

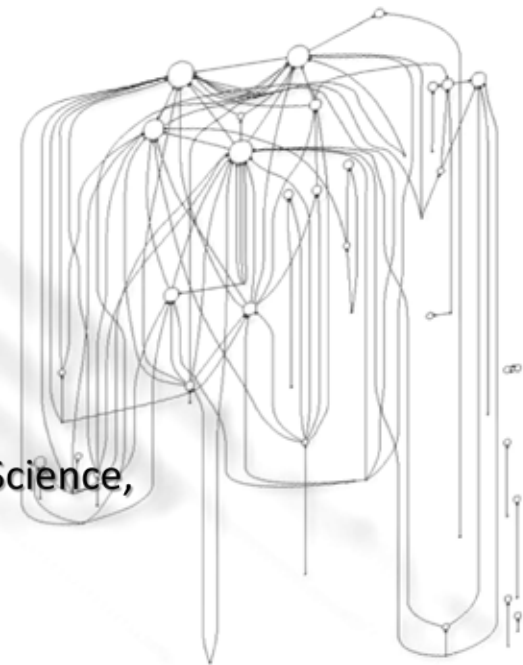
CRIS data for research on research

Gustaf Nelhans,

Swedish School of Library and Information Science,
University of Borås

CRIS2018

Umeå, Sweden June 15th, 2018



Overview

- Introduction
- Researcher's experience
- 'Lost', 'found' and 'made' data
- Results and reflections

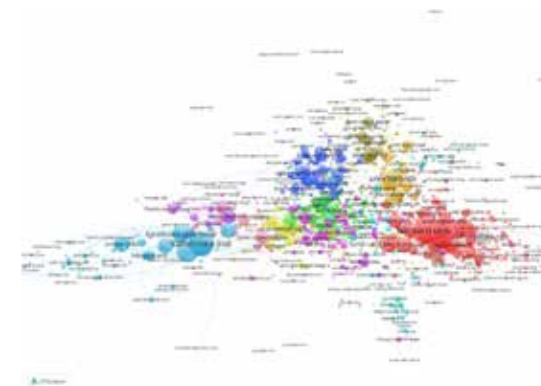
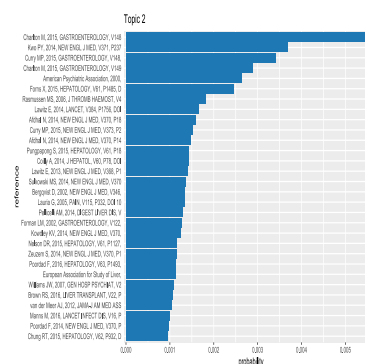
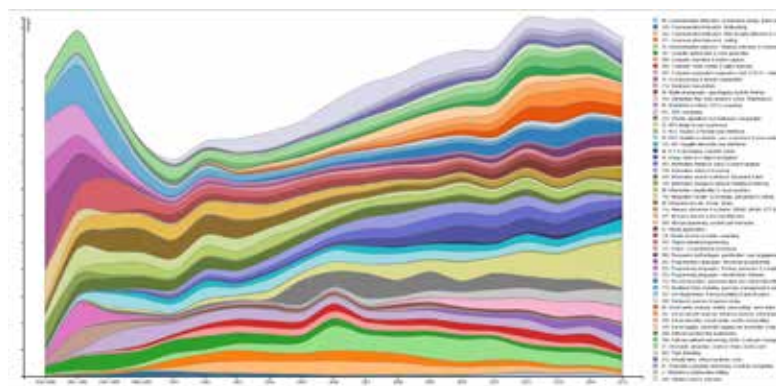
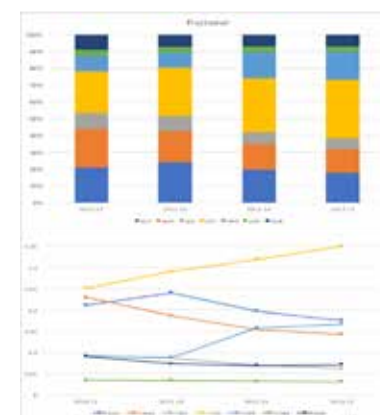
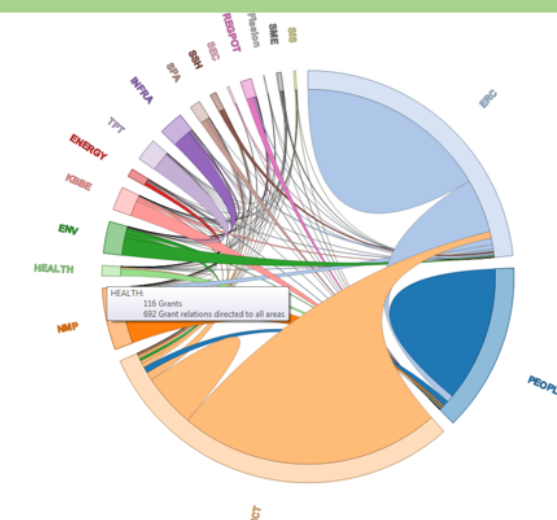
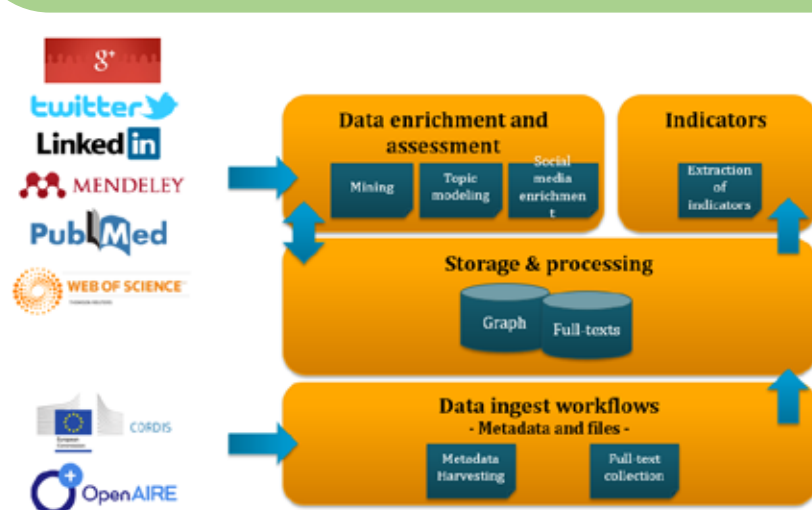
Data for Impact: overview

