1st Conference on Research Data Infrastructure

Humanities and Social Sciences

https://doi.org/10.52825/CoRDI.v1i.393

Published: 07 Sept. 2023

The Open Data Infrastructure for Social Science and Economic Innovations (ODISSEI)

Better Infrastructure, Better Science, Better Society

Tom Emery[1,2][https://orcid.org/0000-0001-6137-9577], Kasia Karpinska[1,2][https://orcid.org/0000-0003-3208-8046], Angelica Maineri[1,2][https://orcid.org/ 0000-0002-6978-5278], and Lucas van der Meer[1,2][https://orcid.org/0000-0003-4415-678X]

[1] Erasmus University Rotterdam – Erasmus School of Social and Behavioral Sciences

[2] Open Data Infrastructure for the Social Science and Economic Innovations (ODISSEI)

Abstract. The Open Data Infrastructure for Social Science and Economic Innovations (ODISSEI) equips social scientists in the Netherlands with the data, tools, and skills that are necessary to answer groundbreaking questions for scientific and policy making purposes. With a variety of use cases to pick from, we aim at engaging in a discussion with other Research data infrastructures to identify synergies but also challenges ahead.

Keywords: Social Sciences; FAIR; data infrastructure.

1. Extended abstract

The Open Data Infrastructure for Social Science and Economic Innovations (ODISSEI) equips social scientists in the Netherlands with the data, tools, and skills that are necessary to answer groundbreaking questions for scientific and policy making purposes. ODISSEI aspires to create a federated data infrastructure which seamlessly connects researchers with the data and facilities provided by member organisations. The experience gained by ODISSEI in the past six years contributes to a wider discussion on the role of Research Data Infrastructure in the Social Sciences. ODISSEI's programme of work is subdivided into four work streams (see Figure 1) which represent four separate ways in which ODISSEI serves the research community [1].

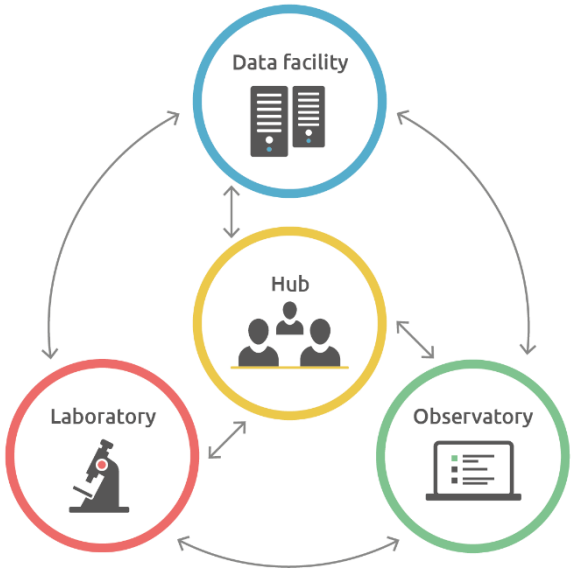

Figure 1. Conceptual representation of ODISSEI Infrastructure

First, the Data Facility ensures that researchers can find, access, and link the data that they need. At the centre of this work is the ODISSEI Secure Supercomputer which allows researchers to analyse complex and rich data from Statistics Netherlands and other ODISSEI data providers in a secure, yet computationally powerful environment (see examples of research conducted on the OSSC in [2] and [3]). ODISSEI also provides grants and support to access this data to facilitate new research. Moreover, the ODISSEI Portal makes diverse data sources available through a unique search interface by leveraging metadata [4]. ODISSEI is also leading the development of a generic Secure Analytical Environment (SANE) which makes it possible for data providers to securely share data with researchers whilst retaining overall control of data [5].

The Observatory supports and maintains key long-standing data collection efforts and participation in the international social science data collections. These include such studies as the European Social Survey, the Survey of Health, Ageing and Retirement in Europe and the Generations and Gender Programme. It also covers the Dutch Election Study which has been collecting data since 1971. This work stream is focused on providing a consistent, stable, and reliable stream of data for social scientists that could then be utilised across the infrastructure.

The Laboratory is where researchers can conduct their own experiments, primarily through the LISS panel, operated by Centerdata. ODISSEI provides financing for the core LISS panel but also provides access to researchers from ODISSEI member organisations to field their own questions to the LISS panel's representative and high-quality sample of over 4,500 households (see e.g. [6] and [7]). The Laboratory is also where future ODISSEI upgrades and enhancements are developed and prototyped.

Finally, the Hub is where researchers are provided with support, expertise, and guidance in the use of ODISSEI services and facilities. It includes an educational programme, community events, remote access grants, data stewardship, as well as a Social Data Science (SoDa) Team at Utrecht University who can provide high quality and intensive support to researchers looking to deploy computational and data science methods within ODISSEI.

As a community, ODISSEI not only delivers services but also develops and promotes standards and best practices for social science research. ODISSEI not only promotes the FAIR principles through the delivery of new search and access services, but also by requiring

adherence to FAIR from ODISSEI users. ODISSEI requires its users to act in accordance with the principles of responsible data science: Fair, Accurate, Confidential and Transparent (FACT). ODISSEI also supports open science by facilitating inclusion, sharing, and equity through its work [8].

ODISSEI also established, develops, and maintains a FAIR Expertise Hub to support communities of data providers in improving their FAIRness. An important instrument for the FAIR Expertise Hub is the FAIR Implementation Profile (FIP), a collection of decisions and plans made by a community about how to achieve FAIRness [9, 10]. A FIP comes with an easy to use wizard and accompanying workshop provided by GO-FAIR. The hub helps data communities in (1) establishing their plans, (2) to agree on their FAIR-enabling resources, and to (3) achieve a substantial increase of FAIRness. The project partners will (4) create alignment with international standards and (5) between communities. Explicit FAIR Implementation Profiles (6) facilitate software developers.

