

# *Fedora Content Modeling for Improved Services for Research Databases*

*Open Repositories 2009*

Mikael Karstensen Elbæk

Alfred Heller

Gert Schmeltz Pedersen

DTU Library, Technical Information Center

@ DTU, Technical University of Denmark

Funded partly by DEFF, Denmark's Electronic Research Library

# *Introduction*

- An experimental re-implementation of the ORBIT research database of the Technical University of Denmark, DTU, is based on Fedora.
- The backbone consists of content models for primary and secondary entities and their relationships.
- By adopting such an abstract data model, the platform enables new and improved services for researchers, librarians and administrators.
- DTU Library has a broad and deep knowledge of Fedora and has contributed to its development with GSearch, the generic search service module.

# Overview

- ➔ Current ORBIT highlights
- ➔ Analysis ...
- ➔ Towards a Fedora-based ORBIT ...
- ➔ Conclusion

# Current ORBIT highlights

- ORBIT<sup>[1]</sup> - **O**nline **R**esearch data**B**ase **I**n **T**echnology - is the official research database of the Technical University of Denmark, DTU.
- It is developed by DTU Library.
- ORBIT provides all information on scientific, popular and educational publications used in research activities at DTU.
- ORBIT serves also as an Institutional repository and a Current Research Information System (CRIS) for statistical and administrative purposes.
- ORBIT is the central tool for departments, sections and the individual researchers to present their publications and externally funded research projects.
- <sup>[1]</sup> <http://orbit.dtu.dk>

- The ORBIT search interface gives public access to information about research and development activities at the University and provides an overview of publications to which DTU researchers have contributed as authors or co-authors.
- ORBIT also provides the possibility of cross search on publications, projects, resources, events, department profiles and personal staff profiles.
- ORBIT exposes data through OAI-PMH and SRU/SRW, delivering research data to the national research database, the European project DRIVER and Google Scholar.
- ORBIT targets researchers, administration and management, the business community, the press, teachers and information brokers.

Login

# ORBIT The DTU Research Database



- Welcome
- Publications
- Projects
- Events
- Resources
- Departments
- Staff
- Search All
- myBit

## Welcome to ORBIT - Online Research Database In Technology

- Help
- Links
- About / Coverage
- Contact

→ Courses in ORBIT

Search Publications	Search Projects
Search Events	Search Resources
Search Staff	Search All

- Departments List
- Personal Collection
- Register Data

Orbit Home - Mozilla Firefox

Ejer Rediger Vis Historik Bogmærker Funktioner Hjælp

http://orbit.dtu.dk/Publications,\$Form.sdirect

Mest besøgte Igang med Firefox Seneste nyheder

Orbit Home DTU link http://orbit.dtu.d...e=download\_records

Login

**ORBIT The DTU Research Database** DTU

Welcome Publications **Projects** Events Resources Departments Staff Search All myBit

**Publications - Basic Search**

Find:

Include:  DTU records  Department records  Personal records

1 records in  format with  rec/page sorted by

Elbæk, Mikael Karstensen; Nondal, Lars 2007

**The Library as a mediator for e-publishing** : A case on how a library can become a significant factor in facilitating digital scholarly communication and open access publishing for less Web--savvy journals  
 In: First Monday, vol: 12(10) (2007). University of Illinois at Chicago University Library  
 Type: Journal article  
 Resolvers: DTU

**File attachments**  
 Post-print (Version 1. Access to All. Size: 81 KB)

- XML
- Short HTML
- Long HTML
- OpenOffice Document (short)
- RIS (Reference Manager)
- RIS (as text)
- BibTeX
- BibTeX (as text)
- Reference (APA)


Færdig

Start Indbakke - Microsoft Ou... 14 Firefox OR09 Microsoft PowerPoint - [...] Conference proposal for... DA 12:46

Orbit Result Page - Mozilla Firefox

http://orbit.dtu.dk/Publications/resultSetTable.recordLink.sdirect?sp=209647

Print

<b>Title</b>	The Library as a mediator for e-publishing : A case on how a library can become a significant factor in facilitating digital scholarly communication and open access publishing for less Web-savvy journals
<b>Resolvers</b>	DTU  link
<b>File attachments</b>	Post-print (Version 1. Access to All. Size: 81 KB)
<b>Person/s:</b>	<p>Author: <a href="#">Elbæk, Mikael Karstensen</a> (Cwisno.: 41294)            Technical University of Denmark, Technical Knowledge Center of Denmark, ePub            Email: <a href="mailto:mke@dtv.dk">mke@dtv.dk</a></p> <p>Author: <a href="#">Nondal, Lars</a>            Copenhagen Business School, Library,            Email: <a href="mailto:ln.lib@cbs.dk">ln.lib@cbs.dk</a></p>
<b>Status</b>	Published. [ISI-indexed: No]
<b>Abstract (English)</b>	Denmark is a small country but with a large and diverse scholarly publishing environment. There are many small journals, mostly in English. A majority of these see the potential in online publishing but do not have the resources and capabilities to do so. Furthermore they have a conservative business model (i.e. finances!) that doesn't encourage an open access publishing strategy. The Copenhagen Business School (CBS) Library provides a low risk environment for small journals related to the business school to make a gradual transition to e-publishing/e-archiving. Whether they at a later stage take the full step towards open access publishing remains to be seen. It is our firm belief that this gradual transition is essential for these smaller journals to eventually arrive online at all.
<b>Publ. in</b>	First Monday
<b>Imprint</b>	University of Illinois at Chicago University Library
<b>Description</b>	2007, Vol: 12(10),
<b>Language</b>	English
<b>Links</b>	<a href="http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1958">http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1958</a>
<b>Keyword/s</b>	Open Access E-publishing
<b>Type</b>	Journal article
<b>Indicators</b>	Popular ; Peer reviewed ; International
<b>DTU Data</b>	Annual report year: 2007 Credited department: DTV
<b>Record Id.:</b>	209647
<b>Owner:</b>	41294

Færdig

Start | Indbakke - Microsoft Ou... | 15 Firefox | OR09 | Microsoft PowerPoint - [... | Conference proposal for... | DA | 12:51

DTU link - Mozilla Firefox

Ejer Rediger Vis Historik Bøgermærker Funktioner Hjælp

http://sfx.cvt.dk:3210/sfx\_local?issn=&volume=12&issue=10&date=2007&title=First Monday&title=The Library as a mediator for e-publishing

Mest besøgte Igang med Firefox Seneste nyheder

Orbit Home DTU link http://orbit.dtu.d...e=download\_records

**DTU** Technical Information Center of Denmark

**DTU link services for this record:**

**Title:** The Library as a mediator for e-publishing  
**Source:** First Monday [1396-0466] yr:2007 vol:12 iss:10

**Full Text**

Full text available via **DOAJ Directory of Open Access Journals**  
 Year:  Volume:  Issue:  Start Page:  **GO**  
 Available from 1996

Full text available via **Free E- Journals** **GO**  
 Available from 1996 volume: 1

**Web Service**

Feedback to **DTU Link** **GO**

**Web Search**

Search for related literature in **Google Scholar**  
 **GO**  
 Search Terms:

Find related information in **a Web Search Engine**  
 **GO**  
 Search Terms:

**Journal info**

Show journal info using **ulrichsweb.com** **GO**

A service provided by The Technical Information Center of Denmark  
 © 2004 SFX by Ex Libris (USA) Inc.

