Towards a Research Information Hub in Finland

Hanna-Mari Puuska, CSC – IT Center for Science, Finland
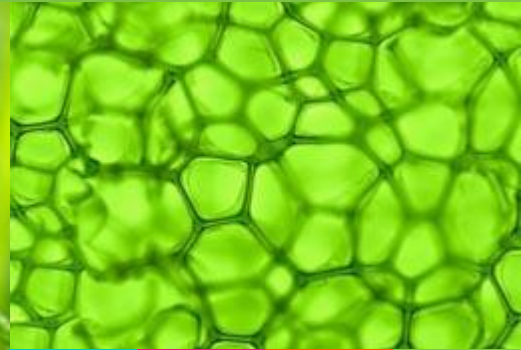
CSC – Finnish research, education and public administration ICT knowledge centre

Contents of the presentation

1. Background: Publication data collection in Finland

2. VIRTA Publication Information Service

3. Towards a Finnish Research Information Hub

4. Extension of VIRTA concept to European level: the ENRESSH-VIRTA pilot

Background: Publication Data Collection in Finland

Background: Publication data collection in Finland

- Since 2011, Ministry of Education and Culture has collected bibliographic information on publications annually from
  - 14 universities and 5 university hospital districts
  - 23 universities of applied sciences
  - 12 state research institutes

- In total, ~50 000 publications per year
  = books, journal articles, conference papers, non-scholarly publications

- The data are publicly available
  - Statistical data: www.vipunen.fi
  - Bibliographic data: www.juuli.fi

JUULI publication portal (www.juuli.fi)

JUULI Julkaisutietoportaali

Suomi Svenska **English**

| | All Fields ⌄ | 🔍 Find | **Advanced** ☑ Retain current filters |

Search:

Showing **1 - 20** of **7,245** for search: '', query time: 0.16s

Sort Title ⌄

Narrow Search

Remove Filters

Field of Science: Natural sciences ✕

Publication Forum Class: 3 ✕

Institution ▲

1 — NO IMAGE AVAILABLE — **A 0.5-Mbp deletion on bovine chromosome 23 is a strong candidate for stillbirth in Nordic Red cattle**
by: Sahana, Goutam; Iso-Touru, Terhi; Wu, Xiaoping; Sander Nielsen, Ulrik; de Koning, Dirk-Jan; Sandø Lund, Mogens; Vilkki, Johanna; Guldbrandtsen, Bernt
Genetics Selection Evolution, BioMed Central, 2016
A1 Journal article (refereed), original research
🔗 Get full text
🔗 Get full text

University — 7,123

Research Institure — 223

Health Care District — 152

2 — NO IMAGE AVAILABLE — **1,25-Dihydroxyvitamin D(3) Influences Cellular Homocystine Levels in Murine Preosteoplastic MC3T3-E1 Cells by Directs Regulation of Cystathionine beta-Synthase**
by: Kriebitzsch C; Verlinden L; Eelen G; van Schoor NM; Swart K; Lips P; Meyer MB; Pike JW; Boonen S; Carlberg C; Vitvitsky V; Bouillon R; Banerjee R; Verstuyf A
JOURNAL OF BONE AND MINERAL RESEARCH, 2011
A1 Journal article (refereed), original research
🔗 Get full text

University of Applied Scien... — 40

Author ▲

Juvela, M. — 126

3 — **1,3-Dipolar cycloaddition of nitrones to a nitrile**

Publications in the performace-based funding model of universities

13 % of universities' core funding (~200 mill. eur) is allocated via publication points:

Publication type	Level 3	Level 2	Level 1	Level 0
Peer-reviewed monograph (C1)	16	12	4	0.4
Peer-reviewed article in journal (A1-2)	4	3	1	0.1
Peer-reviewed article in book (A3)	4	3	1	0.1
Peer-reviewed article in proceedings (A4)	4	3	1	0.1
Peer-reviewed edited work (C2)	4	3	1	0.1
Not-peer-reviewed monographs	0.4			
Not-peer-reviewed articles	0.1			

