which meets all the DigiLab4U requirements.

4 The structure and its transposition

Dataverse is an open-source Web application from the Dataverse Project, developed to share, preserve, cite, explore, and analyze research data. Mainly used by academia, it allows researchers, journals, data authors, publishers, data distributors, and affiliated institutions to access and replicate data from research, ensuring academic credit and Web visibility. A Dataverse repository (wordplay for data-and-universe) is the software installation, which then hosts multiple virtual archives called Dataverse collections, administrated by its creator, who has access to managing all the settings. A Dataverse collection is a container for data sets, each one containing data files (e.g., research data, code, documentation) and related descriptive metadata (e.g., tag and keywords, including documentation and a code that accompanies the data). Once a file is uploaded into a data set, it is no longer possible to eliminate it from the Dataverse. Also, the Dataverses can hold one or more Dataverse collections, which can be set up for individual researchers, departments, journals, and organizations. Dataverses and data sets within the main Dataverse can be created and placed arbitrarily, and they can also be categorized by means of Dataverse categories that identify the type of data hub (e.g., institution, laboratory) and address possible query strings. Hence, the need for a standard framework for uploading and managing content, ensuring ease of retrieving and accessing data. The proposed structure consists of a five-level structure in a father-son manner, from the Dataverse of the single institution within the main DigiLab4U Dataverse (at the top of the hierarchy) to the file attribute (at the bottom of the hierarchy). The structure is provided in Figure 1, with reference to an example in which two partners upload data from experiments performed in their respective remote laboratories. Each level is described in the following. At level 0 it is possible to find every key partner in the Digilab4U network. Every institution will have its own Dataverse collection in the main DigiLab4U Dataverse. Level 1 contains all the labs of every institution. Level 2 refers to all the specific experiments or analyses that can be performed by a laboratory. Level 3 contains all the data sets of a specific type of experiment or analysis. Lower level 4 contains all the files of a specific data set. Three actions are envisaged when uploading content. First, before uploading data, they all need to be renamed using a formatting standard consistent with the uploading session. Second, specific tags need to be applied to every single file. Third, if the files need to be collected and organized in folders according to their characteristics, compressed files can be used, which enables tree visualization.

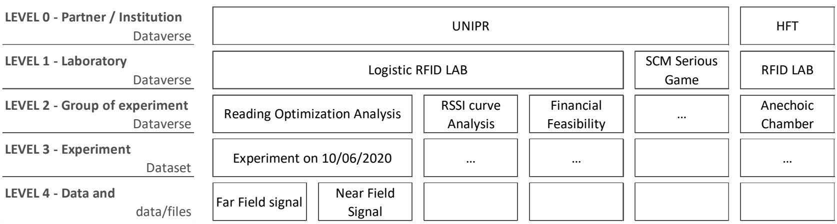

Figure 1: Hierarchical structure for storing Dataverses and data sets

The hierarchical structure so formalized has been discussed by a panel of 8 experts in the field of education and data management from HFT, BIBA, and Parma. The installation and the efficiency of the structure have been discussed, and no concerns arose. Therefore, the demo of the DigiLab4U Dataverse has been officially presented and validated by the experts. The verification of its functionalities and its approval has been achieved, and the system has been recognized as suitable with respect to the original requirements of the DigiLab4U and is now approaching the Go-Live phase.

5 Discussion and Conclusions

In this paper, a hierarchical structure for RDM in Dataverse is presented, referring to the DigiLab4U case. The novelty presented in this paper refers to a more efficient and organized way to store data on the RDM platform Dataverse by means of a five-level structure. This has been chosen as the best compromise between the redundancy of Dataverses and levels of detail in the data catalog. This structure has been verified with respect to (i) the simplicity of the query for retrieving data, and (ii) the suitability of the data catalog structure fostering data consultation. Although this is not evidenced by the paper, it has been discussed, with the several experts involved in the validation and verification process, that companies could be interested in acquiring research results that mostly fit their needs, creating room for the supposed financial sustainability of labs and the network. Future works could analyze this topic, and authors are working on this.