Færdig

Start Indbakke - Microsoft Ou... 14 Firefox OR09 Microsoft PowerPoint - [...] Conference proposal for... DA

Orbit Home - Mozilla Firefox

http://orbit.dtu.dk/Projects,\$Form.sdirect

Mest besøgte Igang med Firefox Seneste nyheder

Login

## ORBIT The DTU Research Database

**DTU**

[Welcome](#)
[Publications](#)
[Projects](#)
[Events](#)
[Resources](#)
[Departments](#)
[Staff](#)
[Search All](#)
[myBit](#)

**Search**

- Basic
- Advanced

**General**

- Help
- Links
- About / Coverage
- Contact
- Courses in ORBIT

**External links**

- Riso publications before 2007
- The Danish Research Database

### Projects - Basic Search

Find:

Include:  DTU records  Department records  Personal records

4 records in  format with  rec/page

<input type="checkbox"/>	<p>Haslev-Hansen, Nicoline</p> <p><b>How can Transport Act as a Future Driver for Economic Growth and Sustainable Development</b></p> <p>From 01-08-2007 to 31-08-2011</p> <p>Type of end date: expected</p> <p>Project status: Current</p> <p>Type: PhD Project</p>	2007 - 2011	
<input type="checkbox"/>	<p>Lossau, NorbertPeters, DaleNeuroth, HeikeElbæk, Mikael Karstensen; Sandfær, Mogens; Heller, Alfred; Pedersen, Gert Schmeltz</p> <p><b>DRIVER-II : Networking European Scientific Repositories</b></p> <p>From 01-12-2008 to 01-12-2010</p> <p>Type of end date: expected</p> <p>Project status: Current</p> <p>Type: Project</p>	2008 - 2010	
<input type="checkbox"/>	<p>Skouby, Knud ErikHenten, Anders; Olesen, Henning; Falch, Morten; Tadayoni, Reza; Sørensen, Lene; Schultz, Nette; Saugstrup, Dan; Jiang, Bo</p> <p><b>My personal Adaptive Global NET</b></p> <p>From 01-01-2004 to 13-12-2005</p> <p>Type of end date: expected</p> <p>Project status: Finished</p> <p>Type: Project</p>	2004 2004 - 2005	
<input type="checkbox"/>	<p>Nielsen, Otto Anker</p> <p><b>Simulation of truck-traffic to the Harbour of Copenhagen</b></p> <p>From 01-01-1996 to 31-12-1996</p>	1996 - 1996	

Færdig

Start OR09 Indbakke - Microsoft Out... 16 Firefox Microsoft PowerPoint - [...]

DA 12:13

Orbit Home - Mozilla Firefox

http://orbit.dtu.dk/AdvancedSearchEvents,resultSetTable.resultSetTable.displayControl.form.sdirect


Mest besøgte Tgang med Firefox Seneste nyheder

Login

**ORBIT The DTU Research Database**

DTU

Welcome Publications Projects Events Resources Departments Staff Search All myBit



**Events - Advanced Search**

**Search**

- Basic
- Advanced

**General**

- Help
- Links
- About / Coverage
- Contact
- Courses in ORBIT

**External links**

- Riso publications before 2007
- The Danish Research Database

Find: All fields

Person

CWIS no.

Title







Keywords

Include:  DTU records  Department records  Personal records

Limit:

From:  To:

4 records in  format with  rec/page sorted by

<input type="checkbox"/>	 Elbæk, Mikael Karstensen DRIVER-II Summit Type: Conference participation (passive)	2008	
<input type="checkbox"/>	 Martinussen, Heidi Drasbek; Jacobsen, Kirsten Suhr; Damgaard, Helle; Pedersen, Jette Thise DEFF Workshop Conference Management Systems Type: Workshop / Seminar	2009	
<input type="checkbox"/>	 Martinussen, Heidi Drasbek; Jacobsen, Kirsten Suhr; Damgaard, Helle; Pedersen, Jette Thise DEFF Workshop Conference Management Systems Type: Workshop / Seminar	2009	

Færdig

Start OR09 Indbakke - Microsoft Out... 16 Firefox Microsoft PowerPoint - [...]

DA 12:36



Orbit Home - Mozilla Firefox


http://orbit.dtu.dk/Person,contextNavigationMenu.link.sdirect?sp=SProjects

ORBIT The DTU Research Database

DTU

Welcome Publications Projects Events Resources Departments Staff Search All myBit

**Staff - Full Profile**




Help  
Links  
About / Coverage  
Contact

→ Courses in ORBIT

**Mikael Karstensen Elbæk**  
Systembibliotekar

Technical University of Denmark  
Technical Information Center of Denmark



→ Key Data  
→ CV  
→ Publications  
→ Projects  
→ Events  
→ All Records



Anker Engelunds Vej 1  
Building 101D, Office/Room 116A  
2800 Denmark  
Phone: 45257446

Initials: miel  
Cwisno.: 41294  
Phone: 45257446  
Email: [mke@dtic.dtu.dk](mailto:mke@dtic.dtu.dk)

Please note: Also personal and records belonging to departments are included in the search results.

1 records in  format with  rec/page sorted by

All

 Lossau, NorbertPeters, DaleNeuroth, HeikeElbæk, Mikael Karstensen; Sandfær, Mogens; 2008 - 2010   
Heller, Alfred; Pedersen, Gert Schmeltz  
DRIVER-II : Networking European Scientific Repositories  
From 01-12-2008 to 01-12-2010  
Type of end date: expected  
Project status: Current  
Type: Project

