euroCRIS Membership Meeting Spring 2019

May 27-29, 2019 at Helsinki, Finland

Program

Monday 27th /

09:00 - 13:00 / Board meeting

13:00 - 14:00 / Board Lunch @ Bröd Punavuori (Bulevardi 26, 00120 Helsinki)

14:00 - 17:00 / Developers' workshop: Implementation of the OpenAIRE CERIF-XML Guidelines
Jan Dvorak, euroCRIS and Joonas Nikkanen, CSC - IT Centre for Science

17:00 - 17:30 / Coffee break

17:30 - 18:30 / CERIF tutorial
Valerie Brasse, euroCRIS

17:30 - 18:30 / Interest group meeting on the updated DRIS
Pablo de Castro, euroCRIS
Following the first round of discussion on the DRIS held with the RIM community at the Warsaw membership meeting in Nov'2018, the euroCRIS Directory of Research Information Systems (DRIS) is getting updated with hundreds of new entries. The Interest Group meeting in Helsinki aims to discuss the progress achieved and to make the snapshot for Nordic countries more comprehensive.

18:30 - 19:00 / Networking / Registration

19:00 - 20:30 / Reception

Tuesday 28th /

08:00 - 09:00 / Registration

09:00 - 09:15 / Words of welcome
Ed Simons, euroCRIS

09:15 - 10:15 / Opening session
[Chair: Hanna-Mari Puuska]

Higher Education and Research in Finland and the Vision of Future / Erja Heikkinen, Ministry of Education and Culture, Finland

European Open Science Cloud - What's Up! / Per Öster, CSC - IT Centre for Science, Finland
The European Open Science Cloud (EOSC) was launched by the EU in a festive event in Vienna November 2018. What was actually launched and what is the opportunity for the higher education institutes? How to engage and where is actually the cloud? The questions are many and Per Öster from CSC - IT Center for Science will try to shed some light by giving insight how CSC, Finland and the Nordic-Baltic countries are commonly engaging with the EOSC.

10:15 - 10:45 / Coffee break

10:45 - 12:30 / Nordic session (Open Science)
[Chair: Jan Dvořák]

Research data - forever FAIR: Finnish perspective / Pirjo-Leena Forsström, CSC - IT Centre for Science, Finland

FAIR already? Principles of reusability and research output – evaluation at a national level / Camilla Lindelöw, National Library of Sweden
The Swedish government gave the Swedish Research Council and the National Library of Sweden parallel assignments to propose criteria and a method for assessing how well research data and scholarly publications produced at Swedish organizations fulfil the FAIR principles. This is in line with the on-going work in the European Union regarding implementation of FAIR. This talk will focus on evaluation at a national level, our recommendations and the possibilities and caveats we have identified during our work. We propose a method based on collecting and evaluating metadata. This will affect the infrastructure handling metadata for research data and scholarly publications.

Redefining Open - National Open Access Monitoring in Finland / Jyrki Ilva, National Library of Finland
The Finnish Open Access monitoring efforts are connected to the national data collection conducted by CSC for the Ministry of Education and Culture. Since 2016, the data collected from the universities and universities of applied sciences has been relatively consistent and reliable, and it shows that the share of openly available publications is increasing. Although the current data is already good enough to be used in the university funding model, there are plans to further enhance the data by moving into an attribute-based model, which would make it easier to make comparisons with other data sources using different definitions of Open Access.

Journal.fi - building infrastructure for open access journals in Finland / Antti-Jussi Nygård, the Federation of Finnish Learned Societies
Journal.fi is a national platform for Finnish online journals. It was founded in 2017 and it is based on Open Journal Systems software. The main aim of the project has been to reduce the cost of editorial work and publishing and to ensure interoperability with national and international standards and systems.