The level is based on journals, series and publishers evaluated by national scholarly panels in **Publication Forum**
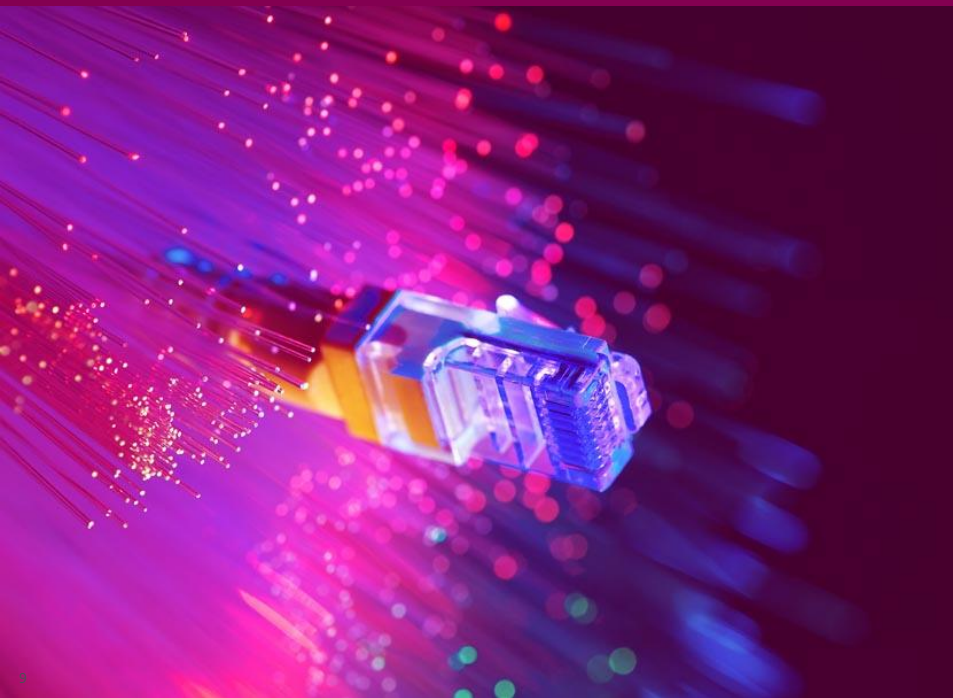
- Level 1 = BASIC scientific publication channels
- Level 2 = LEADING scientific publication channels
- Level 3 = TOP journals and series
- Level 0 = Other identified publication channels

Finnish Publication Forum
www.julkaisufoorumi.fi

- Used for publication channel rankings as part of universities' funding model

- Integrated with other relevant databases (e.g. ISSN, DOAJ and ERIH)

- Contains structural data on ~30 000 journals and series, conference proceedings and book publishers

- Includes information on type (scholarly/non-scholarly), open access policy, peer-review practice, scientific fields, internationality etc…

- E.g. publication channels in "forestry"

SARJA/SERIES	TASO/LEVEL	NIMEKE/TITLE	ISSN1	IS
50681	1	AGRICULTURAL AND FOREST ENTOMOLOGY	1461-9555	14
50682	2	AGRICULTURAL AND FOREST METEOROLOGY	0168-1923	18
50696	1	AGROFORESTRY SYSTEMS	0167-4366	15
51487	1	APPLIED GEOGRAPHY	0143-6228	18
51539	1	APPLIED VEGETATION SCIENCE	1402-2001	16
59639	1	ARBORICULTURE AND URBAN FORESTRY	1935-5297	21
77177	1	AUSTRALIAN FORESTRY	0004-9158	23
52398	2	BIOMASS AND BIOENERGY	0961-9534	18
52945	1	CALIFORNIA AGRICULTURE	0008-0845	21
53968	2	COMPUTERS AND ELECTRONICS IN AGRICULTU	0168-1699	18
54251	1	CROATIAN JOURNAL OF FOREST ENGINEERINC	1845-5719	18
75117	1	ENERGY, SUSTAINABILITY AND SOCIETY		21
55721	2	EUROPEAN JOURNAL OF FOREST RESEARCH	1612-4669	16
56285	1	FOREST CHEMICALS REVIEW	1520-0191	
56286	3	FOREST ECOLOGY AND MANAGEMENT	0378-1127	18
56288	2	FOREST POLICY AND ECONOMICS	1389-9341	18
56289	1	FOREST PRODUCTS JOURNAL	0015-7473	23
56290	1	FOREST SCIENCE	0015-749X	19
59050	1	FOREST, TREES AND LIVELIHOODS	1472-8028	21
56292	1	FORESTRY	0015-752X	14