- Call: CO-CREATION-08-2016-2017: Better integration of evidence on the impact of research and innovation in policy making
- Expected results:
 - **Improved monitoring of R&I activities:** new indicators for assessing research and innovation performance, including the impact of research and innovation policies
 - **Prove value to the society:** determining the societal impact of research and innovation funding in order better to justify research and innovation spending
- **Data driven approach**
- **Partners:**
 - *PPMI* (Lithuania), *CNR* (Italy), *Athena RC* (Greece), *Fraunhofer* (Germany) and, *Univ Borås* (Sweden).

Objective 1: define, develop, analyse new indicators for assessing the performance of EU and national R&I systems.

Objective 2: gather data at input, throughput, output and impact levels, derive facts and understand impact on health-related challenges

Objective 3: perform community-driven validation and develop user-centered tools



Project data

- Cordis (FP7 + H2020)
- **National funders (selection)**

Project data:

- *Call*
- *Project description*
- *Final report*
- *Results*

Methodology (simplified)

Extract information through text mining, manual and automatic annotation.

Create

Concepts, entities, expected results, publications, patents...

Topic modelling



NEWS &

NMTRYPI

Project ID: 6032
Funded under:

New Medicines

From 2014-02-01

Project details

Total cost:
EUR 7 639 176,
EU contribution:
EUR 5 901 411,
Coordinated in:
Italy

Objective

The infectious diseases that threaten people's lives and health need more efficient delivery systems. New medicines are needed now and in the future. This project uses a highly innovative approach to develop new medicines.

