Juuli – a national portal for publication data

Jyrki Ilva / The National Library of Finland

Submitting publication information to each university's own current research information system has long been part of the annual routine for Finnish university researchers. Now, the information collected in these local systems can also be accessed in a shared interface that comprehends all Finnish universities. As part of the larger RAKETTI project co-ordinated by CSC, the National Library of Finland has built the Juuli publication portal (www.juuli.fi), which has been fully accessible since the beginning of June.

Improving the transparency of data collection

The possibilities for establishing a unified national current research system were investigated in the RAKETTI-JURE project which ran from 2009-2011. However, instead of a unified national system, a decentralised model was ultimately adopted. Universities have continued developing their own, local systems, with many of them currently implementing or acquiring new research information systems.

The Ministry of Education and Culture (OKM) has collected information on the volume of university publications for a long time, but the actual reference data stored in local systems has only been collected for the past two years, beginning with publications from 2011. In the new university funding model adopted in 2013 as much as 13% of total university funding (more than 200 million euros a year) is distributed based on the volume of scientific publications. This has substantially increased the importance of publication data collection and data accuracy.

Indeed, a key goal in establishing Juuli was to improve the transparency of data collection. Compiling data into a universally-accessible, shared interface helps in identifying problem areas in data collection and standardising common practices.

Juuli provides access to reference data for both the domestic and foreign publications of the Finnish universities in a single user interface. Large-scale international publication databases, such as Web of Science and Scopus, already contain reference data for most publications appearing in international scientific journals. However, in fields where publishing in, for example, national publication channels or conference publications is common, they are generally far less comprehensive. Also, they differ from Juuli in that they are fee-based, i.e. the data they contain can only be accessed by client organisations.

Of course, not even Juuli is completely comprehensive. Scientific publishing mainly occurs outside universities, and it is quite common that the home university may not be able to collect information on all of the publications produced by its researchers. Particularly problematic are cases where a researcher transfers to another organisation during the publishing process and is no longer present to report on his or her publication when it is finally released. It is also useful to note that, at least at this point, Juuli only contains the publication types A-E of the OKM data collection, which means among other things that dissertations are not included.

Uniform data

Juuli is an interface for presenting publication data, not an actual current research information system. It is based on the open source VuFind software, which uses Apache Solr search technology. The same software is being used by the National Library of Finland in the national discovery interface Finna (www.finna.fi), which is currently in the beta testing phase. Compared to Finna, the user interface of Juuli is much more minimalist and simple in appearance and function.

On the other hand, the data contained in Juuli is much more uniform than that found in Finna, as all of the data imported to the portal is already in the same format specified in the OKM guidelines for data collection. Indeed, it has been possible to customise the user interface functions a great deal, according to the specific attributes of the material it contains. In Juuli, search results can be quickly refined by, for example, university, scientific field, publication language, Publication Forum classification or publication type.

Each publication appears in Juuli only once, i.e. the reference data for a joint university publication collected from several universities is combined to form a single record. Conversely, the authors of the publications are not identified – authors with the same name are not differentiated from one another in the user interface, and the same author name may appear in different forms. The reliable identification of authors would require a national researcher register or the national implementation of international researcher-specific identification codes (e.g. ORCID or ISNI). A preliminary investigation into this is currently underway.

In addition to Juuli, information on the volume of publications produced by the Finnish universities is also available on the Vipunen reporting portal (www.vipunen.fi), which is maintained by CSC. Vipunen contains a large amount of other statistical data related to higher education institutions, as well, and it also allows for data analysis. Juuli, however, offers the possibility of going beyond numerical data to publication-level reference information and, in many cases, directly to the publication itself.

Demand for the portal

Juuli's target public and the assumed user needs associated with it have sparked discussion since its inception. Indeed, in February-March of 2013, a limited test run and online survey was conducted at the Universities of Tampere and Turku to gather university staff opinions on the portal. The survey respondents were chosen from various personnel groups, with a total of 40 responses being received.

The overall impression given by the responses received was very positive. Respondents felt that the open availability of publication information was an extremely important goal and there would be major support for the existence of a national service like Juuli. This was also made apparent in many of the respondents' comments, which emphasised the transparency of research activities and importance of comparisons between universities:

"The portal reveals what universities and their units offer in different research fields, because it's possible to make even very detailed searches, such as who and what organisations are publishing on highly-esteemed publishing channels within a certain field. It also provides excellent exposure for "rank-and-file" researchers, who publish a lot and vice versa; the system is merciless – for the researcher, it's 'publish or perish'."

In free-form responses, the survey respondents expressed not only praise, but also a certain amount of

criticism toward a perceived increase in the amount of measurement and assessment of research activities. Even though Juuli is primarily intended for the examination of publishing at the organisational and disciplinary level, many respondents noted that the user interface made it easy to find the publication information for individual authors, including the publication types and Publication Forum classifications. On the other hand, presenting publication information in a uniform format and in accordance with the quality classification was considered an incentive for researchers to give more thought to publishing and the choice of publishing channels.

The portal prototype was in beta testing phase during the survey. With this in mind, the feedback received from beta users on the portal itself was surprisingly positive. According to the feedback given, many users felt that it took a while to figure out the operating principle of the user interface, but, after doing so, they were primarily satisfied with it:

"I visited the portal a few times before I realised that I could find a lot more information in there than I initially thought. The search menu suggests that the search selection is limited, but after getting inside the information for individual publications, I found that it's possible to make a wider variety of searches," explained one respondent.

Plans for the future

At present, Juuli only contains the publication data of Finnish universities from 2011 (approx. 32,000 entries). There is no comprehensive, uniform data available for publications predating this, i.e. 2011 will probably remain the first year of data found in the portal, even in the future.

For the time being, new data will be added to the portal in a batch process each year. The 2012 publication data collected from universities in the spring of 2013 will be added early this fall, following inspection rounds conducted by the Ministry of Education and Culture, CSC, Federation of Finnish Learned Societies and universities. Responses in the university survey expressed a desire for real-time updating of the data, but this would require both technical upgrades for the local research information systems and a major development of data collection processes. In other words, it will take some time before this can be made a reaility.

Respondents wanted to see publication data for not only universities, but also for universities of applied sciences, state research institutes and hospital districts included in Juuli, so that it would provide a comprehensive picture of all Finnish scientific publications. The publication data collected from universities of applied sciences will be included later this year, beginning with information from 2012. On the other hand, the research institutes are, to a certain extent, a more challenging case, as they operate under several different ministries and there is a need to find common ground regarding data collection principles and the parties responsible for it.

Juuli now offers a freely available user interface for the publication data of the Finnish universities. However, in connection with the project, there has been some preliminary discussion on whether it would be possible to make the publication information more widely available also as open data. This would require a more detailed discussion on what terms and in what form the data should be made available.