

Developing Research Information Hub in Finland, a view from the Ministry

Abstract to euroCRIS2024, 13th Feb 2024

Jukka Haapamäki, Ministry of Education and Culture

1. Introduction

The Research Information Hub is a national service of the Finnish Ministry of Education and Culture. It gathers and shares information on scientific research carried out in Finland in an easily accessible way and format. The Research information hub has four main interlinked goals (Figure 1).



Figure 1. Strategic objectives, benefits and functionalities if the Finnish Research Information Hub.

The visible part of the research information hub is the public Research.fi website. The information displayed on the site and the services provided through it have required a great deal of development and collaboration.

This paper explains how the research information hub came into being and why the Ministry of Education and Culture wanted to implement the service and what are seen as its main goals and benefits nationally.

2. History Part 1: National information basis for higher education institutions

Finland is a relatively small country with 13 universities, 22 universities of applied sciences, 12 state research institutes, two large public research funders, hundreds of private research funding foundations and a number of other research stakeholders. Most of the funding for universities and government research institutes comes from the state [1]. State funding for higher education institutions (HEIs) is based on a set of indicators on the basis of which funding is allocated. For HEIs in particular, publication is a key indicator of funding allocation [2]. Historically, the data used for publication-based funding have been based on data collected by the Ministry of Education and Culture. When the funding share of publications

in the funding model was increased in 2013, the publication data collection was changed from a summary-based data collection to a publication-specific data collection. At the same time, following the examples from Norway and Denmark, Finland introduced the Publication Forum system dividing publication channels into two (later three) levels that partly determined the weight of a publication in the funding model [3]. In practice, universities reported the key metadata for each publication in the Ministry's publication data collection. The data collection was done by entering a CSV form into the Ministry's data collection application.

In the early 2010s, the Virta higher education achievement register was built on top of the study information systems of Finnish higher education institutions. From the beginning, the register had several purposes. It was used for student selection in higher education institutions, where from 2014 onwards applicants could be allocated to quotas according to whether they had a previous higher education entrance qualification. In addition, the Virta register handles several data transfers to authorities and other actors that higher education institutions are obliged to carry out. In addition, the Virta register allowed the Ministry of Education and Culture the flexibility to produce a variety of statistics on higher education. Partly as a by-product, but a very important by-product, the development of the Virta register also created a natural platform for harmonising and improving the quality of data in higher education study information systems.

During the design phase of the Virta study register, the possibility of a similar repository for research information was identified. Thus, in 2015, plans were made to set up a Virta publication information system on top of the publications data collection. The aim was to implement and streamline the collection of statistical data on publications required for the calculation of state funding for higher education institutions, but also to enable the use of publication data for other services. The idea was to implement a service that would make the publications produced by higher education institutions at national level available to various services and other processes through a single, efficient interface service. The Virta publication register was completed at the end of 2016. Soon after that, the research institutions were included in the register. The first service to use the data was Finland's main research funder, the Research Council of Finland (former Academy of Finland), which integrated the register into its own grant system. After the integration, the researcher was able to extract the publications in his/her publication list from the Virta publication register.

The next step was to extend the Virta register to other research metadata. Potential elements were seen as research data, funding decisions, research infrastructures, research organizations, research news, and merit data of researchers. This is how the idea of a Research Information Hub began to take shape.