All

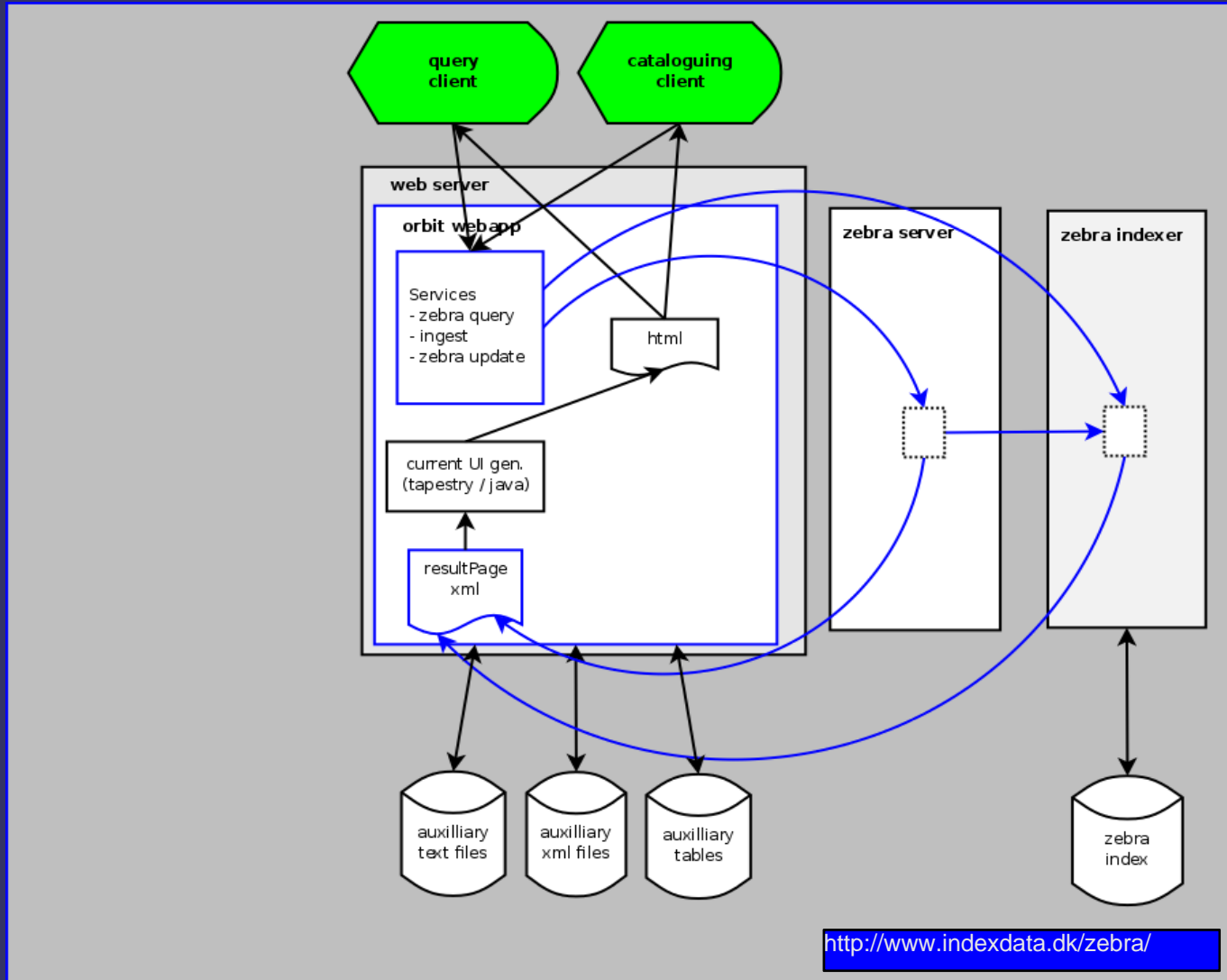
Færdig

Start Indbakke - Microsoft Ou... 16 Firefox OR09 Microsoft PowerPoint - [...] Conference proposal for... DA 15:55

# *Analysis*

- ➔ Current ORBIT Architecture and Data
- ➔ EuroCRIS CERIF Entity-Relationship Model
- ➔ Fedora Content Modelling
- ➔ New ORBIT Architecture and Data

# Current ORBIT Architecture



Publication

Person

Organisation

```
-<record>
+<header><header>
-<metadata>
-<ddf_doc doc_lang="eng" doc_level="pop" doc_review="pr" doc_type="dja" doc_year="2007" format_version="1.2.0" rec_created="2008-0 1-30" rec_id="209647" rec_source="dtu"
rec_status="c" rec_upd="2008-08-25">
-<title>
-<original>
<main>The Library as a mediator for e-publishing</main>
-<sub>
A case on how a library can become a significant factor in facilitating digital scholarly communication and open access publishing for less Web-savvy journals
</sub>
</original>
<title>
-<description>
+<abstract xml:lang="eng"><abstract>
-<subject>
<keyword key_type="fre">Open Access</keyword>
<keyword key_type="fre">E-publishing</keyword>
</subject>
<description>
-<person aff_no="1" pers_role="pau">
-<name>
<first>Mikael Karstensen</first>
<last>Elbæk</last>
</name>
<id id_type="loc_per">DTU-41294</id>
</person>
+<person aff_no="2" pers_role="pau"></person>
-<organisation aff_no="1" org_role="oaf">
-<name>
<level1>Technical University of Denmark</level1>
<level2>Technical Knowledge Center of Denmark</level2>
<level3>ePub</level3>
</name>
</organisation>
+<organisation aff_no="2" org_role="oaf"></organisation>
+<local_field tag_type="4"></local_field>
+<local_field tag_type="4"></local_field>
+<publication></publication>
</ddf_doc>
</metadata>
</record>
```

# Project Person

Mozilla Firefox

Ejer Rediger Vis Historik Bogmærker Funktioner Hjælp

http://orbit.dtu.dk/app?downloadOptionName=XML&service=download\_records

Mest besøgte Igang med Firefox Seneste nyheder

```
<ddfs>
- <ddf id="224274" status="O3" type="pr" user_group="dtu_dtv" user_id="41294" version="2.0">
+ <project authority_object="ddf_object_type" authority_object_dan="Main" authority_object_eng="Main" object="main"></project>
+ <derived></derived>
- <indexing>
+ <instituteowner></instituteowner>
<crossdepartment/>
- <persons>
- <projectleader>
- <name>
<first>Norbert</first>
<full>Lossau, Norbert</full>
<last>Lossau</last>
<name>
<cwisno/>
</projectleader>
+ <projectleader></projectleader>
+ <projectstaff></projectstaff>
- <projectstaff>
- <name>
<first>Mikael Karstensen</first>
<full>Elbæk, Mikael Karstensen</full>
<last>Elbæk</last>
<name>
<cwisno>41294</cwisno>
</projectstaff>
+ <contactperson></contactperson>
+ <contactperson></contactperson>
+ <projectstaff></projectstaff>
</persons>
<cwisno>41294</cwisno>
<cwisno>2225</cwisno>
<cwisno>2685</cwisno>
<cwisno>25085</cwisno>
<pyear>2008</pyear>
<pyear>2009</pyear>
<pyear>2010</pyear>
<prstart>2008</prstart>
<prend>2010</prend>
+ <sorttitle></sorttitle>
<ownerclass>dtu</ownerclass>
<released>2008-10-08T10:00:52-T00:00:00</released>
</indexing>
</ddf>
</ddfs>
```

Færdig

Start OR09 Indbakke - Microsoft Out... 16 Firefox Microsoft PowerPoint - [...]

