

Showing It all – A new interface for finding all Norwegian research Output

Agenda

- CRISTin at a glance
- Old CRIS / New CRIS
- New Search Engine
- Goals
 - In details...
- Summary

What to remember

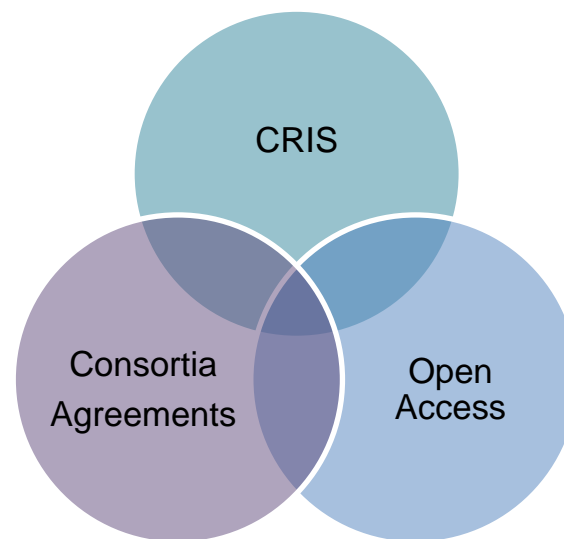
- The Norwegian CRIS is currently being reimplemented and will arise as a brand new application
- The CRIS is being mixed with Open Access throughout the system
- The search index is the (technical) key element in this

CRIS_{tin} at a glance



What we do in CRIS_{tin}:

- Maintaining the national CRIS
- Coordinating Open Access in Norway
- Consortia Licensing Agreements



Old CRIS vs. New CRIS

1. The Norwegian CRIS needs renewal.
The CRIS shall now serve as an *information* system,
rather than a *documentation* system



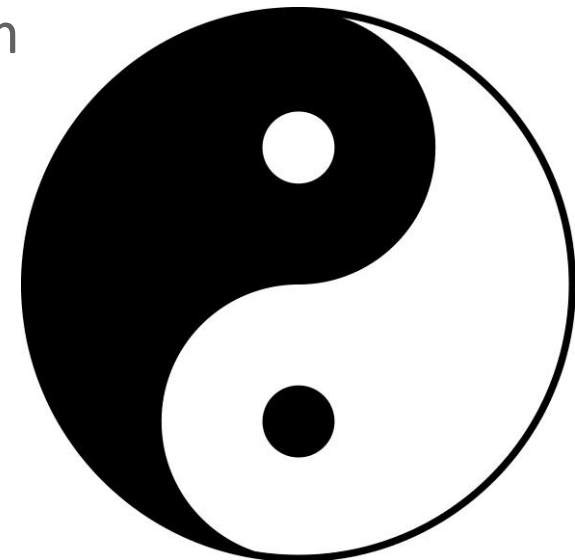
Old CRIS vs. New CRIS



2. Time to merge NORA with the CRIS

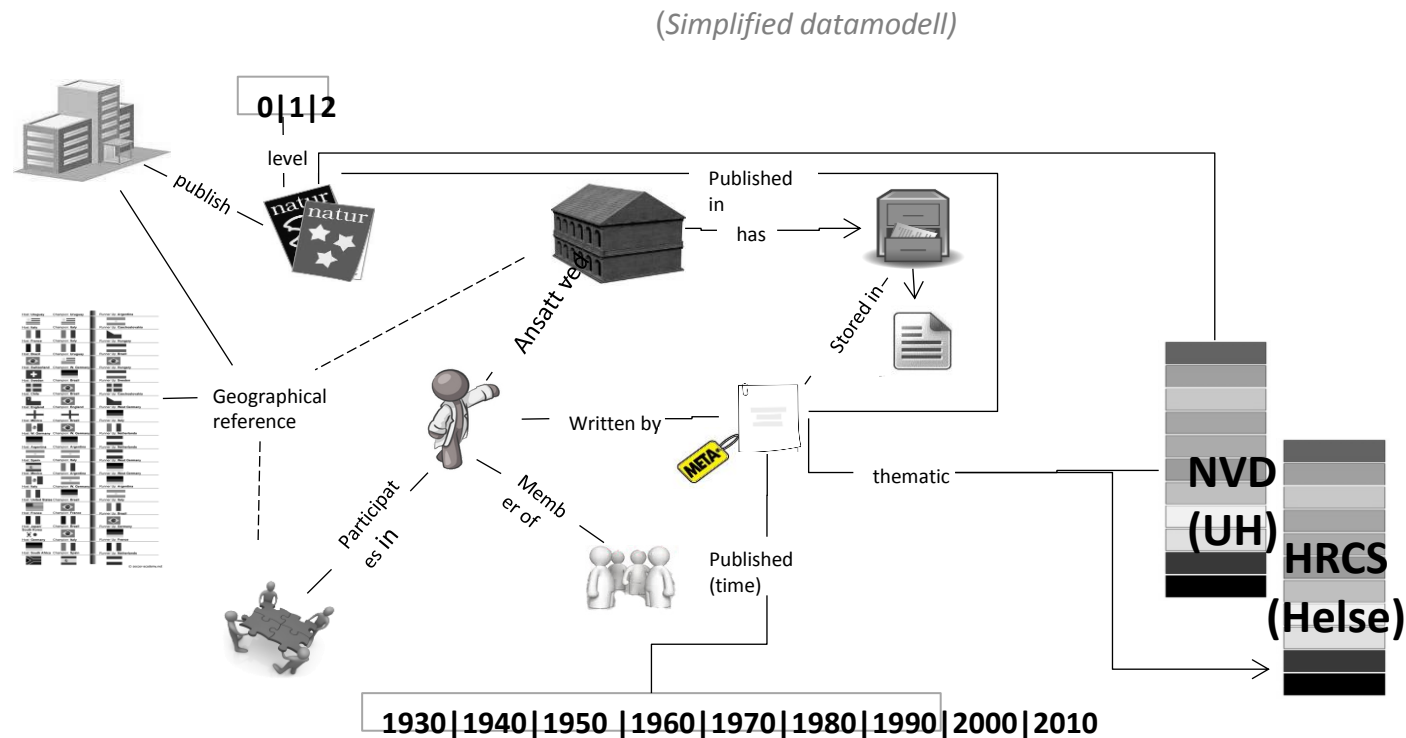
(Metadata in the CRIS with full-text from all Norwegian archives)