CRIS in OpenAIRE - we take you onboard / Aenne Loehden, Jochen Schirrwagen, and Andreas Czerniak, University Bielefeld, Germany
Rich research information in CRIS is a treasure for funders, organizations, and also for researchers. OpenAIRE works together with the research community, research administrators, funders, and with euroCRIS to lift the treasure and make this information visible in the OpenAIRE scholarly communication graph which is a key component of the European Open Science Cloud. The latest release of the OpenAIRE Guidelines for CRIS Managers is a milestone to achieve interoperability between CRISs and OpenAIRE. Early adopters in a number of platforms demonstrate the enormous potential the integration of research information in OpenAIRE can have. We can identify different roles and functions of CRIS with regard to extend and enrich information in OpenAIRE, with e.g. publications, patents, and products sets contributing to research results metadata, and funding and projects set augmenting authoritative data.

We will present progress and challenges of CRIS integration in OpenAIRE in terms of CERIF mapping, validation, aggregation, indexing, and finally discovering research information in the OpenAIRE portal and provide suggestions for improvements. Integration of CRIS in OpenAIRE is of mutual benefit. For instance, OpenAIRE's brokering service allows to enrich and correct metadata located in its content providers. The Usage metrics service is an opportunity to provide standardized information on usage activity for publications and data.

Finally, we'll give an overview of OpenAIRE's strategy to sustain the infrastructure and its human network.

12:30 - 13:30 / Lunch

13:30 - 14:35 / Nordic session (Research Information)
[Chair: Sadia Vancauwenbergh]

The Research Information Hub as an access point to Finnish research / Hanna-Mari Puuska, CSC - IT Centre for Science, Finland

The Research Information Hub gathers and shares data about research conducted in Finland. The information about researchers, publication, research data, ongoing research funding and research infrastructures will be effortlessly found from a single source. The rationalization of information flows lightens the administrative workload and makes the information more accessible and easier to find. The research outcomes will be more accessible to citizens using it, since the hub will provide a single, uniform, open and comprehensive access point available for everyone. The Hub is developed gradually and should be ready by 2020.

Cristin in the Norwegian research landscape / Espen Sandøe Karlsen, UNIT - Direktoratet for IKT og fellestjenester i høyere utdanning og forskning, Norway

Cristin as the national CRIS in Norway is an important hub for metadata in the research ecosystem. We will give an overview of work in progress and show some examples of how Cristin gathers data "in real-time" about Norwegian research publications and research projects and serves as a hub to make metadata about Norwegian research easily available for reuse and further studies. Using reported publication data as a source for OA-monitoring. Working on improved data quality and ease of administration by implementing the new API from NSD to improve data flow between Cristin and The Norwegian Register for Scientific Journals, Series and Publishers.

At CRIS2016 we presented the new integration with The Regional Committees for Medical and Health Research Ethics (SPREK). We will present a work in progress based on this integration to establish a new indicator for hospitals on a subset of health projects (# of patients included in clinical trials).

In other news: We are also on the verge of getting new types of institutions in Cristin. In addition to the hospitals, research institutes and HE institutions, Cristin will include archives, museums and libraries.

The Nordic List – Nordic collaboration for developing a common Nordic publication channel list / Kristin Gåsemyr, NSD - Norwegian Center for Research Data

Since 2015 the Nordic countries have been collaborating to develop a common registry of authorized research publication channels with bibliographic data on journals, series and publishers. Denmark, Finland and Norway have joined their national lists of authorized research publication channels, used for indicators in the national performance-based research funding systems. Sweden, Iceland, Greenland and Faroe Islands do not have a similar funding systems but wanted to use the Nordic list as a tool for a national or Nordic overview of authorized research publication channels. NSD - Norwegian Centre for Research Data, coordinated the technical development of the list and the work was funded by NordForsk and The Presidency of the Nordic Council. The plan for technical development was agreed upon by the collaboration and funding lasted from 2017 until summer 2018. The Nordic list consists of a bibliographic dataset on the publication channels listed in each countries national database. The national lists are joined technically in this system. The aim was to give all the countries a common source of information, bring resource efficiency for the establishes national lists, and a source of information and a tool for the Nordic countries without a list.