Related information

[Result In Brief](#)

[Report Summary](#)

[Open Access publications](#)
(15)

via 

Related information

Result In Brief

[Towards new treatment options for trypanosomatid infections](#)

Report Summaries

[Final Report Summary - NMTRYPI \(New Medicines for Trypanosomatid Infections\)](#)

Open Access publications (15)



[Methoxylated 2'-hydroxychalcones as antiparasitic hit compounds](#)

Borsari, Chiara; Santarem, Nuno; Torrado, Juan; Ollas, Ana Isabel; Corral, Maria Jesús; Baptista, Catarina; Gul, Sheraz; Wolf, Markus; Kuzikov, Maria; Ellinger, Bernhard; Witt, Gesa; Gribbon, Philip; Reinshagen, Jeanette; Linciano, Pasquale; Tait, Annalisa; Costantino, Luca; Freitas-Junior, Lucio H.; Moraes, Carolina B.; Bruno dos Santos, Pascoalino; Alcântara, Laura Maria; Franco, Caio Haddad; Bertolacini, Claudia Danielli; Fontana, Vanessa; Tejera Nevado, Paloma; Clos, Joachim; Alunda, José María; Cordeiro da Silva, Anabela; Ferrari, Stefania; Costi, Maria Paola2017

doi:10.1016/j.ejmech.2016.12.017

[Application of a simple quantum chemical approach to ligand fragment scoring for Trypanosoma brucei pteridine reductase?1 inhibition](#)

Jedwabny, Wiktoria; Panecka-Hofman, Joanna; Dyguda-Kazimierowicz, Edyta; Wade, Rebecca C.; Sokalski, W. Andrzej

Springer International Publishing Journal of Computer-Aided Molecular Design 2017

pmc:PMC5570812, doi:10.1007/s10822-017-1035-4

[The Future of Drug Development for Neglected Tropical Diseases: How the European Commission Can Continue to Make a Difference](#)

Pierce, Raymond J.; Maedaugall, Jane; Leurs, Rob; Costi, Maria Paola2017

doi:10.1016/j.pt.2017.04.007

Coordinator

REGGIO EMILIA

Education Establishments

Costi

icine

CELULAR-IBMC

ERGIA E MATERIAS ASSOCIACAO

DERUNG DER ANGEWANDTEN

FORSCHUNG E.V.

Subjects

[Life Sciences](#)