DA 12:30

Event

Person

```
<ddfs>
- <ddf id="209641" status="O3" type="ev" user_group="dtu_dtv" user_id="41294" version="2.0">
- <event authority_object="ddf_object_type" authority_object_dan="Main" authority_object_eng="Main" authority_type="dtu_event_type_main" authority_type_dan="Konferencedeltagelse (passiv)" authority_type_eng="Conference participation (passive)" object="main" type="conf_particip">
+ <abstract authority_lang="language" authority_lang_dan="Engelsk" authority_lang_eng="English" lang="eng"></abstract>
+ <dates></dates>
<keyword authority_lang="language" authority_lang_dan="Engelsk" authority_lang_eng="English" lang="eng" type="free">Open Access</keyword>
<keyword authority_lang="language" authority_lang_dan="Engelsk" authority_lang_eng="English" lang="eng" type="free">Institutional Repositories</keyword>
<keyword authority_lang="language" authority_lang_dan="Engelsk" authority_lang_eng="English" lang="eng" type="free">Open Access Repositories</keyword>
+ <local></local>
+ <local></local>
+ <local></local>
+ <local></local>
- <name>
<full>DRIVER-II Summit, </full>
<main>DRIVER-II Summit</main>
<name>
<object authority_role="fit_object_role" authority_role_dan="Fulltext" authority_role_eng="Fulltext" filterempty_keepme="1" id="1" object="1" role="fulltext"/>
- <person id="199291" object="aux">
- <contact>
<email>mke@dtv.dk</email>
</contact>
<identifier type="cwisno">41294</identifier>
- <name>
<first>Mikael Karstensen</first>
<full>Elbæk, Mikael Karstensen</full>
<last>Elbæk</last>
</name>
+ <organisation id="199291" object="aux" role="affiliation"></organisation>
<person>
<place>Göttingen, Germany</place>
- <www>
- <text>
http://www.driver-support.eu/multi/DRIVERSummit.php
</text>
</www>
</event>
+ <derived></derived>
+ <indexing></indexing>
</ddf>
</ddfs>
```

Person

Organisation

The screenshot shows a Mozilla Firefox browser window with the address bar containing the URL `http://orbit.dtu.dk/app?downloadOptionName=XML&service=download_records`. The page content is XML data, which is partially highlighted in red. The XML structure is as follows:

```
<ddfs>
  - <ddf id="199291" status="O4" type="pe" user_group="dtu_dtv" user_id="41294" version="2.0">
    - <person authority_object="ddf_object_type" authority_object_dan="Main" authority_object_eng="Main" object="main">
      + <contact></contact>
      + <imageurl></imageurl>
      <keyword authority_lang="language" authority_lang_dan="Engelsk" authority_lang_eng="English" lang="eng" type="free">Institutional repository</keyword>
      <keyword type="free"/>
      - <name>
        <first>Mikael Karstensen</first>
        <full>Elbæk, Mikael Karstensen</full>
        <last>Elbæk</last>
        <name>
      + <organisation employmentid="48633" id="7405" object="aux" position_title="Systembibliotekar" role="affiliation" status="previous" type="cu"></organisation>
      + <organisation employmentid="49292" id="7458" object="aux" position_title="Systembibliotekar" preferred="0" role="affiliation" status="previous" type="cu"></organisation>
      - <organisation employmentid="49292" id="7458" object="aux" position_title="Systembibliotekar" preferred="1" role="affiliation" status="current" type="cu">
        + <contact></contact>
        + <name lang="dan"></name>
        - <name lang="eng">
          - <full>
            Technical University of Denmark, Technical Information Center of Denmark,
          </full>
          <main code="DTU">Technical University of Denmark</main>
          <sub code="DTV">Technical Information Center of Denmark</sub>
          <sub2 code="DTV_none"> </sub2>
          <name>
        </organisation>
      - <profile>
        - <popular>
          - <orbit authority_lang="language" authority_lang_dan="Dansk" authority_lang_eng="Danish" lang="dan">
            Institutional Repository, e-publicering og open access <n> Projektledelse, service management
          <br>
          </orbit>
          <popular>
        </profile>
      + <FIXME></FIXME>
      <identifier type="dtu_initials">miel</identifier>
      <identifier type="cwisno">41294</identifier>
    </person>
    + <derived></derived>
    + <indexing></indexing>
  </ddf>
</ddfs>
```

The browser's taskbar at the bottom shows several open applications: Start, Indbakke - Microsoft Out..., OR09, d2fcris-gsp-at-or09.ppt, fedora-orbit-gsp-at-or09..., and 14 Firefox. The system tray on the right shows the date and time as 11:09.

# *Current ORBIT*

Basic problems:

- ➔ Lack of abstraction level for the data
- ➔ Lack of flexibility towards new requirements / new user services

# New ORBIT

Basic ideas:

- ⇒ Re-implement ORBIT, with Fedora as backend
- ⇒ Provide abstract data model, based on CERIF
  - It is extensible, e.g. for support of full research data life cycle
- ⇒ Realize with Fedora Content Models and RDF
- ⇒ Re-use the current user interface
  - but add re-usable components for extensions of data model

Thus we can have

- ⇒ eXtensible ORBIT ~~~~ *XORBIT*

# euroCRIS CERIF – Common European Research Information Format

**Introduction: euroCRIS - Mozilla Firefox**

File Rediger Vis Historik Bogmærker Funktioner Hjælp

http://www.eurocris.org/cerif/introduction/

ec CERIF 2006: eur... http://w...ORE..xsd Toad Data Model... ec Introducti... Toad Data Model... Toad Data Model... Toad Data Model...

**euroCRIS**  
Current Research Information Systems

**PUBLIC**

- Home
- News
- About euroCRIS
- Organisation
- Charter
- Brochure
- Partners
- Events
- Conferences
- CRIS 2008
- Seminars
- Membership meetings
- Courses
- Join euroCRIS
- Contact us
- Disclaimer and Terms of use

**CERIF**

- Introduction**
- Tutorial
- CERIF Releases
- CERIF 2004
- CERIF 2006
- CERIF 2008

**MEMBERS**

- Members area

**Introduction**

CERIF (Common European Research Information Format) was developed with the support of the EC (European Commission) in two major phases: 1987-1990 and 1997-1999. It is a standard; technically it is an EU (European Union) Recommendation to member states. Since 2002 care and custody of CERIF has been handed by the EC to euroCRIS (www.eurocris.org) a not-for-profit organisation dedicated to the promotion of CRIS (Current Research Information Systems).

Most nation-states have publicly-supported research programmes. It is realised that public sponsorship of research and development leads to wealth creation and improvement in the quality of life. Because public funding is involved, it is necessary for there to be appropriate governance, and also for the information to be available to the public. Broadly, each nation state has a similar research process of: strategic planning; programme announcement; call for proposals; proposal evaluation and awarding; project result monitoring, project result exploitation.

However, research is international. A research project in country A is likely to be based on previous research in several other countries. Many research projects are now transnational: well-known examples include the human genome and climate change but there are many others, especially where expensive infrastructure is utilised such as particle physics or space science. Furthermore, knowledge of the research activity in country A may influence the strategy towards research - including priorities and resources provided - in country B. Thus there is a need to share research information across countries, or even between different funding agencies in the same country.