14:35 - 14:45 / Break

14:45 - 15:30 / ENRESSH session (research information) organized by the European Network for Research Evaluation in Social Sciences and humanities
[Chair: Tina Lingjaerde]

Towards an integrated research information infrastructure at the Nordic level / Gunnar Sivertsen, NIFU, Norway

National bibliographic databases for research output: opportunities and challenges / Linda Sïle, University of Antwerp, Belgium
In more and more countries one can find a national bibliographic database for research output as a recent survey of European national bibliographic for research output in the social sciences and humanities shows. Typically these databases can (and are) used to report on, monitor, and evaluate research activities within one national context. At the same time, with the ubiquity of tools and methods to generate data-driven insights, these databases appear as valuable sources for the furthering of our understanding of research activities more generally. This means going beyond the data use only in national settings. A closer inquiry into the database work underpinning the various European databases, however, points to challenges for the use of data from multiple national databases. For example, one can see differences in criteria that are used to include new database records, the use of different output type and disciplinary classifications, and also different approaches to the record quality control. These challenges, their meaning as well as implications for the use of data from national databases in cross-country settings is the focus of this presentation.

15:30 - 15:50 / Coffee break

15:50 - 16:50 / Panel discussion: The challenges and requirements for international CRIS systems
[Chair: Ed Simons]

Gunnar Sivertsen, NIFU, Norway, **Sadia Vancauwenbergh**, ECOOM-Hasselt, Hasselt University, Belgium, **Linda Sïle**, University of Antwerp, **Andreas Czerniak**, University Bielefeld / OpenAIRE, **Hanna-Mari Puuska**, CSC – IT center for science
With the consolidation of the national-level RIM infrastructure in several European countries, the discussion is gradually moving onto the possibility of setting up international CRIS systems. This panel aims to discuss the current barriers that need to be overcome and the opportunities that such an international development might offer.

16:50 - 17:05 / Update from the euroCRIS Board

17:05 - 18:15 / euroCRIS Business session
[Chair: Valerie Brasse]

Clarivate Analytics, Elsevier, Sigma, 4Science

19:30 / Dinner @ Restaurant Kappeli (Eteläesplanadi 1, 00130 Helsinki)

Wednesday 29th /

09:00 - 10:45 / Lightning talks
[Chair: Jürgen Güdler]

Mapping the NARCIS aggregated domain to OpenAIRE CERIF guidelines / Wilko Steinhoff, DANS, the Netherlands

NARCIS Portal (www.narcis.nl) domain main-entities consist of the following five main "collections": - Organizations, Researchers, Research projects, Publications and Datasets. The Publications and Datasets are being aggregated from different institutional PMH repositories. While the others originate from a relational database, called Nederlandse Onderzoeks Databank (NOD) that are actively maintained by DANS itself. All the Publications metadata comply to a national standard preserved by 'EduStandaard'1. This makes it easier to understand and interpret the collected metadata for an aggregator. Theoretically, all harvested 'collections' are represented as core CERIF entities and should easily be mapped to OpenAIRE CERIF entities. However, to be able to meet with the OpenAIRE Advance Guidelines for CRIS managers we had to re-classify, map and transform the existing model.

In this lightning talk I will focus on some of the issues we ran into while converting these entities:

- Mapping 'EduStandaard' and NARCIS domain specific vocabularies to OpenAIRE (COAR). Different hierarchy of datatypes known only as 'subtypes' in the NARCIS model/domain.
- Name Identifiers: However deprecated (no resolver), the Dutch DAI is still in use to identify researchers and link researchers to their works and organizations.
- NARCIS as an aggregator: How do we deal with (local) identifier-types aggregated from different sources? We need to map federated Identifiers (Persons, Journals, etc.) to local ones to meet the PMH 'referential integrity' mandated by OpenAIRE guidelines.
- Linking research output to research projects.

An "Open" discussion: Preliminary findings from an OCLC survey of open content costs and benefits / Rebecca Bryant, OCLC

Despite growing support for Open Science, Open Access, and Open Government Information, measures of impact, investment, and cost for libraries and research organizations remain largely unknown. How much are research institutions actually investing in infrastructures and services to support open content for academic and non-academic users? Without greater knowledge of the costs of human and financial resources that libraries are allocating to support open content, how can they make informed planning decisions? And, if the benefits for the library and its users are unclear, how can libraries measure how successful they are in realising those benefits?