Vipunen statistics portal (www.vipunen.fi)

Publications A–E

Statistical year	2017

Publications	0=Does not meet the criteria	1=Basic Level	2=Leading Level	3=Top Level	Not evaluated
⊟ University	1,912	15,335	6,696	2,739	10,972
Aalto University	259	1,707	812	292	419
Hanken School of Economics	20	107	52	24	73
Jyväskylän yliopisto	117	1,124	580	233	1,078
Lappeenrannan teknillinen yliopisto	119	481	187	55	73
Tampere University of Technology	151	905	291	88	229
University of Eastern Finland	149	1,396	541	232	786
University of Helsinki	418	4,117	1,905	870	4,330
University of Lapland	57	194	90	12	284
University of Oulu	152	1,440	558	210	439
University of Tampere	87	1,222	502	225	760
University of the Arts Helsinki	25	63	8	9	127
University of Turku	219	1,925	824	337	1,795
University of Vaasa	36	149	55	44	101
Åbo Akademi University	103	505	291	108	478
⊞ University of Applied Sciences	191	384	65	18	5,573
⊞ Research institute	309	1,154	447	142	1,688
⊞ University Hospital	90	3,133	954	437	1,428
Grand total	2,502	20,006	8,162	3,336	19,661

VIRTA Publication Information Service

VIRTA Publication Information Service

o Implemented in 2016

o VIRTA publication information service is a data warehouse, "a data hub", making

   publication information usable for other services

   - o E. g. Publication information automatically to research funding reporting
   - o Implementation in Academy of Finland's online service in 2017
   - o The goal is to extend the concept to other services where researchers need to input data on their publications, such as research data services, research infrastructure services

- The Finnish research organizations use VIRTA to store a copy of publication information in their institutional CRISes or other publication databases
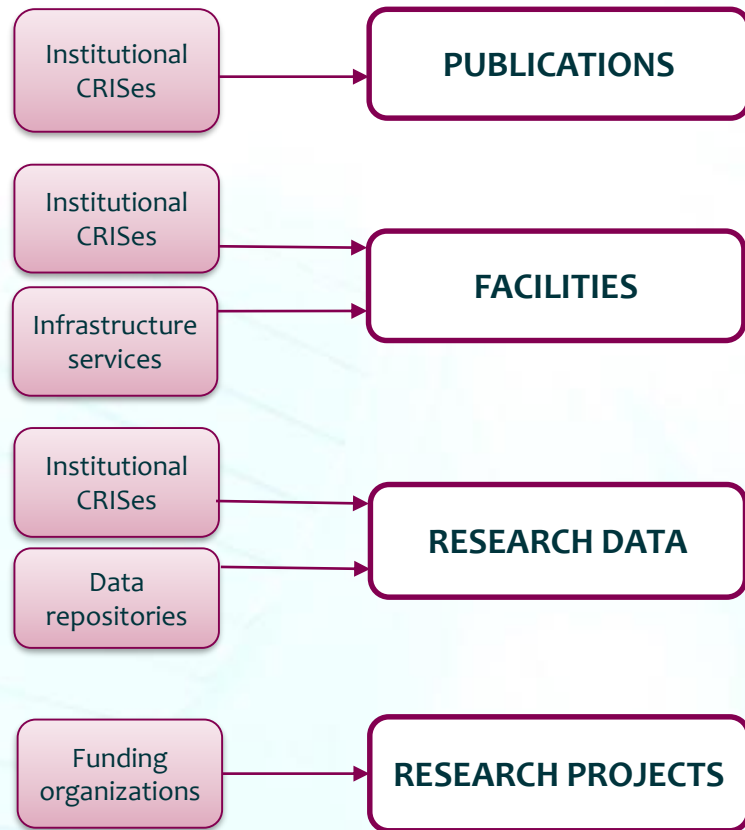
VIRTA Publication Information Service in a nutshell

Data sources	Master data in local CRISes or publication databases of HEIs, university hospitals, state research institutes
Data format	XML files (CSV converter provided for small organizations)
Data contents	The data must include required fields and fulfil certain technical criteria (mandatory fields, correct form of ISSN etc.)
Data transfer	From organizations via a secure and certified connection by using SFTP protocol and SSH authentication keys.
Updates	Updates automatically from local CRISes to VIRTA e.g. once a day. The frequency depends on the organizations, minimum being once a year.
Temporal coverage	All data from previous years to present can be transferred. Statistics compiled once a year.
Data validation	Duplicates, faults, inter-organizational co-publications and Publication forums identified automatically and real time.
Data use and availability	All metadata are published real time in JUULI portal: www.juuli.fi . The data can be used in various systems and services via REST API and OAI-PMH.

