



Workshop: Towards Open Data Practices at the International Conference on the Internet of Everything

Natalie Kiesler^(✉) 

DIPF Leibniz Institute for Research and Information in Education,
Frankfurt am Main, Germany
kiesler@dipf.de

Abstract. The Internet of Everything (IoE) has the potential to transform the way we live, work, and interact with the world around us. It involves the connection of billions of devices and sensors, generating vast amounts of data that can be used to inform decision-making, improve efficiency, and drive innovation. However, realizing the full potential of the IoE requires the adoption of open data practices, which can increase transparency, provenance, and accountability. This workshop aims to discuss pathways toward open data practices in the context of the IoE by bringing together practitioners in the field. The workshop will include intensive discussions with participants about their experiences with the publication of research data (including software) in the IoE context, but also challenges, concerns, and barriers to publishing research data. The workshop further addresses possible publication formats, how and where to publish data, questions of data ownership, as well as common formats and types. By discussing the state-of-the-art within the IoE community, we intend to draw conclusions on a path forward towards open data practices at venues, such as the IoE Con, as these practices have implications for future conference submissions, tracks, and review processes. Finally, by promoting open data practices, we can unleash the full potential of the IoE to create a more connected, smarter, and sustainable future.

Keywords: Internet of Everything · Open Data · Open Science · FAIR principles

1 Introduction

The Internet of Everything (IoE) can be defined as “*a distributed network of connections between people, smart things, processes, and data, whereas these components interact and exchange real-time data*” [13]. The connections between these four pillars comprise people-to-people (P2P), machine-to-machine (M2M), and people-to-machine (P2M) systems. Hence, technologies within the IoE context

are data-intensive with respect to the exchange of information, and the development of new technologies, interfaces, and systems, as they constantly produce research data, including software.

Unfortunately, publishing research data is still a challenge in many scientific disciplines [3]. Especially the submission of research data including software is currently subject to discussions within several areas of computing [2, 8–10]. Among the challenges are the lack of recognition [15], and the lack of FAIRness [1, 6, 16, 19] for primary researchers who need to invest great efforts to produce mature data. At the same time, the majority of published research data (including software) is based on natural language reports limited in length, automatically lacking depth and detail, and adding ambiguity. Further demands towards publishing research data may comprise high (code) quality, documentation [11, 14], or method reports [12] – all of which take time to produce. By the time a researcher reaches the desired maturity, deadlines may have passed, or technologies may be outdated. Moreover, guidelines or procedures for the peer review of software, huge data sets, and/or other technologies have not yet been established. After all, the benefits of publishing such research artifacts remain low. The same is true for the context of the Internet of Everything (IoE) as an emerging field.

The other extreme may be that research data are not even required for a conference publication or journal and that the peer review of submitted data is purely optional [16]. Instructions for authors and reviewers often state: “Reviewers are not required to review your supplementary materials, your paper submission must stand on its own” [4], or “Papers are expected to stand alone, and not to require the reading of supplementary information” [7]. These instructions can be considered a stimulus for authors not to submit their research data at all, not even for the camera-ready version. Eventually, it can lead to less transparent and less comprehensive research studies and publications, which are challenging to reproduce or replicate.

With the International Conference on the Internet of Everything (IoE Con), a new conference is being established that will help shape the direction of this area, and how the community will handle respective research data including software and IT infrastructure. It is, therefore, crucial to start developing guidelines and procedures for (1) the submission of research data to conferences like the IoE Con, (2) the review of research data by peers, and (3) the publication of research data within certain conference tracks and proceedings, or perhaps even alternative formats. Designing a path towards open data practices now will thus strengthen and improve research in the IoE field and the computing community in the long run.

