ORE Experiments presentation

Thomas Place

Tilburg University (NL)
SURFshare (NL) - Driver II (EU)
NEEO (EU)
ORE LG (W)

TILBURG UNIVERSITY

Agenda

- Possible ORE applications based on the Dutch repository infrastructure
- Enhancing publications with data sets
- Detailed relationships between publications and data sets

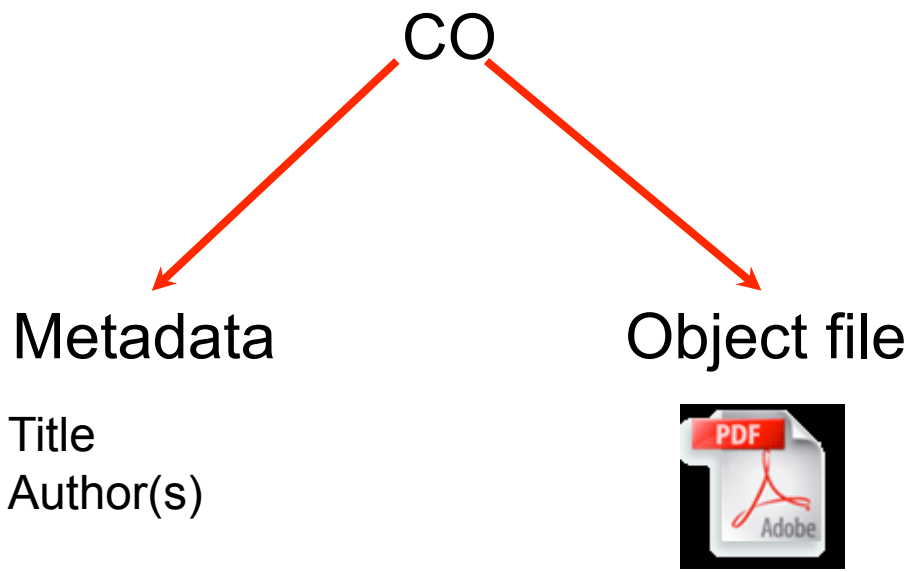
TILBURG UNIVERSITY

Dutch national initiative: DAREnet

- DAREnet is a search service which gives free access to full text publications and other research output in the Netherlands.
- It harvests 21 repositories
- Result of the DARE (Digital Academic REpositories) programme (2003-2006); a national initiative coordinated by the SURF Foundation.
- The follow-up programme is SURFshare (2007-2010); one of the themes is enhanced publications
- The Narcis portal will replace DAREnet

Institute	**DAREnet**	Cream of Science	**Promise of Science**
Amsterdam University Press	64	-	-
Delft University of Technology	8777	1669	3561
Eindhoven University of Technology	15367	3862	2832
Erasmus University Rotterdam	5563	4515	806
International Institute for the Study of Islam in the Modern World	79	-	3
KNAW	1483	3823	-
Leiden University	8571	-	497
Maastricht University	6802	2942	1852
NWO	7247	1601	-
National Institute for Public Health and the Environment	2200	-	-
Naturalis	986	-	-
Open Universiteit Nederland	1003	220	21
Radboud University Nijmegen	7085	3631	1150
Tilburg University	7044	5541	152
University for Humanistics	18	-	1
University of Amsterdam	10321	2581	2170
University of Groningen	19985	1364	1864
University of Twente	6207	4604	947
Utrecht University	20186	3898	1756
Vrije Universiteit Amsterdam	7030	3230	442
Wageningen University & Research Centre	11832	3606	1043
Total	**147850**	47087	**19097**

