

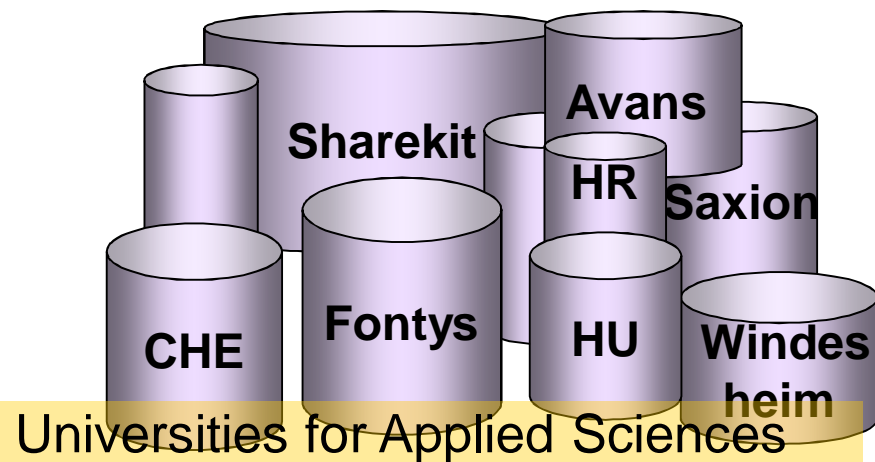
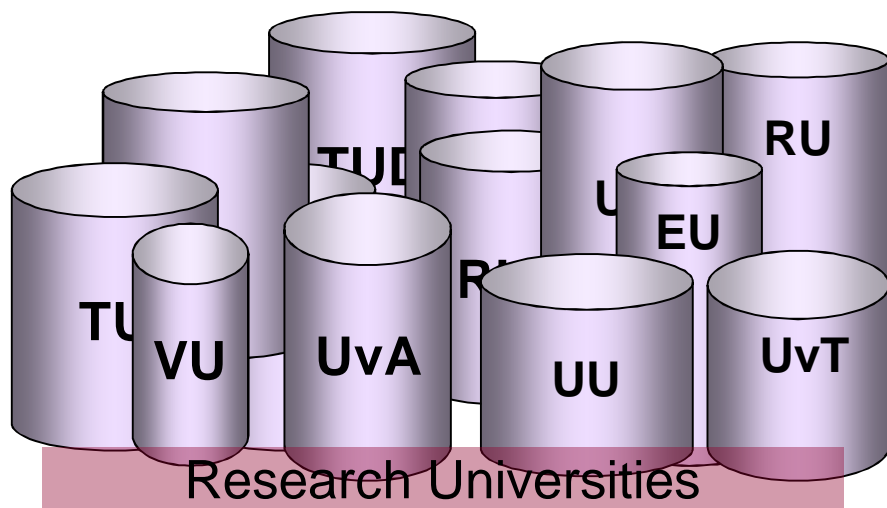
Metadata interoperability in DRIVER and Europeana