2 Goals of the Workshop

As software and research data are not a byproduct of the epistemological process, it only appears natural to make data available to other researchers, e.g., by applying the FAIR principles [19] and making data findable, accessible, interoperable, and reusable. At the same time, open data strategies and guidelines have

been promoted for years in the U.S. and in Europe [5, 17, 18]. However, it is still challenging to transfer these principles to software and technology [8, 9], which is why we need to involve the research community and consider their challenges and barriers.

As one of the first steps on a pathway towards open data practices at IoE Con, the facilitator of this workshop intends to address the following objectives:

1. Increase the community's consciousness of the role and impact of open data practices.
2. Gather the community perspective with regard to challenges, barriers, and possible strategies to overcome them.
3. Develop realistic, and comprehensible requirements for the publication of research data (including software) at IoE Con.
4. Start designing peer review guidelines for the IoE Con program committee.
5. Discuss new publication formats or tracks in the context of the IoE Con.

The expected result of the workshop is the exploration of realistic publication formats and peer review procedures for the IoE community. Gathering this perspective will help inform the IoECon program committee on how to realize open data practices in the near future.

3 Workshop Schedule

The workshop is intended to last three hours and address the aforementioned goals by discussing and elaborating on them with the workshop's participants. The activities will be arranged as follows:

-
- Provide an initial input from the literature about the well-known challenges and barriers for researchers (15 min)
 - Give insights into existing review processes from other conferences (15 min)
 - Discuss the community's prior experience with regard to the publication of research data (including software) based on a set of guiding questions (45 min)
 - Short break (15 min)
 - Use expert groups (or alternatively breakout sessions) to discuss the following three guiding questions (30 min):
 - What are realistic, comprehensive requirements for the publication of research data at IoECon?
 - What are feasible peer review guidelines that can serve as recommendations for the IoECon program committee?
 - Consider new publication formats or tracks for IoECon.
 - Presentation and discussion of each group's input and ideas to derive concrete recommendations and a path for action (45 min, with 3 groups à 15 min)
 - Summary of the main takeaways and wrap-up of the workshop (15 min)
-

The workshop may be realized online or in presence, although participation may be higher in an online format. In that case, discussions and group work will be realized via Zoom breakout rooms.

4 Expected Results and Publication

This new workshop format serves as a networking opportunity within the IoE community, as it allows researchers to connect and exchange their ideas and concerns related to open data and open science practices.

Moreover, all groups will be asked to document their discussions in shared documents (e.g., on a flipchart or in an online file). Moreover, the facilitator of the workshop will track the community's input during the entire workshop, especially as part of the presentation and discussion of the subgroups' results, and as part of the summary of take-aways.

This workshop will further serve as a basis for a submission to next year's IoE Con, as the facilitator will compile the community's ideas and course for action into a workshop report. Upon positive review by the program committee, it will be included in the 2024 conference proceedings.

5 Conclusion

The proposed workshop on open data practices at the International Conference on the Internet of Everything aims to address the challenges and barriers of publishing research data including software from the perspective of IoE researchers and practitioners. The IoE is a data-intensive field that develops new technologies and produces huge amounts of research data. We, therefore, need to start developing a path forward to mitigate the well-known challenges of data publications due to the lack of recognition and FAIRness. The workshop's goals are to increase the community's consciousness of open data practices, gather the community's experience with publishing research data, develop realistic and comprehensible requirements for software publication, start designing peer review guidelines, and discuss new publication formats. The expected result is the exploration of realistic publication formats and recommendations for peer review procedures for the IoE community, helping inform the IoECon program committee on how to realize open data practices in the near future. To conclude, the proposed workshop is an essential step toward open data practices, and it will strengthen and improve research in the IoE field.

