

GROWING OPEN SCIENCE WITH THE COMBINED POTENTIAL OF CITIZEN SCIENCE AND AUTO SCIENCE

Sonja Schimmler^{1,3}, Fabian Kirstein^{1,3}, Sebastian Urbanek^{1,3},
Hannes Wünsche^{1,3} & Manfred Hauswirth^{1,2,3}

¹Weizenbaum Institute for the Networked Society

²TU Berlin

³Fraunhofer FOKUS

Berlin, Germany

{first.last}@fokus.fraunhofer.de

Abstract

In this paper, we present our ideas on how to best support researchers in every phase of the research process when dealing with their research data.

We propose a *Research Data Portal* as the central data infrastructure. With the help of this portal, a researcher can easily manage and update his or her research data, share it with collaborators, and reach out to the public.

We further propose a *Citizen Science Portal*, which includes some new and innovative concepts and methods. In this portal, Citizen Science and Auto Science concepts are applied, and support to bring together the best of both worlds is provided. *Citizen Science* promises to entail the individual (scientists and hobby scientists) to help with research. *Auto Science* is meant to help analyze research data, e.g., to help publish the data and to help improve its quality, by applying methods from artificial intelligence.

Keywords

Research Data; Open Science; Citizen Science; Auto Science

This work has been funded by the Federal Ministry
of Education and Research of Germany (BMBF)

1 Introduction

The Weizenbaum Institute is an interdisciplinary research institute committed to Open Science. In the current start-up phase of the institute, we try to gain an understanding of a researcher's daily life and try to actively support it: (i) We need to understand how the researchers' individual research processes look like, what tools they use, and what research data they produce. (ii) In parallel, we try to design an Open Science-friendly environment, and try to motivate and support the researchers when working in this environment.

Our goal is to achieve *Open Science by Design*, namely to install Open Science as an integral part of our institute's culture.

This has to be done on an organizational as well as on a technical level. In this paper, we will concentrate on the latter. As shown in Figure 1, we will focus on the support for handling research data, and on how scientists and hobby scientists could be better involved.

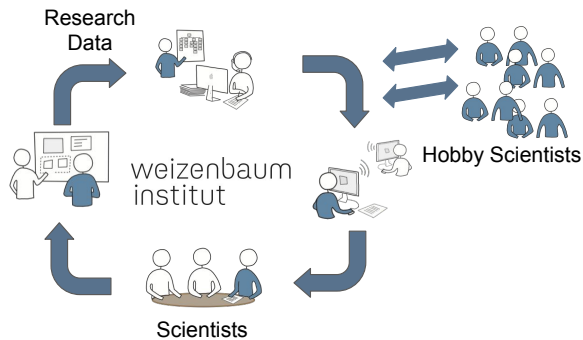


Figure 1: Scientists & Hobby Scientists

2 Vision

Our idea is to create an *Open Science Portal* (see Figure 2), which is tailored to the needs of the Weizenbaum Institute. The key idea is that we design and use this portal in our daily research. It mainly consists of two parts – a *Research Data Portal* and a *Citizen Science Portal*.

2.1 Research Data Portal

We see a *Research Data Portal* as the central point of a researcher's daily life. With the help of the portal, a researcher can easily manage and update his or her research data, share it with collaborators, and reach out to the public.

One essential feature is that it will provide a simple way of exchanging data with other tools, the scientist uses. One of our key design decisions is that a researcher will have a *Researcher's Identity*, and all of his or her research output is tied to this identity, in order to support data sovereignty and data provenance.

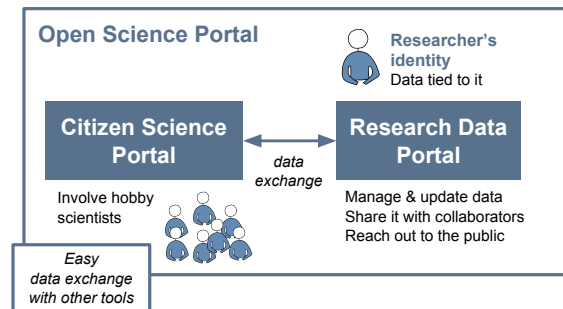


Figure 2: Open Science Portal for the Weizenbaum Institute

Challenges (i) While most research data platforms are tailored to a specific discipline, the platform is planned for an interdisciplinary environment. The main challenge here will be to find a good balance between generality and specificity.