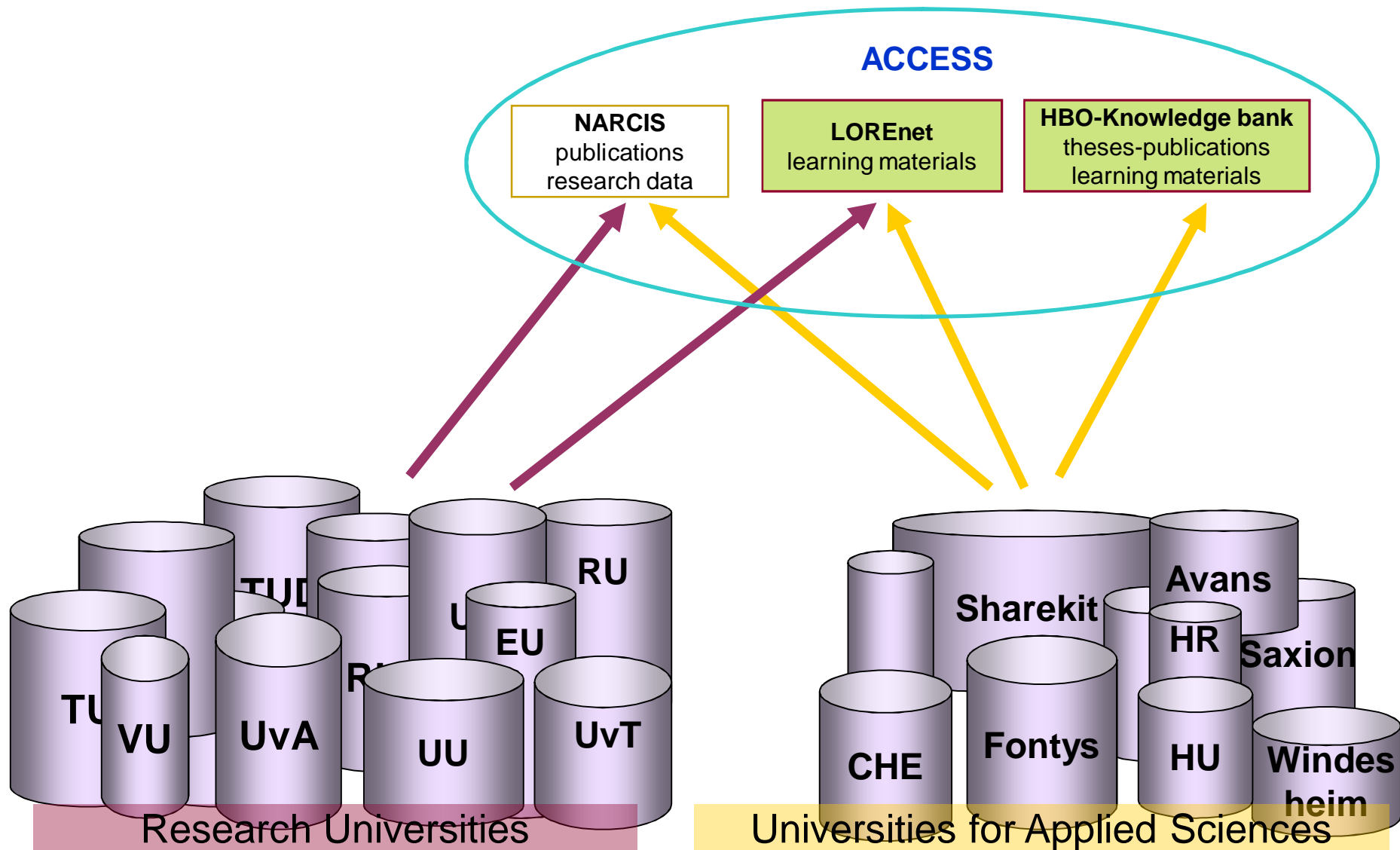
Paul Doorenbosch
KB, national library of the Netherlands