References

Amorim, R. C., Castro, J. A., da Silva, J. R., & Ribeiro, C. (2015). A comparative study of platforms for research data management: Interoperability, metadata capabilities and integration potential. *Advances in Intelligent Systems and Computing*, *353*, 101–111. https://doi.org/10.1007/978-3-319-16486-1_10

Corrall, S., Kennan, M. A., & Afzal, W. (2013). Bibliometrics and research data management services: Emerging trends in library support for research. *Library Trends*, *61*(3), 636–674. https://doi.org/10.1353/lib.2013.0005

Cox, A. M., Kennan, M. A., Lyon, L., & Pinfield, S. (2017). Developments in research data management in academic libraries: Towards an understanding of research data service maturity. *Journal of the Association for Information Science and Technology*, *68*(9), 2182–2200. https://doi.org/10.1002/asi.23781

Esposito, G., Kammerlohr, V., Reverberi, D., Rizzi, A., Romagnoli, G., & Bisaschi, F. (2021). Business Model validation for a marketplace of lab network initiatives. *Proceedings of the 26th Summer School "Francesco Turco"*.

Esposito, G., Mezzogori, D., Reverberi, D., Romagnoli, G., Ustenko, M., & Zammori, F. (2021). Non-Traditional Labs and Lab Network Initiatives: A Review. *International Journal of Online & Biomedical Engineering*, *17*(5).

Galli, M., Mezzogori, D., Reverberi, D., Uckelmann, D., Ustenko, M., & Volpi, A. (2020). Digilab4u: General architecture for a network of labs. *Proceedings of the 25th Summer School "Francesco Turco"*.

Heradio, R., de La Torre, L., Galan, D., Cabrerizo, F. J., Herrera-Viedma, E., & Dormido, S. (2016). Virtual and remote labs in education: A bibliometric analysis. *Computers & Education*, *98*, 14–38.

Kammerlohr, V., Pfeiffer, A., & Uckelmann, D. (2021). Digital Laboratories for Educating the IoT-Generation Heatmap for Digital Lab Competences. In M. E. Auer & D. May (eds.), *Cross Reality and Data Science in Engineering. REV 2020. Advances in Intelligent Systems and Computing, vol. 1231* (pp. 3–20). Springer, Cham. https://doi.org/10.1007/978-3-030-52575-0_1

Wilms, L., K., Stieglitz, S., Buchholz, A., Vogl, R., & Rudolph, D. (2018). Do researchers dream of research data management (2018). *Proceedings of the 51st Hawaii International Conference on System Sciences*, 4411–4420.

Zappatore, M., Longo, A., & Bochicchio, M. A. (2015). The bibliographic reference collection GRC2014 for the online laboratory research community. *Proceedings of 2015 12th International Conference on Remote Engineering and Virtual Instrumentation (REV)*, 24–31.

Authors

Giovanni Esposito, PhD
University of Parma
Parco Area delle Scienze, 181
43124 Parma
https://orcid.org/0000-0001-5150-0855
giovanni.esposito@unipr.it

Giovanni Romagnoli, PhD
University of Parma
Parco Area delle Scienze, 181
43124 Parma
https://personale.unipr.it/en/ugovdocenti/person/96588
giovanni.romagnoli@unipr.it

Davide Reverberi
University of Parma
Parco Area delle Scienze, 181
43124 Parma
https://orcid.org/0000-0001-6768-3932
davide.reverberi@unipr.it

Riccardo Ghinzelli
University of Parma
Parco Area delle Scienze, 181
43124 Parma
riccardo.ghinzelli@studenti.unipr.it

Jens Doveren, Birte Heinemann and Ulrik Schroeder

Towards Guidelines for Data Protection and Privacy in Learning Analytics Implementation

Abstract

While leveraging data collected from learners to improve teaching and learning outcomes has an inherently desirable end goal, Learning Analytics researchers have to be aware of data protection policies and the justified desire for privacy while learning when rolling out such data collection efforts. Successful implementation requires knowing legal frameworks, coordinating with the personnel responsible at the individual institution, and clearly and openly communicating the extent and goal of the data collection effort to the learners and teachers.