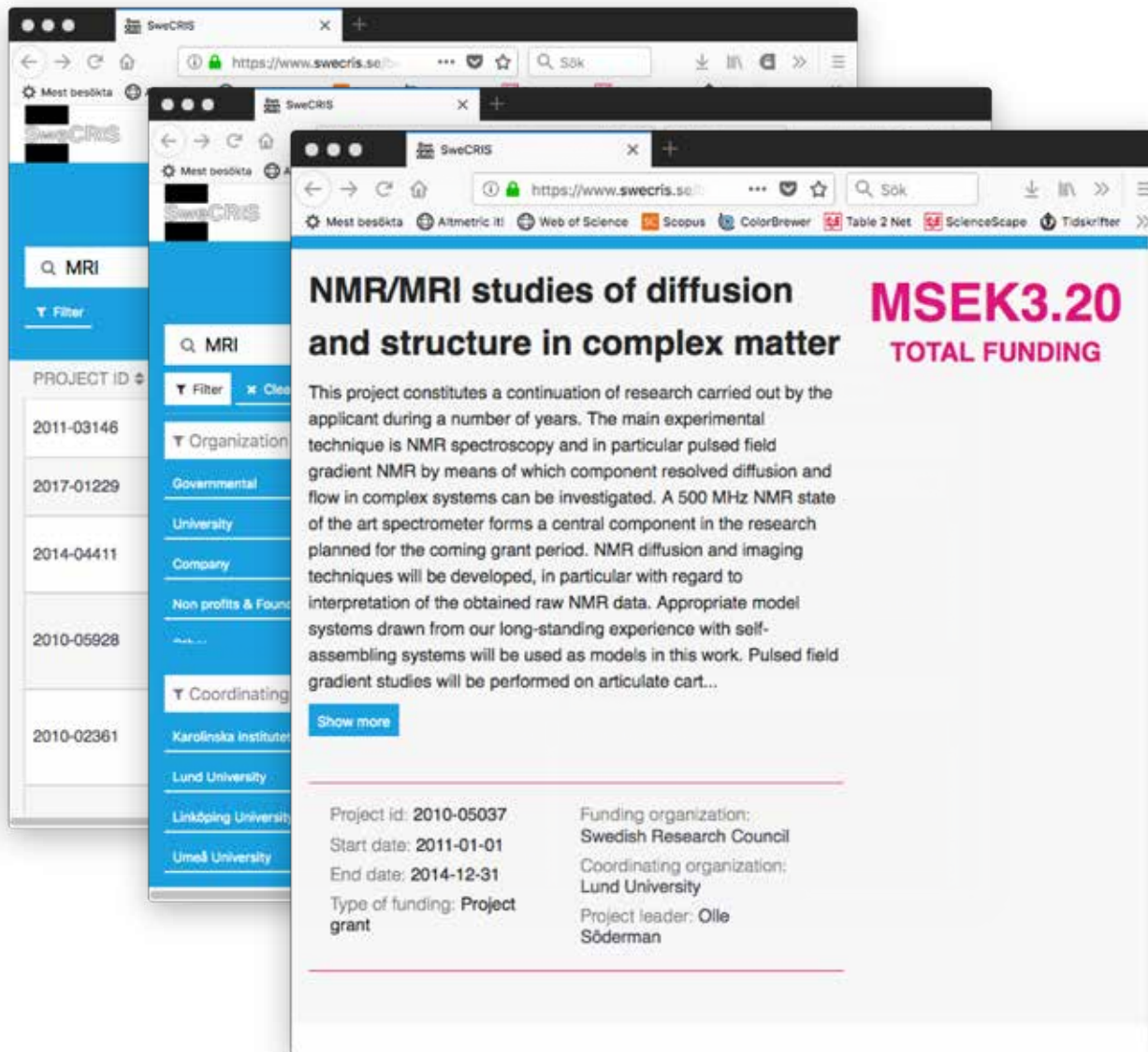
Share this page



‘made’

‘found’

SweCRIS interface



Export fields:

_id
_type
abstract.sv
abstract.en
dates.start_date
dates.end_date
identifier_short
organizations.funding.0.id
organizations.funding.0.sv
organizations.funding.0.en
organizations.coordinating.0.id
organizations.coordinating.0.sv
organizations.coordinating.0.en
people.project_leaders.0.id
people.project_leaders.0.name
people.other_personnel
tags.0.id
tags.0.sv
tags.0.en
title.sv
title.en
total_funding
type_of_awards.sv
type_of_awards.en
Other personel

Issues with SweCRIS

Practice:

- No **Boolean searches** (AND, OR, NOT, NEAR, ...)
- Filters are **ordered** by counts, there is no way to sort them.
- The **subject** codes (based on OECD) are on three levels (260 different codes in all), you cannot select groups of subjects.
- **Export** to JSON or csv should work well.

Metadata:

- **Project ID:s not uniform**
 - Actual project ID:s 2014-24, 2014-240
 - In SweCRIS: 2014-**000**24, or 2014-**00**240 (to match current practice?)
- **Data from earlier databases** is lost or changed
- *Case: FORMAS*

Funder: Formas old database

SweCRIS type	Formas Type	relationship
Project grant	Projektstöd	<i>Similar</i>
Grant for positions or stipends	Mobilitet	<i>specified type: mobility</i>
Project grant	Hållbar stadsutveckling	<i>specified call</i>
Grant for positions or stipends	Forskarassistent	<i>specified position</i>
Project grant	Doktorand	<i>ditto, but called PG in SweCRIS</i>
Project grant	Hållbar stadsutveckling	<i>spec call</i>
Research environment	STARKA FORSKN MILJÖER	<i>Similar</i>

- Project types transformed

- Missing information:

Intrascientific report (in English) (with publ list)

Popular report (in Swedish)

Sex of the applicant

List of projects that the current project is a continuation from. E.g. 2006-958, 2009-1460