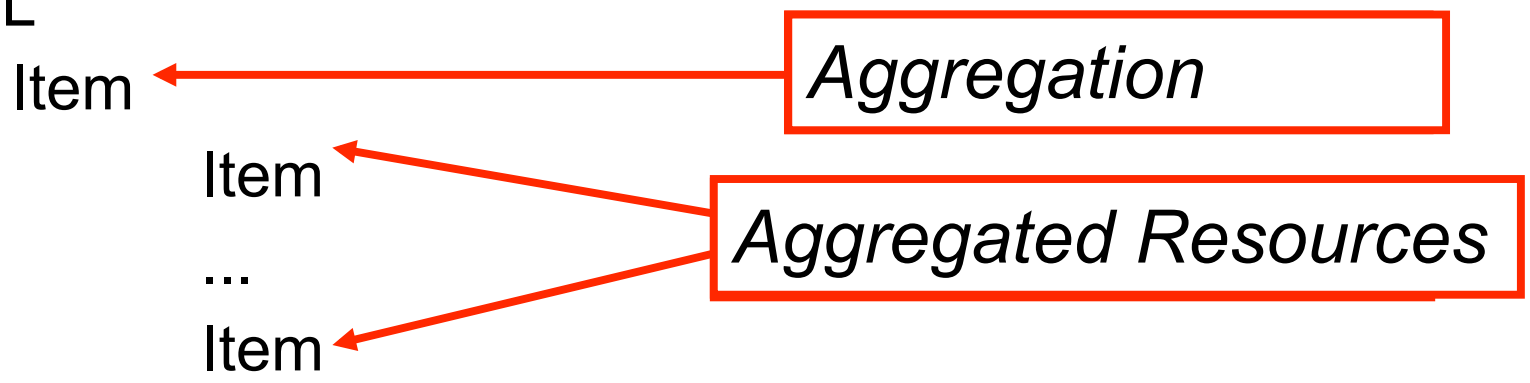
Simple (compound) object model



TILBURG UNIVERSITY

DIDL is used for representing Compound Objects

- DIDL



- Item hierarchy not used; COs graphs not trees

TILBURG UNIVERSITY

Types of objects in CO

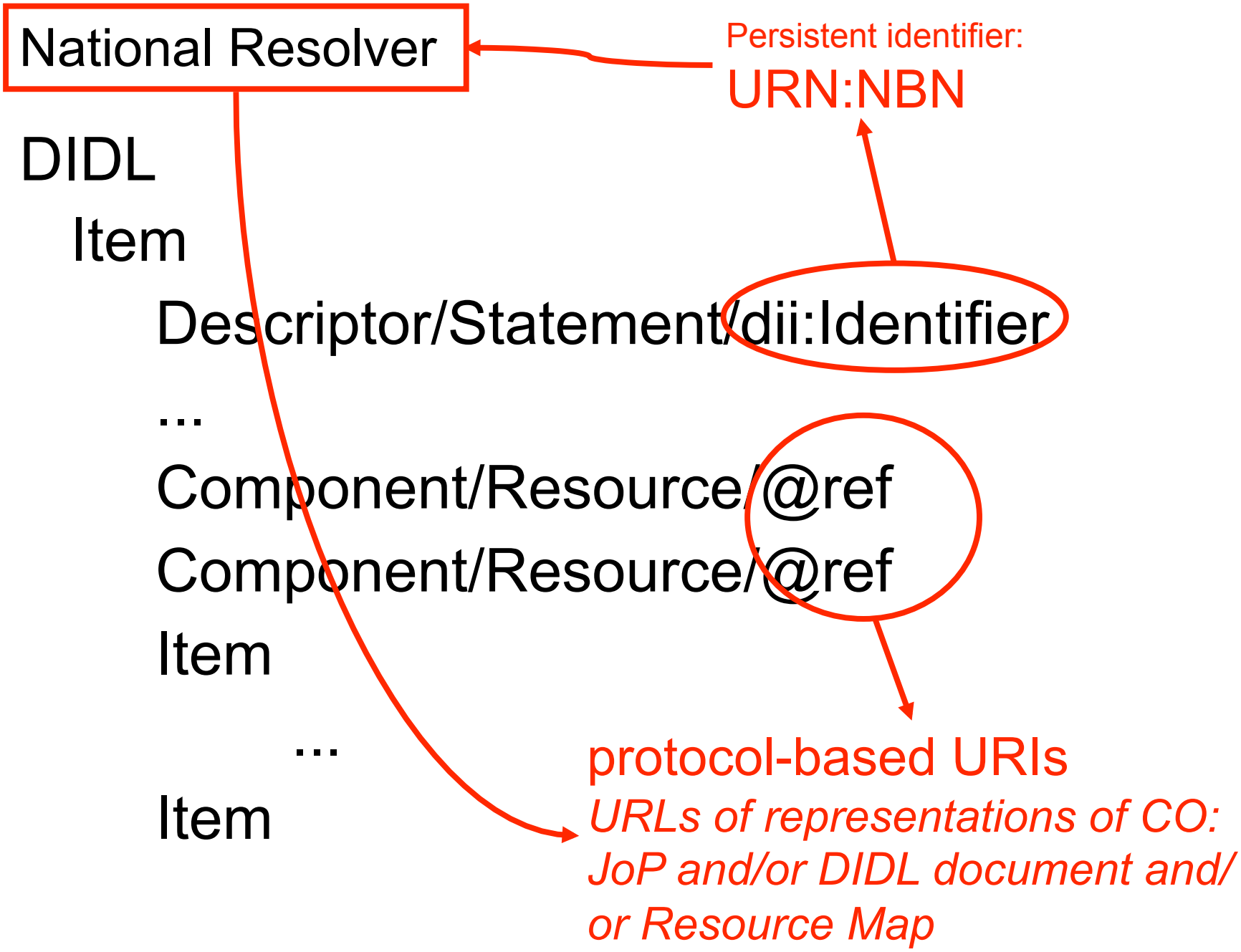
- info:eu-repo/semantics/**descriptiveMetadata** [1, 2, ...]
- info:eu-repo/semantics/**objectFile** [0, 1, 2, ...]
- info:eu-repo/semantics/**humanStartPage** [0, 1]