**Background:**

The information is used by researchers (to find partners, to track competitors, to form collaborations); research managers (to assess performance and research outputs and to find reviewers for research proposals); research strategists (to decide on priorities and resourcing compared with other countries); publication editors (to find reviewers and potential authors); intermediaries / brokers (to find research products and ideas that can be carried forward with knowledge/ technology transfer to wealth creation); the media (to communicate the results of R&D in a socio-economic context) and the general public (for interest).

Within Europe this requirement has long been recognised. In the early eighties Heads of some national research funding organisations initiated a project named IDEAS to investigate linking databases of research information. The follow on project EXIRPTS 1987-1989 extended this - based on the G7 countries - to include USA and Japan. Both projects successfully demonstrated: (a) resolution of schema differences between heterogeneous distributed databases (b) translation and distribution of queries as subqueries to the different target systems (c) integration of the results of the distributed subqueries and presentation back to the queryer. The technology utilised a specially-developed protocol, run over email and file transfer. It

Færdig

Start Indbakke - Microsoft Out... 14 Firefox M:\My Documents\DJIMA... Microsoft PowerPoint - [...]

10:22

## 2. The CERIF 2008 – 1.0 Model

To reduce the complexity of the model towards a better understanding, this introduction and specification document follows a conceptual structure. The conceptual structure allows for different perspectives and views when talking about parts of the model and enables the emphasis to particular model features. This conceptual structure is only a virtual structure and as such not inherent in the physical data model, and therefore, also not incorporated in the physical SQL scripts. It is used for organizing this document and considered an instrument that supports the comprehension of the CERIF model.

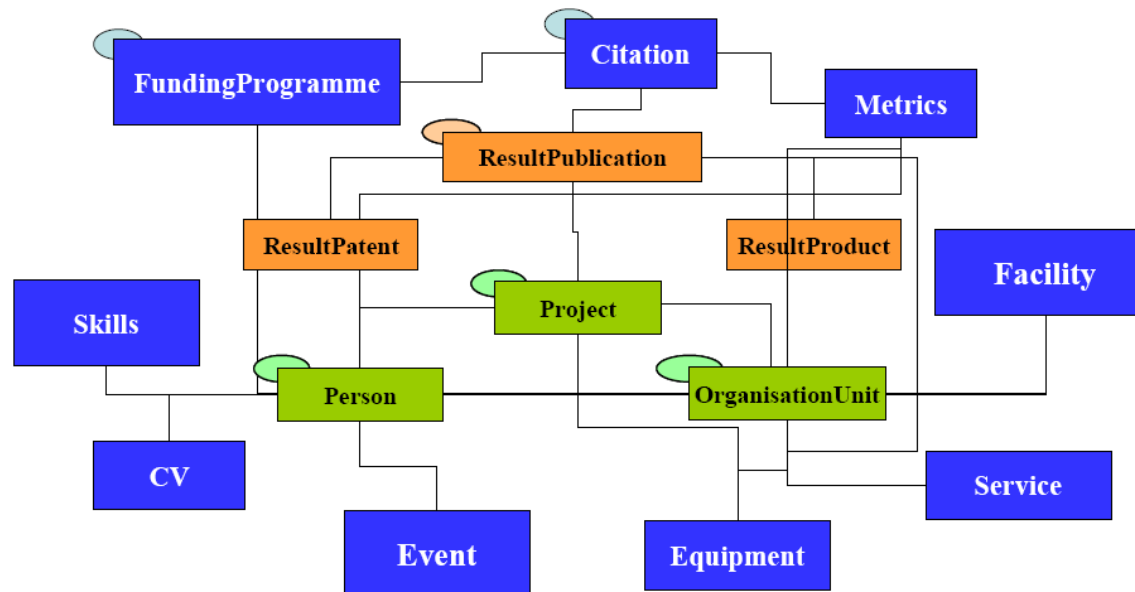


Figure 1: Some CERIF Entities and their Relationships

### 2.1 CERIF Conceptual Structure



CERIF2008\_1.0\_Semantics.pdf (application/pdf Objekt) - Mozilla Firefox

http://www.eurocris.org/fileadmin/cerif-2008/CERIF2008\_1.0\_Semantics.pdf

3.1 CERIF Publication Types

cfPublicationTypes	
	Book
	Book Review
	Book Chapter Abstract
	Book Chapter Review
	Inbook
	Anthology
	Monograph
	Referencebook
	Textbook
	Encyclopedia
	Manual
	Otherbook
	Journal
	Journal Article
	Journal Article Abstract
	Journal Article Review
	Conference Proceedings
	Conference Proceedings Article
	Letter
	Letter to Editor
	PhD Thesis
	Doctoral Thesis
	Report
	Short Communication
	Poster
	Presentation
	Newsclipping
	Commentary
	Annotation

3.2 CERIF Person-Publication Roles

cfPerson-ResultPublicationRoles	
	is author of

4 of 7

Done

Start | Indbakke - Microsoft Out... | OR09 | d2fcris-gsp-at-or09.ppt | fedora-orbit-gsp-at-or09... | 15 Firefox | DA | 12:06

CERIF2008\_1.0\_Semantics.pdf (application/pdf Objekt) - Mozilla Firefox

http://www.eurocris.org/fileadmin/cerif-2008/CERIF2008\_1.0\_Semantics.pdf

Save a Copy | Search | Select | 125% | Sign

**cfPerson-ResultPublicationRoles**

	is author of
	is author (numbered) of
	is author (percentage) of
	is editor (numbered) of
	is editor of
	is subject of
	is reviewer of
	is translator of
	is publisher of
	has commissioned

CERIF 2008 – 1.0 Semantics Page 5

3.3 CERIF OrgUnit-Publication Roles

**cfOrgUnit-ResultPublicationRoles**

	claims IPR of
	is publisher of
	is curator of
	provides reviewer for
	is author of
	has commissioned
	is funded by
	is author institution of
	is publishing institution of
	is external organisation of

5 of 7

Done

Start | Indbakke - Microsoft Out... | OR09 | d2fcris-gsp-at-or09.ppt | fedora-orbit-gsp-at-or09... | 15 Firefox | DA | 12:05



# XML Data Exchange for Interoperability

CERIF2008\_1.0\_XML.pdf (application/pdf Objekt) - Mozilla Firefox

http://www.eurocris.org/fileadmin/cerif-2008/CERIF2008\_1.0\_XML.pdf

```
</cfOrgUnit>
<cfOrgUnit>
  <cfOrgUnitId>0432125</cfOrgUnitId>
  <cfURI>http://www.dfki.de/lt/</cfURI>
  <cfAcronym>LT Lab</cfAcronym>
  <cfHeadCount>40</cfHeadCount>
</cfOrgUnit>
...
</CERIF>
```