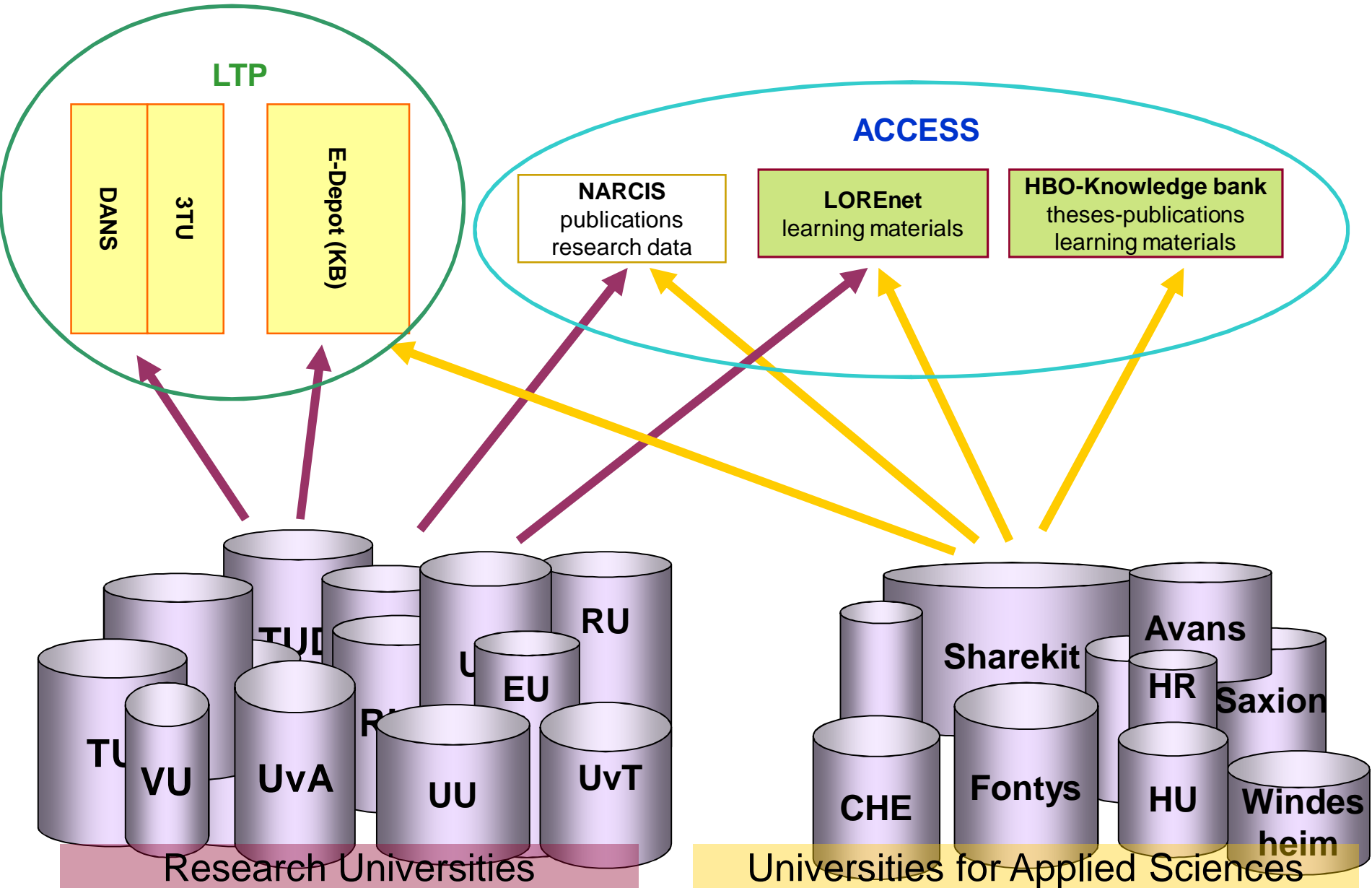


Driver (Digital Repository Infrastructure Vision for European Research): Open Access to European research Material: operational network and service

Europeana: Access to European Culture; network and service, based on The European Library-project and now operational service for integrated access to European national libraries.