Looking ahead, ODISSEI will collaborate with CLARIAH (the Dutch infrastructure for digital humanities) in the SSHOC-NL project (the full proposal is available online, see [11]). SSHOC-NL aspires to develop an SSH-wide digital infrastructure that is interoperable, allowing its data, tools and services to be shared, linked, and combined in imaginative and ground-breaking ways.

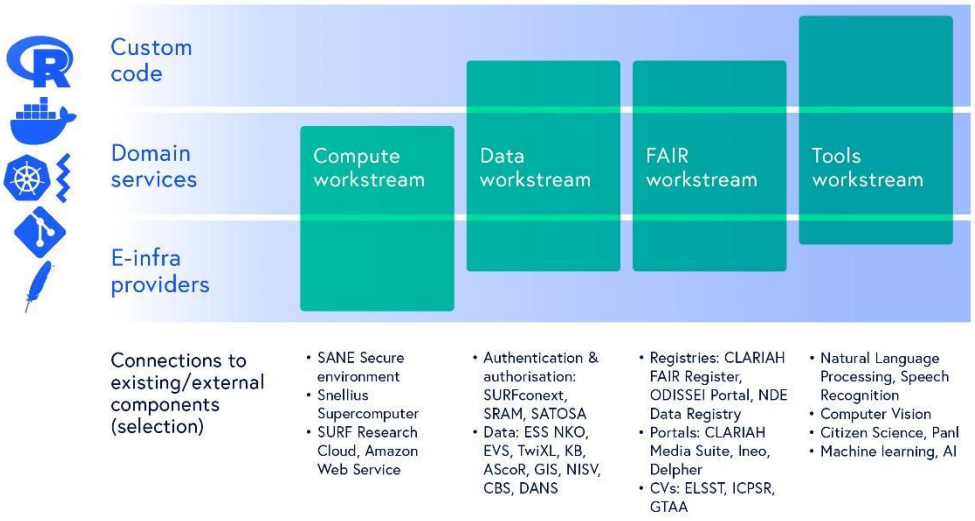

Figure 2. SSHOC-NL Architecture (see [11]).

With a variety of use cases to pick from, we aim at engaging in a discussion with other Research data infrastructures to identify synergies but also challenges ahead.

Data availability statement

The submission is not based on data.

Competing interests

The authors declare that they have no competing interests.

Funding

ODISSEI received a grant by the Dutch Research Council (NWO) under grant number 184.035.014.

Acknowledgement

We are grateful to all the partners in the ODISSEI project and to the members of the Management, Supervisory and Advisory Board.

References

1. T. Emery, P. Dykstra, and L. van der Meer, 'ODISSEI Annual Report 2021-2022', Zenodo, Dec. 2022. doi: https://doi.org/10.5281/zenodo.7413577.
2. E. L. de Zeeuw et al., 'Safe Linkage of Cohort and Population-Based Register Data in a Genomewide Association Study on Health Care Expenditure', Twin Research and Human Genetics, vol. 24, no. 2, pp. 103–109, 2021, doi: https://doi.org/10.1017/thg.2021.18.
3. A. Petrović, 'Multiscale spatial contexts and neighbourhood effects', Delft University of Technology, 2020. Accessed: Nov. 25, 2021. [Online]. Available: https://journals.open.tudelft.nl/abe/article/view/5194
4. T. Emery, R. Braukmann, M. Wittenberg, J. van Ossenbruggen, R. Siebes, and L. van de Meer, 'The ODISSEI Portal: Linking Survey and Administrative Data', Dec. 2020. doi: https://doi.org/10.5281/zenodo.4302096.
5. L. van der Meer, 'SANE - Secure ANalysis Environment @PDI-SSH', Jan. 26, 2023. doi: https://doi.org/10.5281/zenodo.7562026.
6. M. A. Yerkes, C. Remery, S. André, M. Salin, M. Hakovirta, and M. van Gerven, 'Unequal but balanced: Highly educated mothers' perceptions of work–life balance during the COVID-19 lockdown in Finland and the Netherlands', Journal of European Social Policy, p. 09589287221080411, Mar. 2022, doi: https://doi.org /10.1177/09589287221080411.
7. D. Yakar, Y. P. Ongena, T. C. Kwee, and M. Haan, 'Do People Favor Artificial Intelligence Over Physicians? A Survey Among the General Population and Their View on Artificial Intelligence in Medicine', Value in Health, Oct. 2021, doi: https://doi.org/10.1016/j.jval.2021.09.004.
8. User policy
9. E. Schultes, B. Magagna, K. M. Hettne, R. Pergl, M. Suchánek, and T. Kuhn, 'Reusable FAIR Implementation Profiles as Accelerators of FAIR Convergence', in Advances in Conceptual Modeling, G. Grossmann and S. Ram, Eds., in Lecture Notes in Computer Science. Cham: Springer International Publishing, 2020, pp. 138–147. doi: https://doi.org/10.1007/978-3-030-65847-2_13.
10. A. M. Maineri and S. Wang, 'FAIR yes, but how? FAIR Implementation Profiles in the Social Sciences', Dec. 2022, doi: https://doi.org/10.5281/zenodo.74284
11. P. Dykstra et al., 'Social Science and Humanities Open Cloud for the Netherlands (SSHOC-NL)', Feb. 2023, doi: https://doi.org/10.5281/zenodo.7645356.