To help inform meaningful discussions around these challenging questions, OCLC conducted a survey that focused on libraries' ambitions, realities, and investments in support of open content. This assessment collected data about library open content activities across multiple internal silos, including: e-resources management, institutional repositories, research information management/CRIS systems, digitised heritage collections, data archives, born-digital (legal) deposit, and more. This survey represents a significant early effort to qualify and quantify library efforts in the field of open content at an international scale. By late January 2019, 705 responses from 82 countries were submitted, offering a rich pool of information. Seventy-two percent of the responses comes from research and university institutions and from this subset we can distill the relative importance of CRIS-IR in libraries' open content activities.

In this presentation, representatives from OCLC Research will share preliminary findings from this global survey and discuss implications for the euroCRIS community and the broader open content ecosystem.

Deploying CRIS systems in report on (public) research funding / Sadia Vancauwenbergh, ECOOM-Hasselt, Hasselt University, Belgium

In 2011, Peters et al. performed a study on the administrative burden that Flemish research administrations encounter when fulfilling reporting obligations imposed by governmental agencies due to the inefficient collection of research-related information. In line with this report, the Flemish government has issued new regulations anno 2019 that prescribe the use of the Flemish Research Information Portal (FRIS) as the main source of information used for reporting on public research funding. Obviously, these regulations push the requirements of the FRIS portal to an even higher level in terms of the variety of gathered information in relation to the reporting obligations. As such, the requirements include the expansion of the FRIS metamodel with new information objects, the expansion of the metadata for existing information objects and the inclusion of new classification schemes, following the principles of data governance. Furthermore, the direct consumption of FRIS-information for reporting purposes, also imposes even higher quality requirements of the data, not only in terms of the associated business and validation rules, but also in terms of the completeness of the data, the validation thereof and even on the processes associated with data registration. This presentation sheds light on the recent developments in Flanders in order to use FRIS as a reporting system for public research funding.

CRIS enabling open access services – case UTUCRIS at the University of Turku / Anna Salmi and Anna-Kaarina Linna, University of Turku, Finland

The University of Turku has developed a service complex that integrates all-year publication data collection with a green open access service. According to the publication policy of the University, publications produced at the University must be openly available whenever the publisher's terms allow. The Current Research Information System UTUCRIS, based on the Converis system provided by Clarivate Ltd., serves as the repository.

The Library imports approximately 50 % of all publications for the researchers from source databases. The UTUCRIS-system sends the author an automatic message about each validated publication, and the author can submit the accepted manuscript to the Library in answer. The author can also report their publication with a simplified e-form containing a field for the manuscript.

The Library self-archives the manuscript in the system alongside the publication information. As the publisher's embargo period lapses the manuscripts are permanently available online via the UTUCRIS portal.

By 2018 nearly 50 % of all the publications saved in the system were available via green or golden Open Access. The findability of the self-archived manuscripts is currently being developed further; an interface to a DSpace-run publication repository and PubMed via LinkOut service are under construction.

BPM and CRIS working together for the ultimate data reliability: Omega-PSIR at Medical University of Łódź / Wiloletta Michalak, UMED Łodz and Jakub Koperwas, Warsaw University of Technology, Poland

It is crucial for the CRIS/IR system to operate on complete and reliable data. For this sake, it is crucial to establish proper policies, roles, and responsibilities in the organisation. Still, some problems frequently occur like duplicate data, improper author matching, problems in reporting where author has more than one affiliation, not sufficient open access compliance etc. During the deployment at the Medical University of Łodz we are evaluating the approach where the Business Process Management System orchestrates the processes crucial for the reliability of CRIS data delegating tasks to people and other systems at the University. We would like to share our early experiences.