3. History part 2: Towards wider use of national research information

In March 2017, the Ministry of Education and Culture set up a working group with the following objectives:

- Researchers would no longer have to enter their data into several different systems. The metadata collected from the research will be collected in one place in the national repository.
- Permanent identifiers for both researchers and outputs will facilitate the automation of data transfer and the use of researchers' own data in different services.
- The data will be openly available for use through digital services.
- The user of the data does not need to build a data transfer link to each source system (e.g. from a funder to each university)

- Research governance organisations have access to comprehensive, up-to-date and consistent comparative information on research activities at national level

Table 1. Types and sources of information in the Finnish Research Information Hub

Information type	Who can submit information?	How is information submitted?
Publications	Research organizations	XML file to VIRTa Publication Information Service or directly via JUSTUS – Publication Information Reporting Service
Researchers and other research activities	Research organizations, researchers	XML export file to Research Information Hub or through JUSTUS Individual researchers can import information from their ORCID profile
Funding decisions / grants	Funders	XML export file to Research Information Hub
Funding calls	Funders	Direct data transfers from system suppliers Web form for individual funders
Research data	Research organizations, data services and repositories	Fairdata.fi METAX – Metadata Warehouse's REST API or harvesting repositories Submitting information to national Fairdata.fi Tool Qvain service
Research infrastructures	Research organizations, infrastructure providers	Infrastructure web form or export file
Research news	Research organizations, funders and other bodies supporting science/ research in Finland	Harvested from organizational RSS feed
Projects	Research organizations	XML export file to Research Information Hub

The objectives were later transformed to strategic objectives mentioned in the beginning. The task of the research information hub was to achieve these objectives. One key principle was from the beginning that all data comes from local systems, see table 1. Data is not collected directly into the service.

The first part of the information hub was the research data component. This was relatively straightforward, as the research data were already stored in various repositories and there were existing data models for the metadata of the data. However, it should be noted that survey data is still not stored in a very systematic way, which makes it difficult to find.

The research funding decision data was a slightly more challenging package. Finland has two large public research funders, Research Council of Finland and Business Finland, and several hundreds of large and small private foundations that fund research. The aggregation of data required the implementation of a common data model. In addition, individual foundations had to be technically assisted in their data transfer processes to the research information hub. As a side effect of the import of funding decisions into the information hub, a search functionality for open funding calls was also implemented in Research.fi portal. This replaced a previous service, Aurora database, which had been based on a lot of manual work. Also for research infrastructures, the first task was to build a common data model.

So far, the most challenging task has been the implementation of the researcher profiles, including e.g. affiliations, education, description of research, and other research merits. Again, the first task was to build a common data model. But the researcher's data also had to take into account the fact that their uses varied widely between organisations. It was also literally the researcher's own data, so the intention was to include it in the research information hub following a MyData [4] approach, where the researchers had the right to decide what information about them appears in Research.fi. Other data in the research information hub were considered to be general research metadata that could be publicly displayed in various research information systems in accordance with established scientific practice. For the researcher's own data, instead, a section was built in the research information hub where the researcher can control the display of his/her own data.

The public view of the research information hub, the Research.fi portal, was launched in June 2020. A description of the Finnish science system and key statistics were also implemented on the portal. The search functionality of the portal was developed to make use of topic modeling, which will also allow the portal to be searched by different topics.

Currently, all features and content planned for the research information hub in 2017 are technically ready. The data is also be used by various services. However, there are several areas for further development, such as the construction of a CV service, the transfer of data to ORCID, the definition of organisations own research projects and their integration into the information hub. Also, there is still room for improvement in the comprehensiveness of information, especially concerning research data, infrastructures and researchers, as not all HEIs or research institutions transfer these information yet.

4 Lessons learned

From the perspective of the Ministry of Education and Culture, the research information hub has been a successful project. All four objectives mentioned at the beginning have been advanced by the service.

" Provide landscape of Finnish research ecosystem and basis for science policy". The service has brought together data that were in different places. The pooling of data has enabled new methods of analysis to be applied to the data. In particular, this has been useful for research funding decisions which until now have

only been available in the databases of individual funders. In addition, the research.fi portal, built on top of the research information hub, is a cost-effective platform for describing the Finnish research system.

"Ability to search for Finnish research and researchers". The service has improved the visibility of Finnish research. Improving visibility is particularly important as the Finnish government has pledged to significantly increase public funding for R&D in the 2020s in order for Finland to reach the target of 4% of GDP spent on R&D. A significant use of public funds requires that research is also made visible. A research information hub will also facilitate a more diversified use of expertise by the media.