Towards a Finnish Research Information Hub

Current situation in Finland: national databases as silos

| Institutional CRISes | → | **PUBLICATIONS** | **VIRTA Publication Information Service**
 Publication information from 54 Higher education institutions and state research institutes. Metadata: www.juuli.fi and statistics: www.vipunen.fi |

| Institutional CRISes
 Infrastructure services | → | **FACILITIES** | **Finnish Research Infrastructures**
 The infrastructures database compiles metadata on national and institutional research facilities, equipment, materials and services. Currently at piloting stage. http://infras.openscience.fi/ |

| Institutional CRISes
 Data repositories | → | **RESEARCH DATA** | **Etsin - Research Data Finder**
 Etsin contains metadata on research datasets. Integrated part of the Finnish Fairdata Service which include also a preservation service, a metadata warehouse, and a metadata input service. The research data finder can be used at https://etsin.avointiede.fi |

| Funding organizations | → | **RESEARCH PROJECTS** | **National Research Funding Database**
 The database aims to gather all information related to research funding process. The information will be harvested from the funders' registers. In terms of funding decisions, the service will be implemented by the end of 2018. https://wiki.eduuni.fi/display/CSCTTV/National+Research+Funding+Database |

INTEROPERABILITY OF SERVICES

"Services have up-to-date information available"

Metadata can be used and enriched by national and local systems (e.g. funders). Common vocabularies and data models ensure the consistency of information.

Connected

Metadata

INFORMATION PROVIDERS

"Once the information has been input it can be utilized in several systems"

UNIVERSITIES

UNIV. OF APPLIED SCIENCES

RESEARCH INSTITUTIONS

FUNDERS

PUBLISHERS

INFRASTRUCTURES

REPOSITORIES

ORCID

Automated information flows

INFORMATION USERS

"Open access to public information"

STATISTICS AND VISUALIZATION
Up-to-date facts and figures about research

NEW APPLICATIONS

Open information to be utilized in new applications supporting the researchers work. Eg. CV-tools

RESEARCH INFORMATION HUB 2020

Connected metadata for publications, research data and other research activities, research infrastructures, funding, researchers and organizations
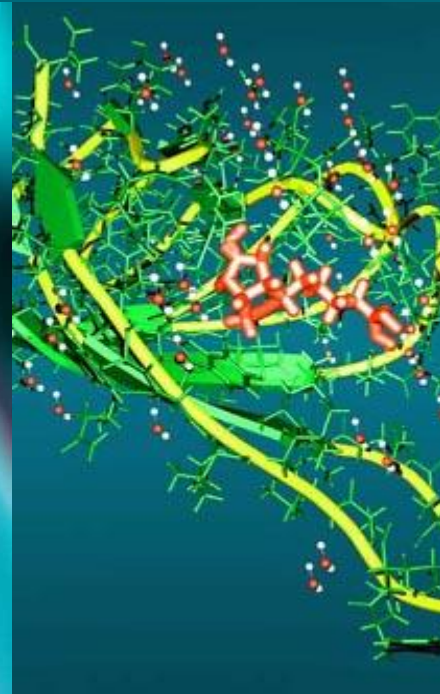
INFORMATION RETRIEVAL

"Overview for research in Finland"

Research.fi portal will provide a comprehensive picture of the research being conducted in Finland.

Search criteria can be specified in, for example, field of science, thematic areas, ongoing projects or experts.