References

1. Barker, M., et al.: Introducing the fair principles for research software. *Sci. Data* **9**(1), 622 (2022). <https://doi.org/10.1038/s41597-022-01710-x>
2. Beardsley, M., Hernández-Leo, D., Ramirez-Melendez, R.: Seeking reproducibility: assessing a multimodal study of the testing effect. *J. Comput. Assist. Learn.* **34**(4), 378–386 (2018). <https://doi.org/10.1111/jcal.12265>
3. Borgman, C.L., Pasquetto, I.V.: Why data sharing and reuse are hard to do (2017). <https://escholarship.org/uc/item/0jj17309>
4. CHI: Papera - Quick Facts (2022). <https://chi2022.acm.org/for-authors/presenting/papers/>

5. European Union: European open science cloud (2023). <https://eosc-portal.eu/>
6. Hong, N.P.C., et al.: Fair principles for research software (FAIR4RS principles) (2022). <https://doi.org/10.15497/RDA00065>
7. ITiCSE: Call for papers (2023). <https://iticse.acm.org/2023/call-for-papers/>
8. Jay, C., Haines, R., Katz, D.S.: Software must be recognised as an important output of scholarly research. arXiv preprint [arXiv:2011.07571](https://arxiv.org/abs/2011.07571) (2020)
9. Katz, D.S., Gruenpeter, M., Honeyman, T.: Taking a fresh look at fair for research software. *Patterns* **2**(3), 100222 (2021). <https://doi.org/10.1016/j.patter.2021.100222>
10. Katz, D.S., Psomopoulos, F.E., Castro, L.J.: Working towards understanding the role of FAIR for machine learning. In: DaMaLOS, pp. 1–6 (2021). https://www.danielskatz.org/papers/FAIR4ML_for_DAMALOS.pdf
11. Kiesler, N.: Dataset: recursive problem solving in the online learning environment codingbat by computer science students (2022). <https://doi.org/10.21249/DZHW:studentsteps:1.0.0>
12. Kiesler, N.: Daten- und Methodenbericht Rekursive Problemlösung in der Online Lernumgebung CodingBat durch Informatik-Studierende. Technical report (2022). [https://metadata.fdz.dzhw.eu/public/files/data-packages/stu-studentsteps\\$/attachments/studentsteps_Data_Methods_Report_de.pdf](https://metadata.fdz.dzhw.eu/public/files/data-packages/stu-studentsteps$/attachments/studentsteps_Data_Methods_Report_de.pdf)
13. Kiesler, N., Impagliazzo, J.: Perspectives on the internet of everything. In: Pereira, T., Impagliazzo, J., Santos, H. (eds.) IoECon 2022. LNICST, vol. 458, pp. 3–17. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-25222-8_1
14. Kiesler, N., Pfüll, B.: Higher education programming competencies: a novel dataset. In: Iliadis, L., Papaleonidas, A., Angelov, P., Jayne, C. (eds) ICANN 2023. LNCS, vol. 14261, pp. 319–330. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-44198-1_27, <https://github.com/nkiesler-cs/HEPComp-Dataset>
15. Kiesler, N., Schiffner, D.: On the lack of recognition of software artifacts and IT infrastructure in educational technology research. In: Henning, P.A., Striewe, M., Wölfel, M. (eds.) 20. Fachtagung Bildungstechnologien (DELFI), pp. 201–206. Gesellschaft für Informatik e.V., Bonn (2022). <https://doi.org/10.18420/delfi2022-034>
16. Kiesler, N., Schiffner, D.: Why we need open data in computer science education research. In: Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education, ITiCSE 2023, vol. 1. Association for Computing Machinery, New York (2023). <https://doi.org/10.1145/3587102.3588860>
17. National Science Foundation: Open data at NSF (2013). <https://www.nsf.gov/data/>
18. Pilat, D., Fukasaku, Y.: OECD principles and guidelines for access to research data from public funding. *Data Sci. J.* **6**, OD4–OD11 (2007). <https://www.oecd.org/sti/inno/38500813.pdf>
19. Wilkinson, M.D., et al.: The fair guiding principles for scientific data management and stewardship. *Sci. Data* **3**(1), 1–9 (2016). <https://doi.org/10.1038/sdata.2016.18>