"Support for the reuse of research information in other services and systems". Data from the research information hub are already used in a number of services, such as the Academy of Finland's funding applications and the transfer of data to OpenAire. As expected, the use of the information hub in different services has been the slowest to progress, as this area also requires the capabilities of the third party services to using the data. This is by far the area where the greatest benefits are yet to be realised.

"Merit researchers and make their research output and activities more visible". Supporting diversity in researchers' career development has become an increasingly important theme in the 2010s and, at the latest with COARA [5], a key science policy objective. The research information hub is a good technical platform to highlight the diversity of researchers' careers. Without the service, the promotion of diversity in merits would risk becoming a desk-top activity.

In addition to these stated objectives, the service has produced many spin-off benefits. For its part, it has created a culture of collaboration and networks that facilitate cooperation. It has created a forum for discussion of research metadata, the end result of which is not just a theoretical exercise in classifying, for example, researchers' activity data, but is actually intended to produce harmonised data. The quality of all research metadata has improved. The quality of the data used by the Ministry in its steering has improved, as has the quality of the data used in research evaluation.

Building a service such as the research information hub at national level makes sense in many respects. The abovementioned objectives are certainly shareable from the point of view of an individual research organisation, but it would not be possible to implement them in the same way at the level of an individual organisation. The result would be piecemeal solutions with less impact. There would also be a risk of creating solutions that are not interoperable. A national solution is also cost-effective compared to solutions at the level of individual organisations.

On the other hand, the research information hub is built to operate on top of individual organisations' own research information systems. The key principle is that the researcher does not enter any data directly into the information hub, but all data comes from the source systems. This gives organisations the flexibility to use the CRIS systems they need. The research information hub also aims to contribute to the creation and support of a research infrastructure independent of commercial operators. It is important for national science policy to make use of the widest possible knowledge base. It is not necessarily feasible for commercial operators to produce such a broad knowledge base. The research information hub has been implemented in-house, which has offered a flexible way to create a solution that meets many different needs. At the moment, there are no ready-made products available that would directly meet the needs at the national level, which also involves the facilitation of broad cooperation between information users, data sources and other stakeholders.

References

- [1] OECD (2023), "The future of Finland's funding model for higher education institutions", OECD Education Policy Perspectives, No. 76, OECD Publishing, Paris, <https://doi.org/10.1787/3d256b59-en>.
- [2] Nielsen et al. (2023) "Evaluation of the governance and funding practices used by the Ministry of Education and Culture for steering Finnish Higher Education Institutions. Final report. Technopolis Group. https://api.hankeikkuna.fi/asiakirjat/ce143d5f-5338-495b-a311-9b052f49037f/101ef200-d993-48f6-9675-e4f2b3630063/MUISTIO_20230616061550.pdf accessed 13 Feb 2024.
- [3] Auranen, Otto & Pölönen, Janne (2012). Tieteellisten julkaisukanavien tasoluokitus: Julkaisufoorumi-hankkeen (2010–2012) loppuraportti. Tieteellisten seuran valtuuskunnan verkkojulkaisuja 2012: 1. <https://julkaisufoorumi.fi/sites/default/files/2020-08/Julkaisufoorumi-hankkeen%20%282010-2012%29%20loppuraportti.pdf> accessed 13 Feb 2024.
- [4] Poikola, Kuikkaniemi, Honko (2016). MyData - A Nordic Model for human-centered personal data management and processing. <https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/78439/MyData-nordic-model.pdf> accessed 13 Feb 2024.
- [5] CoARA—Coalition for Advancing Research Assessment (2022) Agreement on Reforming Research Assessment. <https://coara.eu/app/uploads/2022/09/2022_07_19_rra_agreement_final.pdf> accessed 13 Feb 2024.