### 7.1.2 CERIF Result XML Entities (XML Examples)

```
<?xml version="1.0" encoding="UTF-8"?>
<CERIF
  xsi:schemaLocation="http://www.eurocris.org/cerif/cfResPubl-CORE cfResPubl-CORE.xsd"
  xmlns="http://www.eurocris.org/cerif/cfResPubl-CORE"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  release="2006" date="YYYY-MM-DD" sourceDatabase="euroCRIS">
  <cfResPubl>
    <cfResPublId>publication-joerg-et-al</cfResPublId>
    <cfURI>http://www.eurocris.org/fileadmin/Upload/Events/
      Conferences/CRIS2008/Papers/cris2008_Joerg.pdf </cfURI>
    <cfResPublDate>2008</cfResPublDate>
    <cfStartPage>107</cfStartPage>
    <cfEndPage>123</cfEndPage>
    <cfISBN>978-961-6133-38-8</cfISBN>
  </cfResPubl>
  <cfResPubl>
    <cfResPublId>publication-veca-c-storey</cfResPublId>
    <cfURI>http://www.springerlink.com/content/j23263j02m850617/</cfURI>
    <cfResPublDate>1993</cfResPublDate>
    <cfNum>4</cfNum>
```

Done

Start Indbakke - Microsoft Out... OR09 d2fcris-gsp-at-or09.ppt fedora-orbit-gsp-at-or09... 15 Firefox DA 12:36

Alternative:  
CERIF realized  
in relational  
database tables

Toad Data Modeler - HTML Report - Mozilla Firefox

http://www.dfki.de/~brigitte/CERIF/CERIF2008\_1.0FDM/Physical/default.html

CERIF 2008: euroCRIS

## Physical model - tables

www.quest.com

model info ER diagram tables columns indexes alternative keys relationships data dictionary

cfOrgUnit\_Event  
cfOrgUnit\_ExpSkills  
cfOrgUnit\_Facil  
cfOrgUnit\_FundProg  
cfOrgUnit\_OrgUnit  
cfOrgUnit\_PAddr  
cfOrgUnit\_Prize  
cfOrgUnit\_ResPat  
cfOrgUnit\_ResProd  
cfOrgUnit\_ResPubl  
cfOrgUnit\_Srv  
cfOrgUnitKeyw  
cfOrgUnitName  
cfOrgUnitResAct  
cfPAddr  
**cfPAddr\_Class**  
cfPers  
cfPers\_Class  
cfPers\_Country  
cfPers\_CV  
cfPers\_DC  
cfPers\_EAddr  
cfPers\_Equip  
cfPers\_Event  
cfPers\_ExpSkills  
cfPers\_Facil  
cfPers\_FundProg  
cfPers\_Lang  
cfPers\_OrgUnit  
cfPers\_PAddr  
cfPers\_Pers  
cfPers\_Prize  
cfPers\_Qual  
cfPers\_ResPat  
cfPers\_ResProd  
cfPers\_ResPubl  
cfPers\_Srv  
cfPersKeyw  
cfPersName  
cfPersName\_Pers  
cfPersResInt

### cfPers

#### Columns

Key	Column name	Domain	Data type	Not null	Unique	Check	Default	Comments
PK	cfPersId	ID	Char (32)	YES	NO			person unique identifier (cfPersId) that propagates to cfPersonName cfPersonName_Person cfPersonResearchInterest cfPersonKeywords cfPerson_Person cfPerson_ExpertiseAndSkills cfPerson_Qualification cfPerson_Event cfPerson_Facility cfPerson_PrizeAward cfPerson_ResultProduct cfPerson_ResultPublication cfPerson_PostAddress cfPerson_FundingProgramme cfPerson_Language cfPerson_OrganisationUnit cfPerson_Classification cfPerson_Service cfPerson_Country cfPerson_Equipment cfPerson_ElectronicAddress cfPerson_CurriculumVitae cfProject_Person
	cfSex		Char (1)	NO	NO	u, m, f	u	The gender of a person f=female m=male u=unknown (default)
	cfURI		Char (128)	NO	NO			

#### Relationships

Constraint (relationship) name	Relationship type	Parent table	Child table	Card.
cfPers-PersKeyw (cfPers-PersKeyw)	Identifying	cfPers	cfPersKeyw	1:N
cfPers-PersResInt (cfPers-PersResInt)	Identifying	cfPers	cfPersResInt	1:N
cfPers-Pers_Class (cfPers-Pers_Class)	Identifying	cfPers	cfPers_Class	1:N
cfPers-Pers_Pers2 (cfPers-Pers_Pers2)	Identifying	cfPers	cfPers_Pers	1:N
cfPers-Pers_Pers1 (cfPers-Pers_Pers1)	Identifying	cfPers	cfPers_Pers	1:N
cfPers-Pers_OrgUnit (cfPers-Pers_OrgUnit)	Identifying	cfPers	cfPers_OrgUnit	1:N
cfPers-Pers_ResPubl (cfPers-Pers_ResPubl)	Identifying	cfPers	cfPers_ResPubl	1:N
cfPers-Pers_EAddr (cfPers-Pers_EAddr)	Identifying	cfPers	cfPers_EAddr	1:N

http://www.dfki.de/~brigitte/CERIF/CERIF2008\_1.0FDM/Physical/TablesB.html#Table\_147

Start Indbakke - Microsoft Out... OR09 d2fcris-gsp-at-or09.ppt fedora-orbit-gsp-at-or09... 15 Firefox DA 12:58

Selected alternative: Realize with Fedora Content Model Architecture

tutorial1.pdf (application/pdf Objekt) - Mozilla Firefox

http://www.fedora-commons.org/documentation/3.0/userdocs/tutorials/tutorial1.pdf

```
graph TD; Data[Data] -- fedora-model:hasModel --> ContentModel[Content Model]; ContentModel -- fedora-model:hasService --> ServiceDefinition[Service Definition]; ServiceDefinition -- fedora-model:isDeploymentOf --> ServiceDeployment[Service Deployment]; ServiceDeployment -- fedora-model:isContractorOf --> ContentModel;
```

**Figure 5: Fundamental CMA Relationships**

Figure 5 illustrates the required relationships between fundamental object types in the CMA. In the CMA object serialization these relations are asserted as RDF statements in the digital objects' RELS-EXT Datastream. These relations are asserted only in the object at the origin of each arrow though typically these relations will be harvested and indexed within utilities such as Semantic Triplestores or relational databases to enable fast query over them, or into caches which permit rapid access to their functionality.

The "hasModel" relation identifies the class of the Data object. There may or may not be a Fedora digital object that corresponds to the identifier. If the identifier refers to an object it must be a CModel object and contain the base content model document. It is expected that many Data objects conform to a single Content Model (and have a relation asserted to the same CModel object). The Content Model characterizes the Data objects that conform to it.

27 of 29

Start | Indbakke - Microsoft Out... | 14 Firefox | M:\My Documents\DIM... | Microsoft PowerPoint - [...]

10:09

tutorial1.pdf (application/pdf Objekt) - Mozilla Firefox

http://fedora-commons.org/confluence/download/attachments/4718930/tutorial1.pdf?version=1

http://www.fedora.i...sext-ontology.rdfs

Save a Copy Search Select 150%

scholarly objects

commercial web content

museum and library objects

Figure 4: A Network of Digital Objects

Digital object relationship metadata is a way of asserting these various kinds of relationships for Fedora objects. A default set of common relationships is defined in the [Fedora relationship ontology](#) (actually, a simple RDF schema) which defines a set of common generic relationships useful in creating digital object networks. These relationships can be refined or extended. Also, communities can define their own ontologies to encode relationships among Fedora digital objects. Relationships are asserted from the perspective of one object to another object as in the following general

17 of 29

Start OR09 Indbakke - Microsoft Out... 14 Firefox Microsoft PowerPoint - [...]