In this paper, we present existing community guidelines and our own experiences from a rollout of Learning Analytics in the DigiLab4U project.

Keywords

Learning Analytics, Data Protection, Privacy

1 Introduction

The field of Learning Analytics seeks to leverage quantitative data about learning processes to improve teaching efficacy and learning outcomes. These improvements can derive from data directly, e.g., by presenting learners with insights into their own behaviour, or more indirectly, e.g., by informing decisions about how to improve future iterations of a course.

While such improvements to teaching are inherently desirable for students, gathering data about learning processes while they happen requires a certain degree of interference with the learner's privacy. Depending on the specific learning environment, the kind of data collection, as well as the mode of a collection, can vary widely, but examples include clickstream analysis in learning management systems or gaze and movement analysis in virtual reality applications. Some users might perceive such data collection as surveillance, which might in turn have an adverse effect on their willingness to engage with the learning environment or might even lead to them refusing to engage at all.

Another complicating factor in data collection for Learning Analytics purposes, especially when working in an environment involving institutions from different jurisdictions, are different data protection policies and legal requirements. The intricacies of different data protection policies require communication and individual clearing with every institution that is involved.

Regardless of individual regulations, learners must be able to make an informed decision about whether they consent to the collection of their data. From this requirement follow two implications: one technical and one communicational. On the technical side, systems must be designed in a way that respects users' consent or the lack thereof, i.e., they must provide baseline functionality for users that have opted out of data collection, and they must be able to delete user data should consent to be withdrawn. On the communications side, challenges include finding ways to explain what pieces of data are collected and processed to potentially not particularly tech-savvy learners, as well as clearly communicating the potential benefits learners might reap from participation.

We aim to develop guidelines for researchers that intend to implement Learning Analytics data collection in real-world scenarios. These guidelines will be informed by previous work in that field, our own experience in rolling out Learning Analytics in various institutions of a multi-national research consortium, as well as a series of interviews with researchers and students about their expectations towards Learning Analytics, attitudes towards privacy and how they value the trade-off between the two.

2 Background

2.1 Data Protection Regulation

The European Union and its member states have long enacted regulations regarding the processing of personal data. As per (REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL, 2016), all handling of personal data must guarantee the following:

- Lawful processing
- Specified, explicit, and legitimate use
- Protection from secondary use
- Protection from inadequate and irrelevant processing

- Protection from the use of outdated information
- Protection from unnecessarily long data retention

In addition, the General Data Protection Regulation (GDPR) explicitly codifies the following people's privacy rights (What Is GDPR, the EU's New Data Protection Law?, 2020):

1. The right to be informed
2. The right of access
3. The right to rectification
4. The right to erasure
5. The right to restrict processing
6. The right to data portability
7. The right to object
8. Rights in relation to automated decision-making and profiling

The intricacies of how these principles influence the implementation of Learning Analytics are beyond the scope of this paper but are summarized in (Lukarov, 2019), as well as (Hoel et al., 2017). Underneath the overarching European regulation, there is an entire stack of more specific regulations and ordinances from a national down to an institutional level. Considering these highly heterogeneous regulations and institutional practices, the only general advice one could give to the aspiring Learning Analytics implementer is to communicate with the data protection officer responsible.

2.2 Community Experience

We are far from the first group to implement Learning Analytics in a real-world context, and the issue of privacy and data protection looms over all these efforts. After some rather unfortunate learning experiences, such as the failure of the Gates-funded inBloom (Singer, Natasha, 2014), where overzealous and ill-communicated collection of learning data from sensitive subjects led to a very public backlash, the Learning Analytics community has developed guidelines and checklists for effective communication and stakeholder involvement, such as the DELICATE checklist in (Drachsler & Greller, 2016) and (Schumacher & Ifenthaler, 2018).