NEEO proposal:
If Item == **objectFile**

- info:eu-repo/semantics/**publishedVersion**
- info:eu-repo/semantics/**authorVersion**

More types to be added:

- Data set
- Chapter
- etc


National Resolver
Persistent identifier:
URN:NBN
DIDL
Item
Descriptor/Statement/dii:Identifier
...
Component/Resource/@ref
Component/Resource/@ref
Item
...
Item
protocol-based URIs
URLs of representations of CO:
JoP and/or DIDL document and/
or Resource Map

TILBURG UNIVERSITY

ReM service based on URN:NBN resolver

- Resolver harvest regularly all repositories →resolver has up-to-date information about content of the repositories
- Resolver can convert DIDL docs to ReMs and give URLs to the ReMs:

  http://<resolver>/rem/atom/<identifier>
  http://<resolver>/rem/rdf/<identifier>

- Repository content can now be published anywhere on the Web

TILBURG UNIVERSITY

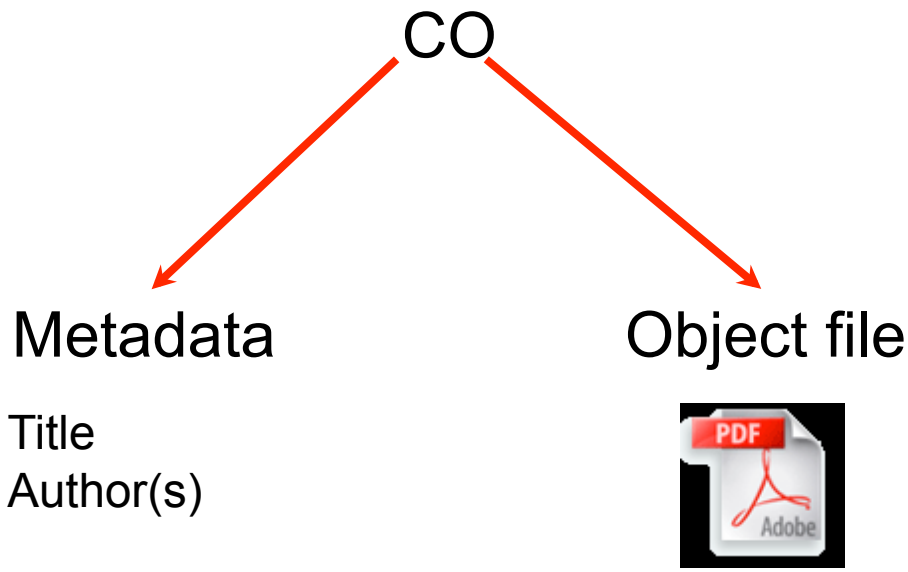
ReM service based on URN:NBN resolver (2)

- In one go, the Web is populated with 150.000 ReMs
- The protocol-based URIs of the ReMs are persistent as long as the central service is in place: no changing URLs; objects can move, repositories can move, ...
- The component objects, like the metadata (included by-value in the DIDL) become real citizens of the Web: Web resources with URIs
- This is proposal that hopefully becomes reality.