VIVO, REF and KDSF – Local initiatives in a configurable CRIS / Miguel Garcia, Clarivate Analytics

Large communities of research-intensive institutions tend to establish standards in order to communicate among them or with third-parties. EuroCRIS is an example of that, but countries tend to also use standards to find ways of retrieving research data from their numerous institutions. In this session we will explore how Converis, a configurable CRIS system, is supporting some initiatives namely the KDSF (Germany) and the REF

(U.K.) in order to help institutions to comply with them with minimal effort. In addition, we will explore how the flexibility of Converis' data model compares and aligns with the VIVO ontology and how institutions using Converis are currently integrating with this open source research portal.

10:45 - 11:15 / Coffee Break

11:15 - 13:15 / Jostein Hauge session
[Chair: Mark Cox]

Updated/extended data model for the German KDSF (Research Core Data Set) / Sebastian Herwig and Sophie Biesenbender, Univeristy of Münster, Germany

The German science and higher education system is characterized by federalism, multi-level governance and interwoven regulatory competences of different levels of government. Establishing binding standards and harmonized policies for German higher education institutions and non-university research institutions is a complex task that requires concerted action and co-operation between the federal and state governments. This talk is intended to provide insights from the specification of the German KDSF – the core data set research. Particular emphasis is on the support structures currently in place to facilitate the dissemination and implementation of the KDSF in the federal and multi-level governed German science system as well as the plans for future development perspectives of the standard.

How opening up data on publicly funded research became a driver of innovation in Flanders. The use case of FRIS. / Leen Van Campe, Vlaanderen, Belgium

33.000 research projects. 80.000 researchers. 378.000 scientific publications. This is the amount of data from scientific research that are being collected by Flanders Research Information Space (FRIS), a program run by the Flemish Government (Department Economy, Science of Innovation). Information comes directly from 34 research institutes whose processes and systems are aligned to support the government's strategy that opening up data is essential for driving knowledge transfer to the industry and the broad public. Data is being fed continuously and in an automated way into the Flemish information space.

Our aim? To find expertise in Flanders, lead to new insights and create an innovation-driven economy. That's why we promote openness of research and redesign some of the existing information flows in Flanders, reduce the administrative burden of reporting by requesting information only once, make innovation strategies more efficient and effective, and make all our research data accessible and publicly available for reuse via open API'S. This is FAIR data at work! #winner of an e-gov Awards "Open Data" and "Best Project 2018".

SIGMA CRIS, an implementation of VIVO / Anna Guillaumet, Sigma

The growing complexity of the digital research environments has highlighted the importance of having powerful tools for searching and discovering research information.

The advance of Open Science, that includes not only Open Access, but also Open Data and the need to make the data FAIR, has shown that it's necessary to evolve the way we search and visualize the research results. So, when the scientific information is correctly stored in the CRIS, it must also be easy to search and find, and must guide the discovery of the information.

This is one of the objectives within SIGMA's Strategic Plan for the Research area. Since the beginning of 2018, SIGMA is working to improve the search tools for scientific production, whose first results we will present here.

This project began by evaluating the best ontology to translate the current data model for SIGMA CRIS, to a semantic model. To do this, we studied some of the semantic engines that exist and finally decided to test the VIVO solution.

We found that VIVO ontology fits more than 80% with the Spanish model for Research and took into account that VIVO is an OpenSource software supported by members and an ontology to represent the scholarship that is used by relevant universities, highly positioned in the international rankings and, finally, we also value the large community of members behind VIVO and Duraspace.

For these reasons, SIGMA decided to join the DuraSpace community, participating in the governance of the VIVO project.

With this experience, we have now an adapted SIGMA ontology based on the VIVO ontology and a Semantic Expert Guide, working on SIGMA CRIS. With our involvement in the governance of the VIVO project, we have participated in the definition of the VIVO product direction for 2019, among other tasks.

Developing Common Specification on Research Information and Data (CS) based on the CERIF model /
Renata Arovelius, Swedish University of Agricultural Sciences
Research information and data are part of the business information of Swedish universites and other research institutions. Research records are consequently subject to implemention of the Principle on Public Access to Records (OP). It means, universities and other research institutions, being part of the government, must ensure accuracy, reusability and longevity for the long term access to this information. Common Specifications for e-archiving and e-records management (CS) can be a good facilitator for the access to research information and data in ongoing activities and over time. CS outline requirements for metadata and appropriate keys for interoperability between different systems.