Mozilla Firefox

Eller Rediger Vis Historik Bogmærker Funktioner Hjælp

http://www.fedora.info/definitions/1.0/fedora-relsext-ontology.rdfs

Mest besøgte Igang med Firefox Seneste nyheder

tutorial1.pdf (application/pdf Objekt) http://www.fedor...t-ontology.rdfs

```
- <rdf:RDF base="info:fedora/fedora-system:def/relations-external#">
+ <rdf:Property rdf:ID="fedoraRelationship"></rdf:Property>
- <rdf:Property rdf:ID="isPartOf">
  <rdfs:label xml:lang="en-US">Is Part Of</rdfs:label>
  - <rdfs:comment xml:lang="en-US">
    A definition of the generic part/whole relationship between fedora objects. The subject is a fedora object representing a part and the predicate
    is a fedora object representing a whole.
  </rdfs:comment>
  <rdfs:subPropertyOf rdf:resource="#fedoraRelationship"/>
</rdf:Property>
+ <rdf:Property rdf:ID="hasPart"></rdf:Property>
+ <rdf:Property rdf:ID="isConstituentOf"></rdf:Property>
+ <rdf:Property rdf:ID="hasConstituent"></rdf:Property>
+ <rdf:Property rdf:ID="isMemberOf"></rdf:Property>
+ <rdf:Property rdf:ID="hasMember"></rdf:Property>
+ <rdf:Property rdf:ID="isSubsetOf"></rdf:Property>
+ <rdf:Property rdf:ID="hasSubset"></rdf:Property>
+ <rdf:Property rdf:ID="isMemberOfCollection"></rdf:Property>
+ <rdf:Property rdf:ID="hasCollectionMember"></rdf:Property>
+ <rdf:Property rdf:ID="isDerivationOf"></rdf:Property>
+ <rdf:Property rdf:ID="hasDerivation"></rdf:Property>
+ <rdf:Property rdf:ID="isDependentOf"></rdf:Property>
+ <rdf:Property rdf:ID="hasDependent"></rdf:Property>
+ <rdf:Property rdf:ID="isDescriptionOf"></rdf:Property>
+ <rdf:Property rdf:ID="HasDescription"></rdf:Property>
+ <rdf:Property rdf:ID="isMetadataFor"></rdf:Property>
+ <rdf:Property rdf:ID="HasMetadata"></rdf:Property>
+ <rdf:Property rdf:ID="isAnnotationOf"></rdf:Property>
+ <rdf:Property rdf:ID="HasAnnotation"></rdf:Property>
+ <rdf:Property rdf:ID="hasEquivalent"></rdf:Property>
</rdf:RDF>
```

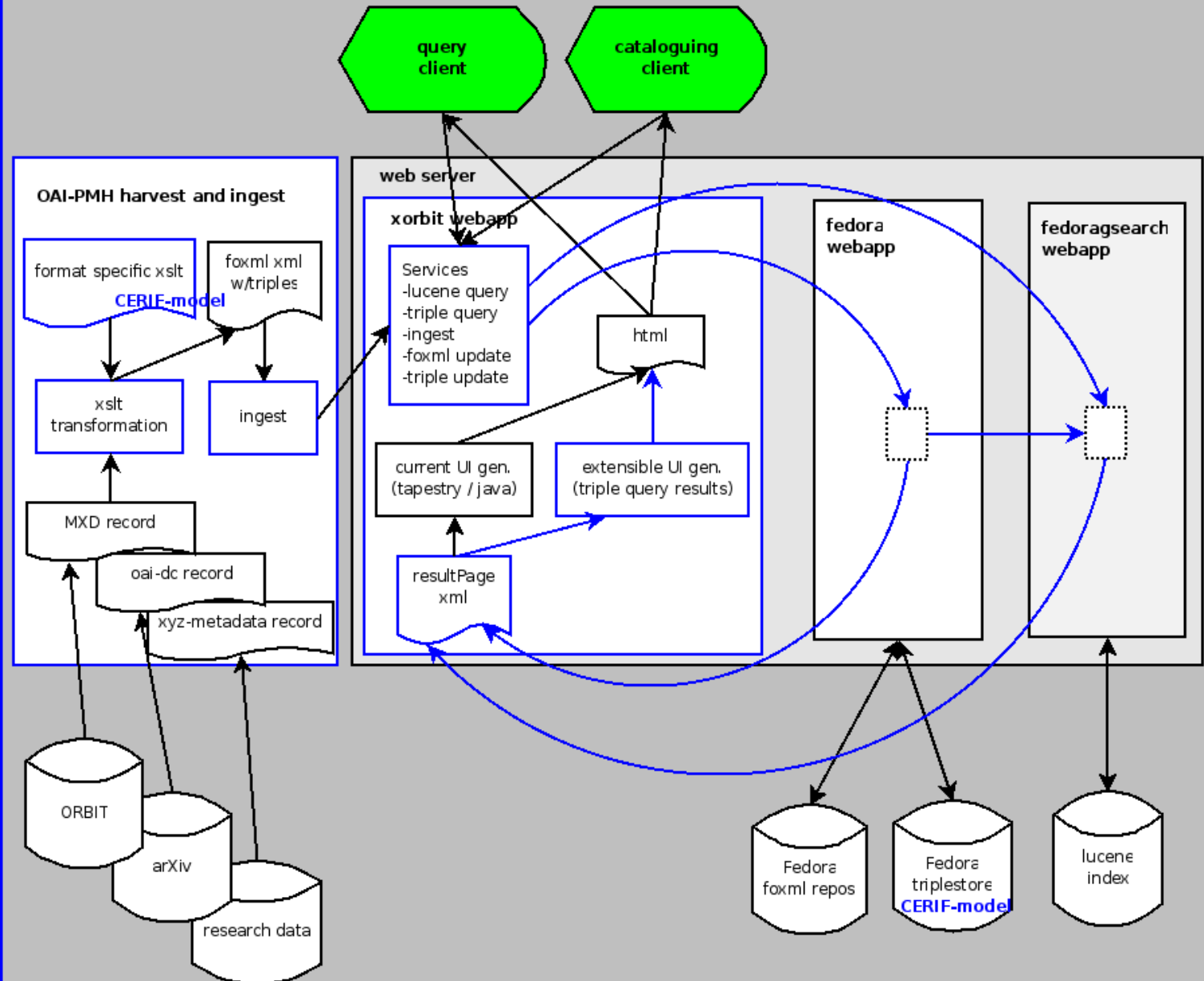
Færdig

Start OR09 Indbakke - Microsoft Out... 14 Firefox Microsoft PowerPoint - [...]