- The CRIS is a national database for Norwegian research output
- NORA (Norwegian Open Research Archives) is a search index containing all fulltext-content from all Norwegian institutional repositories.



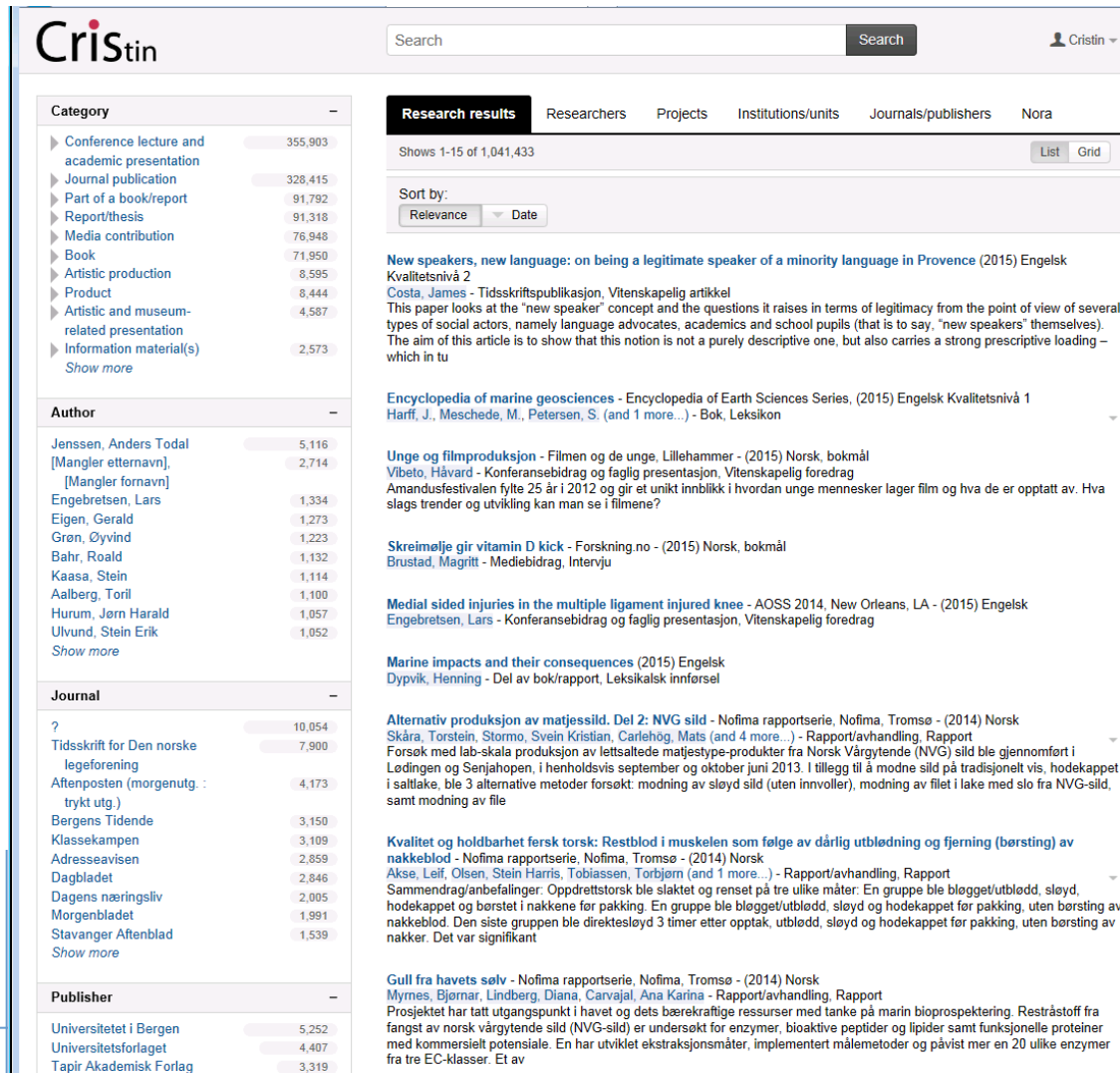
New Cris: Main objectives

- A suitable tool for the researcher to administer his/her own information
- A better way of finding and displaying Norwegian research output
- To showcase Norwegian publications, projects and researchers *in context*
- A smooth dissemination to national and global services.



By the numbers

Cristin



- 1 000 000 + results (330 000 are scientific/peer-reviewed)
- 56 000 researchers
- 10 000 projects
- 5000 institutions/departments etc.
- 46 000 journals/publishers
- 80 000 full-text documents (NORA)

New Search Engine



Dedicated software

- Very fast compared to SQL-based search

Technology SOLR:

- Open Source-framework with a rock-solid eco-system
- By far the most popular tool
- Faceted search
- Scalability



Goals

A search index that meets the following requirements:

- a) Integrating NORA with the CRIS
- b) Identifying the potential of Open Access in the CRIS
- c) Building a tool for validating data posts and data integrity
- d) A flexible basis for developing new services in the future



a) Integrating NORA / CRIS

- A challenging task
- CRIS as a uploading service for the archives
(The metadata is already present in the CRIS)

Fulltext Available	
<input type="checkbox"/> False	1,037,980
<input type="checkbox"/> True	3,453



b) Identifying the potential of Open Access in the CRIS

Two new ingredients needed:

- Sherpa/Romeo (Green Open Access-journals)
- DOAJ (Gold Open Access-journals)

RoMEO		–
green		91,314
yellow		38,155
white		25,612
gray		10,918
blue		7,593
'Open Access'		–
True		12,492



c) Building a tool for validating data posts and data integrity