CSC

www.research.fi

ENRESSH-VIRTA pilot:
Applying VIRTA concept to European level

VIRTA-ENRESSH-POC

- One of the goals of the European Network on Research evaluation in Social Sciences and Humanities (www.enressh.eu) is to design a roadmap for a European database for SSH output

  ➤ In view of this task, a proof of concept VIRTA-ENRESSH-POC of a European database for publications was set up

- VIRTA-ENRESSH-POC is a collaborative pilot project exploring a potential cost-efficient solution for the integration of European research information
  - Founded on the efforts made at national level
  - Especially for SSH but not excluding other fields
  - Carried out between 3/2017-3/2018
  - Involved partners from Belgium, Finland, Norway, and Spain
  - The technical solution builds on the strengths of the Finnish VIRTA Publication Information Service

VIRTA-ENRESSH-POC– practical issues

Participating countries reported their complete publication metadata from the **years 2014-15**:
- **Norway**: University of Oslo
- **Flanders**: University of Antwerpen
- **Spain**: Universidad Carlos III de Madrid (UC3M)
- **Finland**: University of Helsinki, University of Jyväskylä, Tampere University of Technology

- 52 948 publications in total

- Finland and Oslo cover all fields, Antwerpen and Madrid only SSH

- The data collected in the pilot has its highest quality and consistency in terms of the bibliographic data meanwhile the classifications, semantics and typologies vary
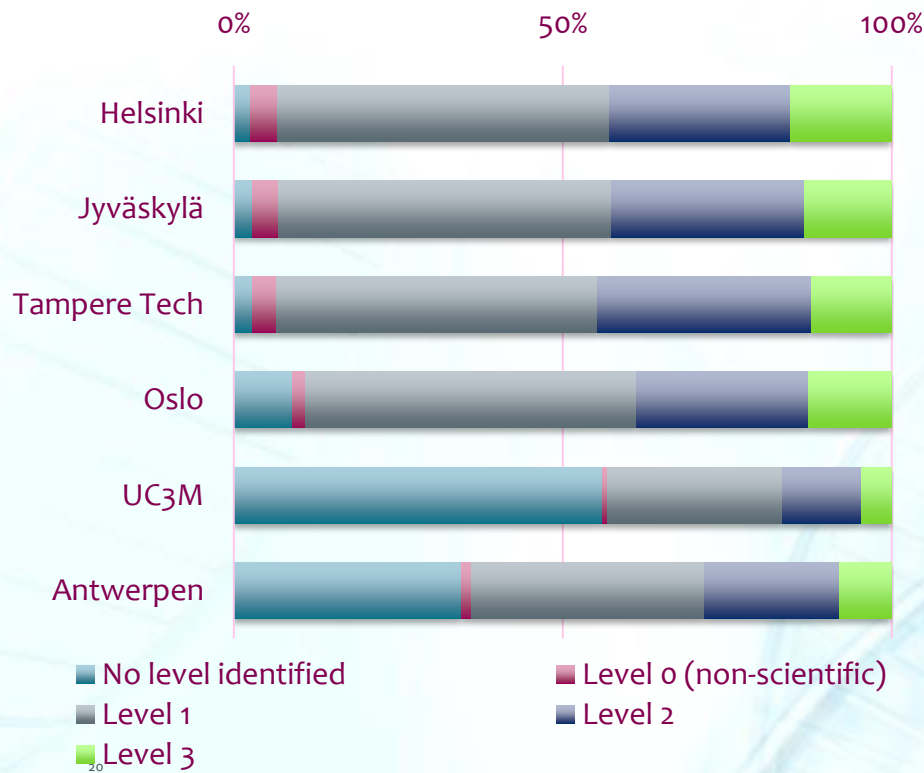
Issues of data comparability: Publication type mapping