TILBURG UNIVERSITY

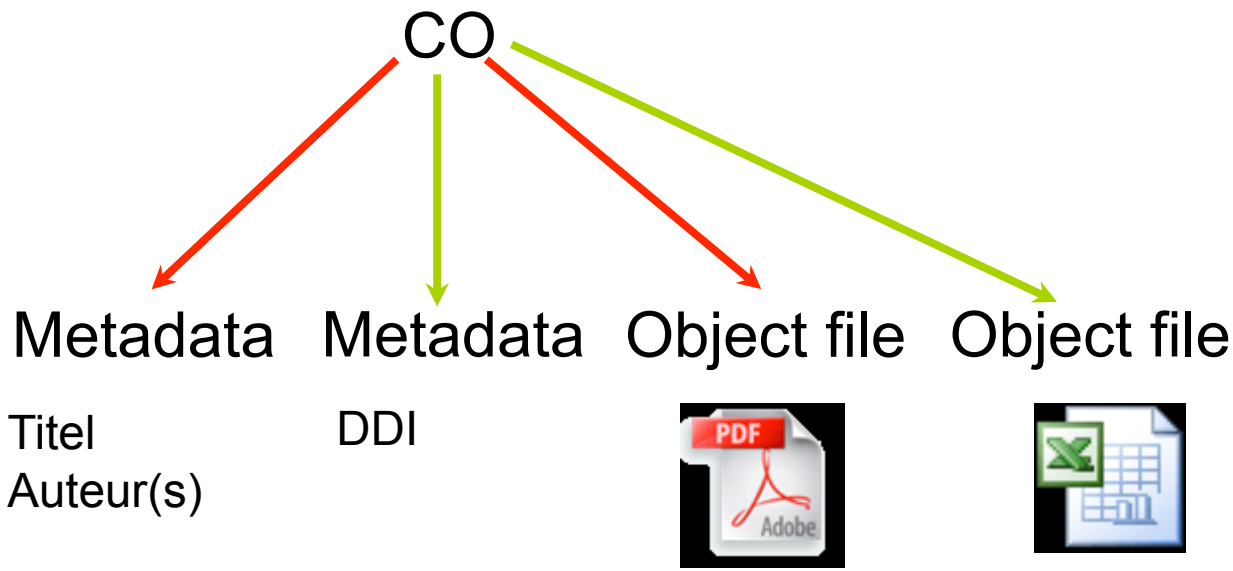
DAI based Aggregations

- In the Netherlands all authors get a Digital Author Identifier which is included in the repository records.
- Search or **aggregation service**: ReM of Aggregation of all publications of an author; publications can be distributed over repositories: *http://<author service>/rem/atom/<dai>*
- New publications will be added automatically to the Aggregation and become visible in the feed.
- Aggregation of Aggregations