Current Status In a first step, we have analyzed different open source research data platforms (cf. Wilkinson et al. 2016; Amorim et al. 2017) and evaluated three promising candidates in more detail: DSpace¹, Zenodo² and CKAN³. Based on this evaluation, we decided to use DSpace (Smith et al. 2003), as it

¹ duraspace.org/dspace

² zenodo.org

³ ckan.org

matches most of our requirements, and as it is widely used among our partner institutions. Following this first step, we have started to tailor this open source research data portal to our specific needs. We can base this on our experience in developing open data platforms, e.g., the European Open Data Portal⁴, which was launched 3 years ago, has about 30.000 visitors a month, and encompasses over 860.000 data sets.

In parallel, in order to tackle challenge (i), we will stay in close contact with the different research groups and the different protagonists of the Weizenbaum Institute.

2.2 Citizen Science Portal

To further support researchers when dealing with their data, we want to apply new and innovative concepts and methods, which are integrated in a *Citizen Science Portal*. Our main idea is to apply Citizen Science and Auto Science concepts, and to bring together the best of both worlds. *Citizen Science* promises to entail the individual (scientists and hobby scientists) to help with research. *Auto Science* (Weber 2017) is meant to help analyze research data, e.g., to help publish the data and to help improve its quality, by applying methods from artificial intelligence.

Challenges (i) At the moment, many Citizen Science projects face the challenge that neither a critical mass of hobby scientists, nor people with the “right” background are attracted. (ii) Furthermore, hobby scientists are mostly involved in a crowdsourcing-fashion, namely to collect and clean data. Prominent examples from the USA are Galaxy Zoo (astronomy), Foldit (biochemistry) and Polymath (mathematics) (Franzoni and Sauermann 2014). (iii) Last but not least, new and innovative concepts and methods are needed in

order to realize our *Citizen Science Portal* as envisioned.

Current Status To this point, we have completed our first proof-of-concept implementation of the Citizen Science Portal. In order to address challenge (i), the main idea is that large interactive screens will be placed in public, and that the Citizen Science projects will be executed in this environment. There exist similar ideas from the crowdsourcing-domain, e.g., (Goncalves et al. 2013).

As next steps, we will try to address challenges (ii) and (iii) by experimenting with some new and innovative interaction techniques as well as with Citizen Science and Auto Science concepts. In order to gain first insights, we plan to perform 2-3 Citizen Science projects from the research agenda of the Weizenbaum Institute.

3 Conclusions

In this paper, we have presented our ideas on how to best support researchers in every phase of the research process when dealing with their research data.

From our point of view, the most innovative potential of our proposal lies in the in-depth integration of *Research Data Portal* and *Citizen Science Portal*, and in the combination of Citizen Science and Auto Science concepts.

⁴ www.europeandataportal.eu

4 References

1. Amorim, Ricardo Carvalho et al. (2017). “A Comparison of Research Data Management Platforms: Architecture, Flexible Metadata and Interoperability”. In: *Universal Access in the Information Society* 16.4, pp. 851–862.
2. Franzoni, Chiara and Henry Sauermann (2014). “Crowd Science: The Organization of Scientific Research in Open Collaborative Projects”. In: *Research Policy* 43.1, pp. 1–20.
3. Goncalves, Jorge et al. (2013). “Crowdsourcing on the Spot: Altruistic Use of Public Displays, Feasibility, Performance, and Behaviours”. In: *2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, pp. 753–762.
4. Smith, MacKenzie et al. (2003). “DSpace: An Open Source Dynamic Digital Repository”. In: *D-Lib Magazine* 9.1.
5. Weber, Silke (2017). “Der Start-up Mediziner (Madisch, Research Gate)”. In: *Zeit Online*.
6. Wilkinson, Mark et al. (2016). “The FAIR Guiding Principles for Scientific Data Management and Stewardship”. In: *Scientific Data* 3.