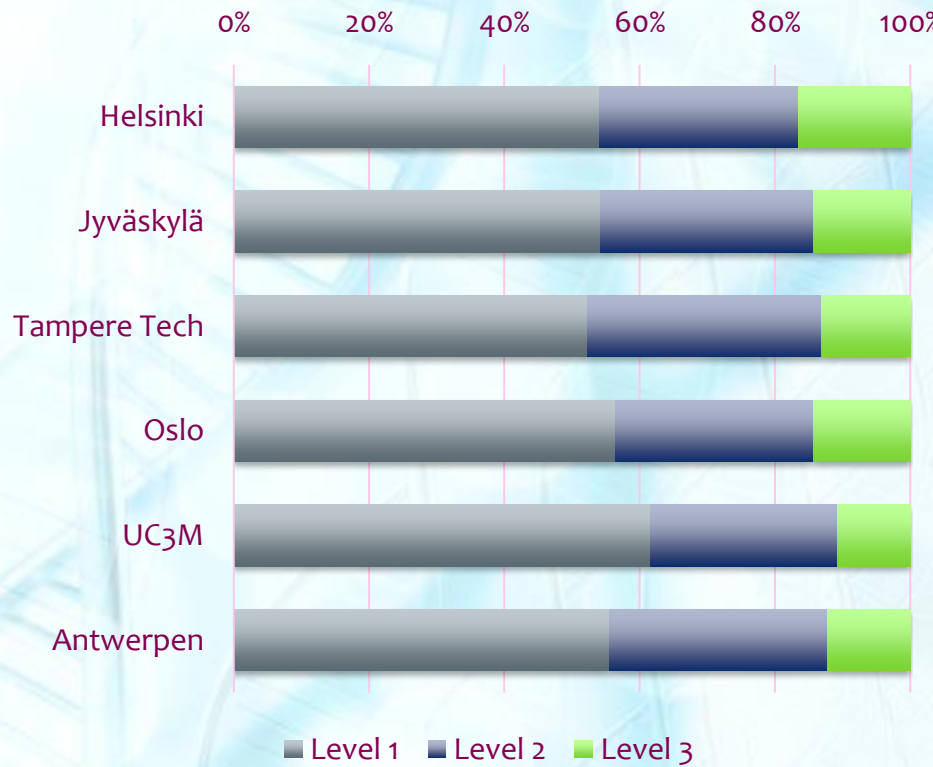|	Finland / Madrid	Flanders 1=peer-reviewed / 0 = non peer-reviewed		Norway
Peer-reviewed articles	A1 Journal article, original research	VABB-1: journal article	1	3= Article in series (ISSN)
	A2 Review article			
	A3 Book section	VABB-4: book chapter	1	2= Article in book (no ISSN)
	A4 Conference proceedings	VABB-5: proceedings paper	1	
Non peer-reviewed articles	B1 Non-refereed journal articles	VABB-1: journal article	0	
	B2 Book section	VABB-4: book chapter	0	
	B3 Non-refereed conference proceedings	VABB-5: proceedings paper	0	
Monographs	C1 Book	VABB-2: monograph	1	1= Monograph
	C2 Edited book	VABB-3: edited book	1	
Professional	D1 Article in a trade journal			
	D2 Article in a professional book			
	D3 Professional conference proceedings			
	D4 Development or research report			
	D5 Textbook, professional manual or guide			
	D6 Edited professional book			
Popular	E1 Popularised article, newspaper article			
	E2 Popularised monograph	VABB-2: monograph	0	
	E3 Edited popular book	VABB-3: edited book	0	

Authorized publication channel registries as a solution for more structured and comparable data?

- For all publications reported in the POC, the publication channel was automatically detected against the Finnish Publication Forum database (JUFO)

- Corresponding registries for publication channels are maintained also in other countries, such as Norway, Denmark and Belgium (Flanders).

- "The Nordic List" funded by NordForsk has implemented a common Nordic registry of authorized research publication channels integrating databases in Norway, Finland and Denmark.

VIRTA-ENRESSH-POC: Publications by Finnish Publication Forum levels

Journal articles, all

Journal articles, identified as level 1-3

Legend:
- No level identified
- Level 0 (non-scientific)
- Level 1
- Level 2
- Level 3

Next steps

- Collaboration to be continued both in the framework of 1) ENRESSH and 2) Nordic countries

- In a Nordic meeting in Finland in May 2018, the stakeholders of national CRIS systems in Nordic countries decided to continue to

  1. both contextual and technical development of ENRESSH-VIRTA
  2. further development og the Nordic list and
  3. produce a biannual bibliometric analysis comparing Nordic institutions

- The cooperation at Nordic level does not exclude other European countries and the goal is also to extend the POC to more countries.

- The next phase also includes investigation of the use of CERIF by EuroCRIS in import and export

- Cooperation to be strengthened also with other initiatives that aim at the integration of publication metadata at European level

Hanna-Mari Puuska

Development Manager
hanna-mari.puuska@csc.fi
+358 50 3818 568

https://www.facebook.com/CSCfi

https://twitter.com/CSCfi

https://www.youtube.com/c/CSCfi

https://www.linkedin.com/company/csc---it-center-for-science