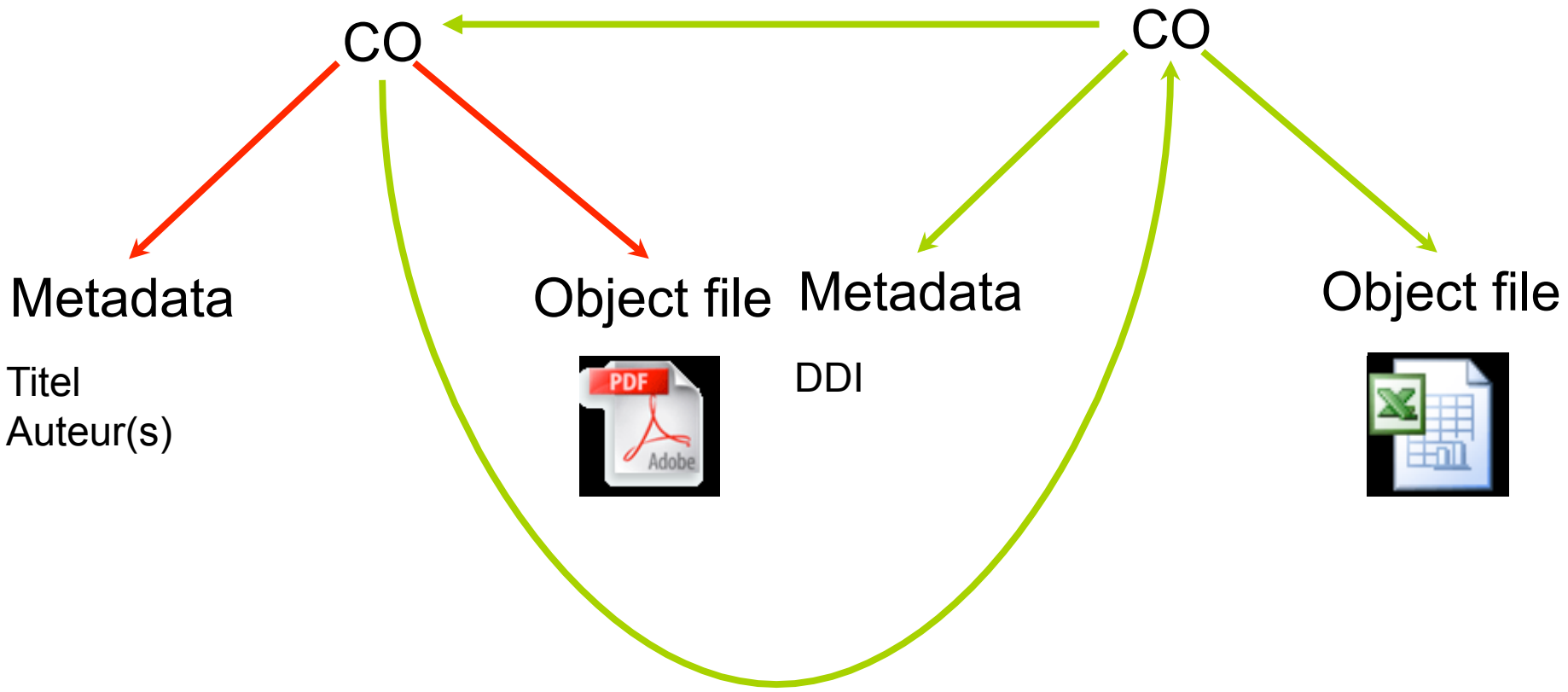

CO
Metadata
Object file
Title
Author(s)
PDF
Adobe

TILBURG UNIVERSITY

Enhancing publication with data set



- What belongs to what? Extra semantics
- Same data can be used by other publications. Separate CO for data set
- Note that for some researchers the data set and the publication belong together; without the publication the data set misses its context.



CO can be component of another CO

Aggregation ore:aggregates / ore:isAggregatedBy Aggregation

TILBURG UNIVERSITY

Going inside: the case of the European Values Study

- International survey study on the Values of Europeans. Each 8 or 9 years data are collected in all European countries.
- 800 publications used the data of the EVS
- The co-ordinator of EVS (Tilburg University): the relations between the publications and the EVS data sets must become more explicit. Information must be added to the publications.
- E.g., direct links to the questions in all languages

EVS

- Unit of analysis for the enhancement of the publications: tables, diagrams, ..., i.e., relations between concepts = statements
- Web of relations including (aggregation) levels
  - Publications (800) – citation links
  - Statements
  - Concepts
  - Variables (internal and external)
  - Data

TILBURG UNIVERSITY

EVS

- With ORE, statements, concepts, ... can become Web citizens in themselves.
- We can refer to them, we can interrelate them in new ways.
- Combination between ORE and DDI 3.0
- This is all very early work, but exciting!