- The SOLR index serves as a listening post for every record in the CRIS/NORA-systems
- Each post validated according to integrity rules
- Tells data owner what should or could be corrected.



c) Examples

(Research Results)

Cristin

Year	
420	2
2209	1
2202	4
2105	1
2000	

Journal	
?	10,054
Tidsskrift for Den norske legeforening	7,900

Author	
Jenssen, Anders Todal	5,116
[Mangler etternavn], [Mangler fornavn]	2,714

d) A flexible basis for new services in the years to come

- Data from the index can be used in various ways:
 - RSS-generator/Web-services
 - Map services
 - Timelines
 - Tag-clouds
 - Landing-pages within the app
- And perhaps more important
 - All the services we haven't thought of



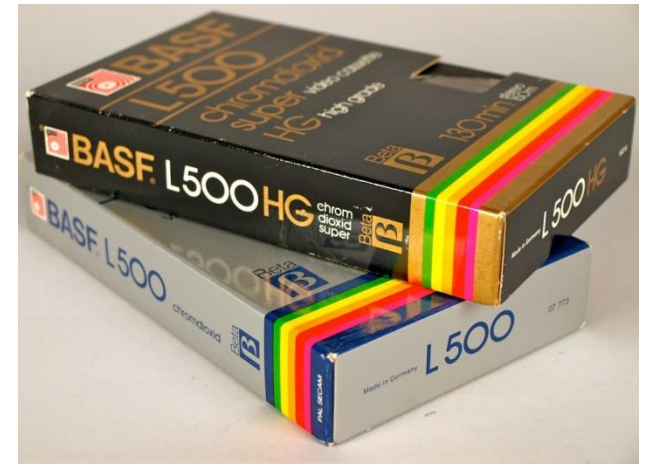
Challenges

- Technical issues are not the main challenge
- Missing data in the CRIS
(from an *information* system point of view)
- Different practices for registration of data
(Archives annotate data in many different ways)
- Difficult task to identify overlap between CRIS and NORA (Overlap is 3400, it should be 40 000)



Release dates

- Fully functional beta implemented
- Service will be released together with the rest of the CRIS
(When? We don't quite know yet)
- Log-server under implementation



Summary

What to remember (again)

- The Norwegian CRIS is currently being reimplemented and will arise as a brand new application
 - Release date not decided
- The CRIS is being mixed with Open Access throughout the system
 - Great potential in merging metadata in the CRIS with the full text documents in the IRs, making Norwegian research output more searchable and visible
- The search index is the (technical) key element in this
 - Validating data-tool essential for improving the CRIS
 - Potential for many new services in the CRIS in the years to come.