The privacy implications of different technology stacks and processing methods are the subject of ongoing research in the Learning Analytics community, having led to the use of such elaborate methods as differential privacy (Steil et al., 2019). A survey of the available literature can be found in (Ciordas-Hertel et al., 2019).

Apart from careful technology choice, open and honest communication is the most important ingredient to a successful rollout of Learning Analytics. That entails explaining what is stored, for how long, and how it is being used, but also making sure learners see a tangible benefit from having their data analysed.

3 Own Contributions

3.1 Choice of Data Warehousing Solution

When gathering data from learners, the technical implementation of how that data is stored is of particular importance when considering data ownership and privacy. In order to be able to make any guarantees with regard to retention policies, deletion of data upon user request, and such, a thorough understanding of the selected data warehousing solution is required.

As it would be dishonest to promise users properties that we ourselves have no way of ensuring, we only considered self-hosted solutions that are free and open source. All hosting and maintenance of the data warehousing solution in the DigiLab4U project were done at RWTH Aachen University.

Although that might seem counterintuitive to those who have never implemented a software system with data persistence, architecting such a system in a way that ensures data integrity, prevents accidental loss of data, and enables arbitrary user data to be deleted at will is a surprisingly tricky endeavour. Hence, many implementations of xAPI learning record stores do not allow the true deletion of data, which we did not deem satisfactory.

These considerations and a thorough survey of the options available led to the choice of *Learning Locker* as the learning record store in the DigiLab4U project. A more detailed description of the decision-making process, as well as an overview of the privacy implications of many ready-to-use Learning Analytics data warehousing solutions on the market can be found in (Lukarov et al., 2020).

3.2 Stakeholder Survey

One of the lab environments enhanced with Learning Analytics as part of the DigiLab4U project is the RFID measuring chamber at HFT Stuttgart. In late 2021, the students of the bachelor program in Information Logistics were asked to take a survey containing questions of interest to various research groups in the DigiLab4U project. We were particularly interested in students' pre-existing experience with Learning Analytics, the value they see in it, as well as their attitude toward sharing their data to enable Learning Analytics not only for them personally but also for fellow students.

Of the 41 participants, 34 filled in the survey completely. The results are hence to be treated more as anecdotal data but can nevertheless give us pointers on how to improve the rollout of Learning Analytics in future courses.

The two aspects relevant to the acceptance of data collection for Learning Analytics purposes in this context are the perceived value for the learners and whether it is great enough to overcome an inherent reluctance towards data sharing. In order to judge the degree to which attitudes toward Learning Analytics results might be tainted by a general lack of statistical literacy, we asked participants to specify their pre-existing knowledge in statistics in general and data visualization interpretation in particular.