Alternative steps (ongoing, Vinnova)

- Contact funders, ask for Final reports
 - *Principle of public access to official documents.* – Offentlighetsprincipen
 - Works well for single requests
- Quite unsuccessful for larger sets of data
 - Pdf format, scanned, (not) OCR:ed
- GDPR!
 - ORCID is *personally identifiable information*

Match publications from CrossRef

JSON	Rådata	Rubriker
Spara	Kopiera	
▼ country:		
resource:	"http://sws.geonames.org/2661886/"	
▼ inScheme:		
resource:	"http://data.crossref.org/fundingdata/vocabulary"	
created:	"2014-05-28T14:49:36.000000"	
▼ prefLabel:		
Label:		
literalForm:		
lang:	"sv"	
content:	"Vetenskapsrådet"	
about:	"http://data.crossref.org/fundingdata/vocabulary/Label-769410"	
modified:	"2017-08-17T14:54:58.000000"	
▶ altLabel:	[...]	
id:	"http://dx.doi.org/10.13039/501100004359"	
▼ continuationOf:		
resource:	"http://dx.doi.org/10.13039/501100006310"	
fundingBodyType:	"gov"	
fundingBodySubType:	"government non-federal"	

JSON	Rådata	Rubriker
Spara	Kopiera	
status:	"ok"	
message-type:	"work-list"	
message-version:	"1.0.0"	
▼ message:		
facets:	{}	
total-results:	12988	
▼ items:		
▼ 0:		
indexed:	{...}	
reference-count:	257	
publisher:	"Wiley"	
license:	[...]	
▼ funder:		
▼ 0:		
DOI:	"10.13039/501100004359"	
name:	"Vetenskapsrådet"	
doi-asserted-by:	"publisher"	
▼ award:		
0:	"2016-01381"	
▶ 1:	{...}	
▶ 2:	{...}	
▶ content-domain:	{...}	
▼ short-container-title:		
0:	"Acta Physiol"	
DOI:	"10.1111/apha.13080"	
type:	"journal-article"	

CrossRef data

Amount of projects and publications for each funder:

Funder	Projects (expected)	Unique DOI:s
SRC	1016 (2520)	1852
Formas	203 (198)	255
Vinnova	64 (1634)	104

SRC Swedish Research Council

Formas The Swedish Research Council Formas

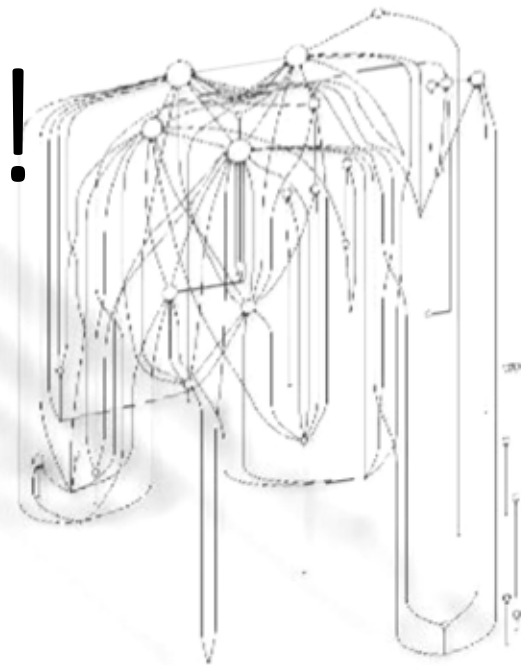
Vinnova VINNOVA (Sweden's innovation agency)

Reflections

- How well does the SweCRIS provide relevant information ("made") data
 - How well does SweCRIS give opportunities for "found" data.
 - What gets lost?
-
- Jensen, K.B. 2012: Lost, Found, and Made – Qualitative Data in the Study of Three-Step Flows of Communication In I. Volkmer (Ed.) *The Handbook of Global Media Research*, 433–50. Wiley-Blackwell <http://dx.doi.org/10.1002/9781118255278.ch25>.

Thank you!

gustaf.nelhans@hb.se



Data for Impact has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 770531.