

Openness and Research Infrastructures: review of the recent FIRI grant and interview with Prof. Leif Schulman (FinBIF)

In January 2015 the Academy of Finland has granted 19.3 million euros in funding to Finnish research infrastructures, FIRI. About 60 potential research infrastructure projects applied for two FIRI funding calls launched in 2014. These calls consequently targeted at high priority national research infrastructures and national ESFRI nodes. With the success rate of 27%, forty-eight projects were successfully supported by this grant.

In large context, any modern-day research infrastructure (RI) presents a bundle of complex relations and connections between people, technologies and high society' demands. Research data, scientific instrumentation, ICT services are tightly coupled with high-level professional skills and knowledge. That's why all these projects are quite various and refer to different aspects of RIs. Among the FIRI grant's winners you could find libraries and databases (*Developing the Research Services of the Saami Culture Archive of University of Oulu*); biological archives (*Finnish Biodiversity Information Facility – FinBIF*), integrated upgrade of research equipment (*Biocenter Finland* of the University of Helsinki), high-speed communication networks and highly distributed capacity and capability computing facilities (e.g., *Finnish Grid and Cloud Infrastructure (FGCI)* and *Upgrading of the Finnish National Seismological Network – a subproject of FIN-EPO*); data infrastructure together with observation and laboratory facilities (*Integrated Atmospheric and Earth System Science Research Infrastructure (INAR RI)*); advanced facilities for high energy physics (*Detector systems for nuclear and particle physics at FAIR and CERN*), as well as infrastructural centers of competence which provide a service for the wider research community based on an assembly of techniques and know-how (e.g., *Development of Otaniemi Micro- and Nanotechnology Infrastructure - phase II*).

All these funded projects have not only been responsible for some of the greatest scientific discoveries and innovation developments, but are also much influential in attracting the best researchers from around the world and in building bridges and networking between national research communities and scientific disciplines making their research data and publications more opened and accessible. Some of the granted RIs are directly tighten to the objectives of openness in science and research. One of them is the Finnish Biodiversity Information Facility ([FinBIF](#)), a national data center for biodiversity information, set up by the Finnish Museum of Natural History through a three-year project in 2015-2017. The project is a component of a larger aggregate project Envibase administrated by the Finnish Environment Institute (SYKE). FinBIF operates as a national service supporting species-related research, management, teaching, and life-long learning. It collects, archives, and distributes Finnish electronic datasets on biodiversity and provides visualization and analysis tools in one single portal. The FinBIF project consortium consists of a few research teams of the University of Helsinki, Turku University, University of Oulu, and University of the Eastern Finland. Professor Leif Schulman, a Director of the Finnish Museum of Natural History of the University of Helsinki, leader of the University of Helsinki group for FinBIF has kindly agreed to participate in our short Q&A tour exploring the current state of open science in the context of research infrastructure.

Q: What are the data and data practices (in storing, preprocessing, analyzing, tracking) in your field of research, biodiversity science? Are there suitable tools, standards or best practices helping with the above?

A: FinBIF creates a novel, integrated e-infrastructure combining three kinds of biodiversity data (specimens, observations, and DNA), links them with modern research tools enabling rapid generation of further biodiversity data, and brings these all together on an open-access basis into virtual e-Labs making possible their simultaneous analysis together with

environmental data. The current, and historical, tradition within the field of FinBIF – biodiversity data – is not completely open. Although many practices within the field have traditionally been in line with modern open data policies, such as the free sharing of specimens and information between natural history museums, there is also a need for a culture reform especially as regards research data. Therefore, FinBIF does not only construct ICT solutions for biodiversity data management but also develops, publicizes and promotes best practices of open data management. It promotes open science and effective use of data by society at large, as well as data quality control, descriptions, and accumulation. FinBIF follows current international trends and best-practice examples in doing this, as well as develops them according to national needs.

Q: On which research infrastructure do you rely for data management and sharing? Do you consider the research infrastructure you use is well prepared to meet your requirements? What are the main bottlenecks or other constraints you could diagnose?

A: FinBIF is a data management and sharing infrastructure in itself. We aim to promote the use of FinBIF among Finnish researchers as a data infrastructure for the data produced in research projects. With the current project funding (1.4 M€ from the Academy of Finland and 2.1 M€ from the Ministry of Finance for the years 2015-2017) there are no immediate cost constraints. Privacy, data ownership, and legal issues will need focus during the project work, but we do not see these as overwhelming obstacles hindering the development of the RI.

Q: What is your expectation of data being reused within your research community and outside of the community? Which values have already been generated by reuse of your data (research discovery, publications etc.)?

A: Finnish biodiversity data (a part of it) is already being distributed via the international service [GBIF](#), which is an open data source. The use of GBIF mediated data is described on GBIF's homepage, so values of re-use can be evaluated on the basis of information provided there. As to FinBIF, we expect the data to be used in both basic and applied research as well as in the national environmental administration for management, conservation, and sustainable use of renewable natural resources and, e.g., the control of invasive alien species (see [vieraslajit.fi](#), which is one example of how FinBIF services will be provided).

Q: In Finland, are any barriers to the more open sharing of data cultural or technical?

A: We see them as almost purely cultural. Naturally technical barriers also exist in the sense that current information systems are not always fit-for-use as reg. open data sharing, but this is merely a question of a lack in implementing existing technical solutions, rather than a question of lack of suitable technologies. Implementing the technology only requires some money and time, but unless the emerging new thinking favoring open data sharing spreads further, open data sharing will not become common place.

Q: What do you feel about the state of the art in open access publishing in Finland? What are impediments in open publishing domain?

A: One impediment is lack of funds reserved for this: in many cases researchers/projects do not have the funds required for open access publishing fees.

Q: Could you indicate some reasons behind reluctance to promote open science among researchers?

A: One is jealousy as reg. "own" data, that is, researchers want to reserve data for their own research use, rather than sharing it. In many cases this extends far beyond reasonable embargo times and is based on unrealistic estimates of coming personal use of the data. In some cases reluctance to share data also comes from a notion that commercial actors (enterprises) should not be allowed to use data collected by researchers free of charge.

Q: Whose responsibility is it to create a more open science?

A: It is a joint endeavor: researchers have a personal responsibility to change their attitudes; research funders should promote open data sharing both through incentives and through compulsory measures; research institutions should develop and implement data policies promoting open sharing; data users should apply ethical data use, e.g., through meticulously following data use licenses and citing data sources, etc.

Yes, we agree here with Prof. Leif Schulman. The coherence of the different actors that determines any research infrastructure is a common agenda for promoting openness. As we have seen from the interview, there are many factors - policy, funds, culture momentum - that shape emerging research infrastructure and its openness capability. FinBIF aims to deploy an e-infrastructure built by federating and integrating existing European and Finnish infrastructures and resources in which all its constituent parts are part of a large whole. But the openness of this large-scale system depends both on its technical potentials and organizational and social frameworks and policies driven at national and international levels. We wish to FinBIF to successfully combine biodiversity science and open e-research and to extend open science practices within the Finnish research community and beyond.

March 2015, by Marina Bouianov (CSC)