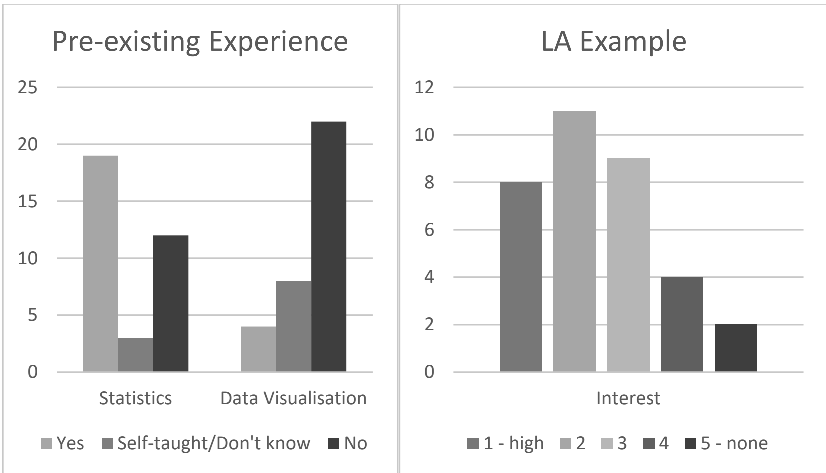

Figure 1 – Student Survey Results

As regards experience with statistics, of the 34 participants, 19 reported having taken a statistics course, 3 reported to be self-taught, and 12 reported no experience. In that same cluster of questions, 4 out of 34 students explicitly reported experience with interpreting data visualizations, 8 did not know, and 23 reported no experience.

Presented with a fictional example of data visualizations for learning feedback and asked whether they would find such feedback interesting, 8 respondents gave a 1 (very interested) on a five-step Likert scale, 11 gave a 2, 9 gave a 3, 4 gave a 4, and the remaining 2 gave a 5. When asked whether they would be willing to share their learning data to help in the generation of such feedback, only one participant gave an unconditionally positive response, 20 specified that they would require their data to be anonymized, and 13 did not give any response.

4 Closing Thoughts

The results of the learner survey suggest a certain degree of hesitation towards sharing their data. One possible cause may be a failure to see how they might personally benefit from feedback generated using Learning Analytics. Over 40% of participants reported only middling or no interest in the examples that were provided, which might explain the lack of enthusiasm.

These experiences once again underline the importance of early communication with learners as stakeholders. Only they can articulate their needs and the value that they attach to any given form of feedback, which in turn must be evaluated by teachers with respect to its didactical value.

Another factor that was beyond the scope of this paper is the issue of scaling up Learning Analytics infrastructures—moving from a smaller, more experimental rollout to a larger, institution-wide one often requires the use of different, industry-grade big data processing frameworks, which again come with their own privacy considerations. The bigger such a project grows, the higher the incentive for standardization, which on one hand facilitates collaboration and exchange of knowledge, but on the other hand, to a certain degree limits technological choices.

Acknowledgements

This work has been funded by the German Federal Ministry of Education and Research (BMBF) through the project DigiLab4U (No. 16DH2114).

References

Ciordas-Hertel, G.-P., Schneider, J., Ternier, S., & Drachsler, H. (2019). Adopting Trust in Learning Analytics Infrastructure: A Structured Literature Review. Journal of Universal Computer Science, 25(13), 1668–1686. https://doi.org/10.3217/jucs-025-13-1668

Drachsler, H., & Greller, W. (2016). Privacy and analytics: It's a DELICATE issue a checklist for trusted learning analytics. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge — LAK '16*, 89–98. https://doi.org/10.1145/2883851.2883893

Hoel, T., Griffiths, D., & Chen, W. (2017). The Influence of Data Protection and Privacy Frameworks on the Design of Learning Analytics Systems. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 243–252. https://doi.org/10.1145/3027385.3027414

Lukarov, V. (2019). *Scaling up learning analytics in blended learning scenarios*.

Lukarov, V., Ehlenz, M., & Schroeder, U. (2020). Building a data warehouse for multimodal learning analytics research projects. *Companion Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, 25–28.

REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. (2016). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2016:119:FULL&from=EN

Schumacher, C., & Ifenthaler, D. (2018). Features students really expect from learning analytics. *Computers in Human Behavior*, *78*, 397–407. https://doi.org/10.1016/j.chb.2017.06.030

Singer, Natasha. (2014, April 21). InBloom Student Data Repository to Close. *The New York Times*. https://bits.blogs.nytimes.com/2014/04/21/inbloom-student-data-repository-to-close/?_r=0

Steil, J., Hagestedt, I., Huang, M. X., & Bulling, A. (2019). Privacy-aware eye tracking using differential privacy. *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*, 1–9. https://doi.org/10.1145/3314111.3319915

What is GDPR, the EU's new data protection law? (2020). https://gdpr.eu/what-is-gdpr/