DA 14:15

The architecture of the Fedora-based ORBIT implementation:

- The Fedora web application with its repository of *foxml* objects and its *triplestore*, supplemented by the *fedoragsearch* web application with its *lucene* index.
- OAI-PMH harvest and ingest, which wraps all sorts of metadata into *foxml* objects, including the generation of triples in the *RELS-EXT* datastream, representing entities and relationships of the CERIF-model.
- The *xorbit* web application, which services query clients and cataloguing clients.



# Transform ORBIT-XML to FOXML

```
!MXD!
</foxml:xmlContent>
</foxml:datastreamVersion>
</foxml:datastream>

<foxml:datastream ID="RELS-EXT" STATE="A" CONTROL_GROUP="X" VERSIONABLE="true">
  <foxml:datastreamVersion ID="RELS-EXT.0" LABEL="RDF Statements about this object" MIMETYPE="application/rdf+xml">
    <foxml:xmlContent>
      <rdf:RDF xmlns:fedora-model="info:fedora/fedora-system:def/model#"
        xmlns:cf2008-model="info:cf2008-model/fedora-system:def/model#"
        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:rel="info:fedora/fedora-system:def/relations-external#">
        <rdf:Description>
          <xsl:attribute name="rdf:about">info:fedora/<xsl:value-of select="$PIDVALID"/></xsl:attribute>
          <fedora-model:hasModel rdf:resource="info:fedora/cf2008:cfResultPublication"></fedora-model:hasModel>
          <cf2008-model:cfHasMainOriginalTitle>
            <xsl:value-of select="normalize-space(//mxd:title/mxd:original/mxd:main)"/>
          </cf2008-model:cfHasMainOriginalTitle>
          <xsl:for-each select="//mxd:person[@pers_role='pau']">
            <cf2008-model:cfHasAuthor>
              <xsl:attribute name="rdf:resource">info:fedora/cfPerson:<xsl:value-of select="normalize-
            </cf2008-model:cfHasAuthor>
          </xsl:for-each>
          <xsl:for-each select="//mxd:organisation[@org_role='oau']">
            <cf2008-model:cfHasOrg>
              <xsl:attribute name="rdf:resource">info:fedora/cfOrganisation:<xsl:value-of select="tran
            </cf2008-model:cfHasOrg>
          </xsl:for-each>
        </rdf:Description>
      </rdf:RDF>
    </foxml:xmlContent>
  </foxml:datastreamVersion>
</foxml:datastream>

<xsl:if test="$PDFREF">
  <foxml:datastream ID="PDF" CONTROL_GROUP="E" STATE="A" VERSIONABLE="true">
    <foxml:datastreamVersion ID="PDF1.0" LABEL="PDF full text" MIMETYPE="application/pdf" >
      <foxml:contentLocation TYPE="URL">
        <xsl:attribute name="REF"><xsl:value-of select="$PDFREF"/></xsl:attribute>

```



File Edit View Search Tools Documents Help



New



Open



Save



Print...



Undo



Redo



Cut



Copy



Paste

FOXML

\*oai.ddf.deff.dk\_165948

```

1180 </rdf:Description>
1181 </rdf:RDF>
1182 </foxml:xmlContent>
1183 </foxml:datastreamVersion>
1184 <foxml:datastreamVersion ID="RELS-EXT.4" LABEL="" CREATED="2008-12-17T12:37:27.876Z" MIMETYPE="text/xml">
1185 <foxml:xmlContent>
1186 <rdf:RDF xmlns:cf2008-model="info:cf2008-model/fedora-system:def/model#"
1187   xmlns:fedora-model="info:fedora/fedora-system:def/model#" xmlns:rdf="http://www.w3.org/1999/02/22-
system:def/relations-external#" >
1188   <rdf:Description rdf:about="info:fedora/oai.ddf.deff.dk:165948">
1189     <fedora-model:hasModel rdf:resource="info:fedora/cf2008:cfResultPublication"/>
1190     <cf2008-model:cfHasMainOriginalTitle>Positivity of the spherically averaged atomic one-electron
1191     <cf2008-model:cfHasAuthor rdf:resource="info:fedora/cfPerson:Fournais"/>
1192     <cf2008-model:cfHasAuthor rdf:resource="info:fedora/cfPerson:Hoffmann-Ostenhof"/>
1193     <cf2008-model:cfHasAuthor rdf:resource="info:fedora/cfPerson:Sørensen"/>
1194     <cf2008-model:cfHasOrg rdf:resource="info:fedora/cfOrganisation:Videnbasen_for_Aalborg_University"/>
1195     <cf2008-model:cfHasOrg rdf:resource="info:fedora/cfOrganisation:CNRS_and_Laboratoire_de_Mathématiques"/>
1196     <cf2008-model:cfHasOrg rdf:resource="info:fedora/cfOrganisation:Universität_Wien"/>
1197     <cf2008-model:cfHasOrg rdf:resource="info:fedora/
cfOrganisation:Universität_Wien_and_Erwin_Schrödinger_International_Institute_for_Mathematical_Physics"/>
1198   </rdf:Description>
1199 </rdf:RDF>
1200 </foxml:xmlContent>
1201 </foxml:datastreamVersion>
1202 </foxml:datastream>

```

## Triple query

```
select $pub $title from <#ri> where
  $pub
  <info:fedora/fedora-system:def/hasModel>
  <info:fedora/cf2008:cfResultPublication>
and
  $pub
  <info:fedora/cf2008:def/cfMainOriginalTitle>
  $title
and
  $pub
  <info:fedora/cf2008/cfPresentedAt>
  $event
and
  $event
  <info:fedora/cf2008/cfYear>
  "2008"
```

## Result

```
<sparql>
  <head>
    <variable name="pub"/>
    <variable name="title"/>
  </head>
  <results>
    <result>
      <pub uri="info:fedora/oai.ddf.deff.dk:169697"/>
      <title>Biomass Co-Firing in Suspension-Fired Power Plants</title>
    </result>
    <result>
      <pub uri="info:fedora/oai.ddf.deff.dk:165948"/>
      <title>Positivity of the spherically averaged atomic one-electron
        density</title>
    </result>
  </results>
</sparql>
```

# *Summing up*

- The backbone of eXtensible ORBIT is the abstract data model:
  - It serves as a high level abstraction
  - It determines the set of relationships used in the triples
  - It is used in the format specific xslts that transform metadata records into triples
  - The model may evolve to accommodate new requirements
  - Then the format specific xslts are edited to generate triples that represent the new model

# *Perspectives*

- Examples of new services that are facilitated by the new implementation are:
  - flexible and extended tools for reporting, making customized reports possible
  - interlinking between research entities like projects and publications, thus:
    - enhancing information retrieval providing several entries to research, i.e. from project to publications to person to related events etc
    - enabling author tools that makes creating relations between research entities easy

# Conclusion

- Abstract data models to play significant role for applications of Fedora in general?

*YES, we think so*

- Towards eXtensible applications of Fedora in general?

*YES, we think so*

- *What do \_you\_ think?*