There is a significant lack of semantic and technical interoperability between systems within research administration, e.g., personnel information systems, accounting systems, records management and publication systems. Implementing the CERIF model together with the CS for research records and data gives the possibility for mapping crucial identifiers and a standardised description including an XML schema for transfer between different systems supporting management of research records as well as systems for long term preservation and access.

Four Swedish universities supported by other universities in Sweden and the Swedish Research Council recently undertook an initiative on development of CS for research records and data based on the CERIF model. This presentation will discuss some of the results on possibilities and limits within the implementation of the model together with the CS.

Developing an Open Repository into Full Service platform for Open Publishing - The Case of 25 Universities of Applied Sciences in Finland / Minna Marjamaa, Laurea University of Applied Sciences & Tiina Tolonen, Oulu University of Applied Sciences, Finland
The integration of a simple CRIS and a long-term preservation into open repository saves resources for 25 universities of applied sciences. The model will reorganize publication, publicity issues, archiving and reporting in UASs of Finland as well as it effectively promotes open publishing in Finnish UASs.

13:15 - 14:15 / Lunch

14:14 - 15:15 / Presentation session (Research funding)
[Chair: Jonas Gurell]

Using SweCRIS-data for a more strategic investment approach in a fragmented R&D-funding landscape /
David Jonason, VINNOVA - Swedish Innovation Agency
Given Sweden's "rather fragmented landscape of over 20 mostly mid-sized agencies funding innovation-related activities [...], horizontal coordination is a major concern" (OECD, 2016). As one of those agencies, Vinnova

(Sweden's Innovation Agency) is experimenting with building internal tools to make public funding data (made available by SweCRIS and others) more accessible and actionable for our program managers, thus providing the basis for more informed, data-driven and portfolio-based investment approach. We will discuss our take on horizontal coordination, the potential for more data-driven public R&D-funding in Sweden and how this fits into Vinnova's shifting bottom-up approach.

Project funding data of the Academy of Finland as an element of science policy knowledge base / Otto Auranen, Academy of Finland

Mission of the Academy of Finland (the Finnish national research funding agency) is to fund high-quality scientific research, provide expertise in science and science policy, and strengthen the position of science and research. The funding is allocated on a competitive basis based on international peer-review. Funding applications submitted to the Academy and project reports after the funding period provide a rich data source for analyzing various aspects of research activity, its trends and impact. Statistical, bibliometric and text-mining analyses of this data can provide vital information on Finnish research to decision-makers and scientific community. Presentation will give an overview of current and future efforts of the Academy in analyzing application and funding data.

The use of researchfish® by the Novo Nordisk Foundation / Henrik Barslund Fosse, Novo Nordisk Foundation

The Novo Nordisk Foundation uses researchfish® as its reporting platform for recipients of research grants in order to collect information about research activities, outputs, outcomes and impact. These data constitute the base layer for many of the analyses that the Foundation performs to demonstrate the impact of the Foundation's grants. The presentation will describe the Foundation's funding policies, the workflows for information collection on research outputs through researchfish® and supplementary data, and how the Foundation carries out its analyses. Finally, some results from impact analyses of Novo Nordisk Foundation-funded research will be shown.

15:15 – 15:45 / Coffee break

15:45 - 16:30 / Panel discussion: the role of research funders
[Chair: Pablo de Castro]

Jonas Gurell, Vetenskapsrådet, Sweden, **David Jonason**, Vinnova, Sweden, **Otto Auranen**, Academy of Finland, **João Moreira**, Foundation for Science and Technology, Portugal
The gradual emergence of Open Research platforms and the recent issuing of Plan S show areas where research funders exert a key influence in shaping fairer scholarly communications. The implementation of solid RIM workflows should also very much help research funders achieve their objectives in aspects like reporting and assessing research impact. In order to discuss these issues, this panel discussion will bring together a number of representatives from funders in various European countries.

16:30 - 17:45 / Closing