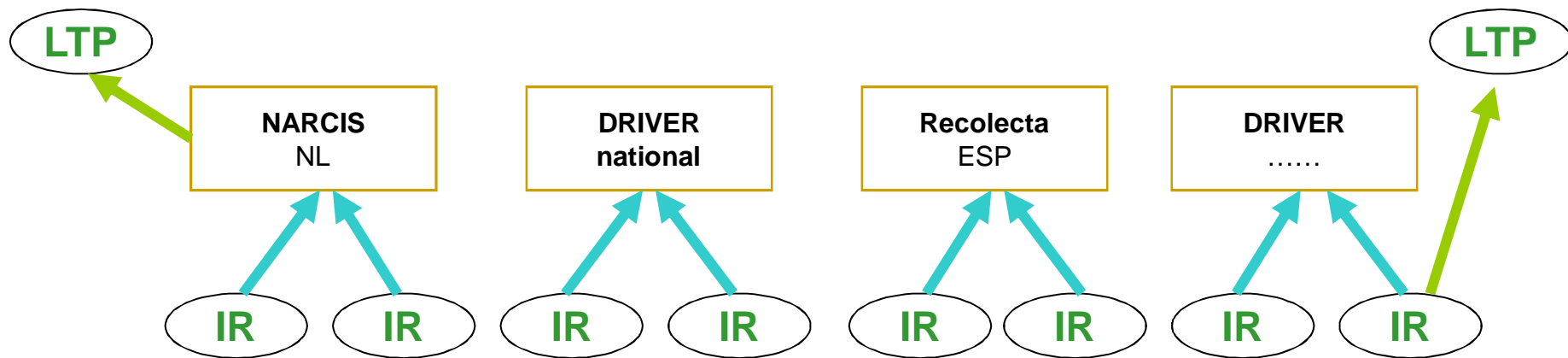






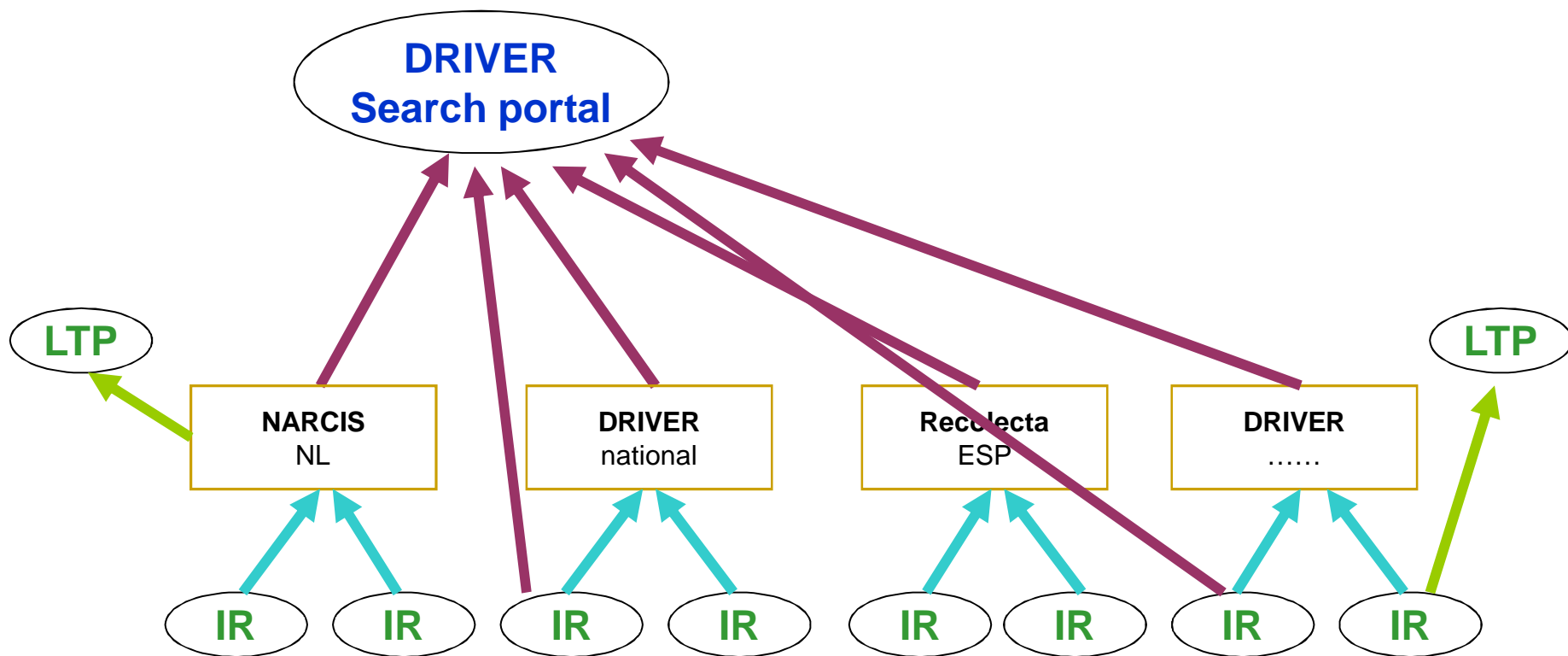


USER



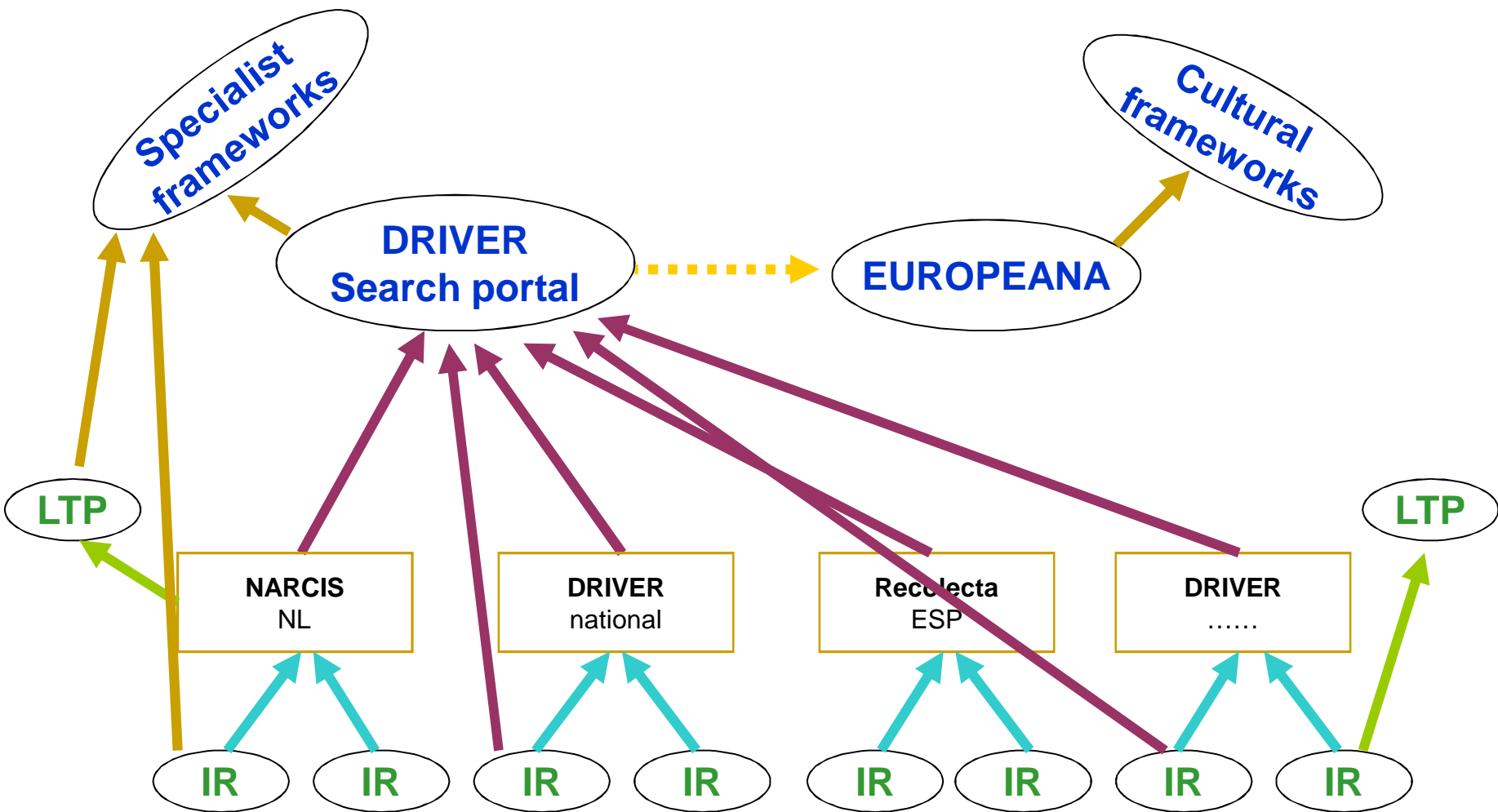


USER





USER



- **Open Access; centralized access point**
- Institutional repositories and some thematic repositories
- Mainly publications; in a lot of different formats; growing amount of research data are handled differently; sometimes normalized
- **Publications:** one framework, one access point, harvesting; **one standard for metadata**
- **Data:** no framework, no common standard; posting
- Enhanced Publications (ePrint + research data = publication)
- Integration, searching and navigation via metadata
- **Resource ('object') stays at IR**
- **Long Term Preservation is part of the infrastructure;** both for publications and for research data; organised locally, regional or national and in different ways.

IR is responsible for

- Making and maintaining metadata
- Normalisation of metadata
- Making it harvestable
- Copyright clearance (Open Access)
- Stable access to linked resource (objects)
- Organizing Long Term Preservation and Access

Central organisation is responsible for

- Harvesting of metadata
- Integrated Access to metadata
- Exploring new services

Institutional Repository

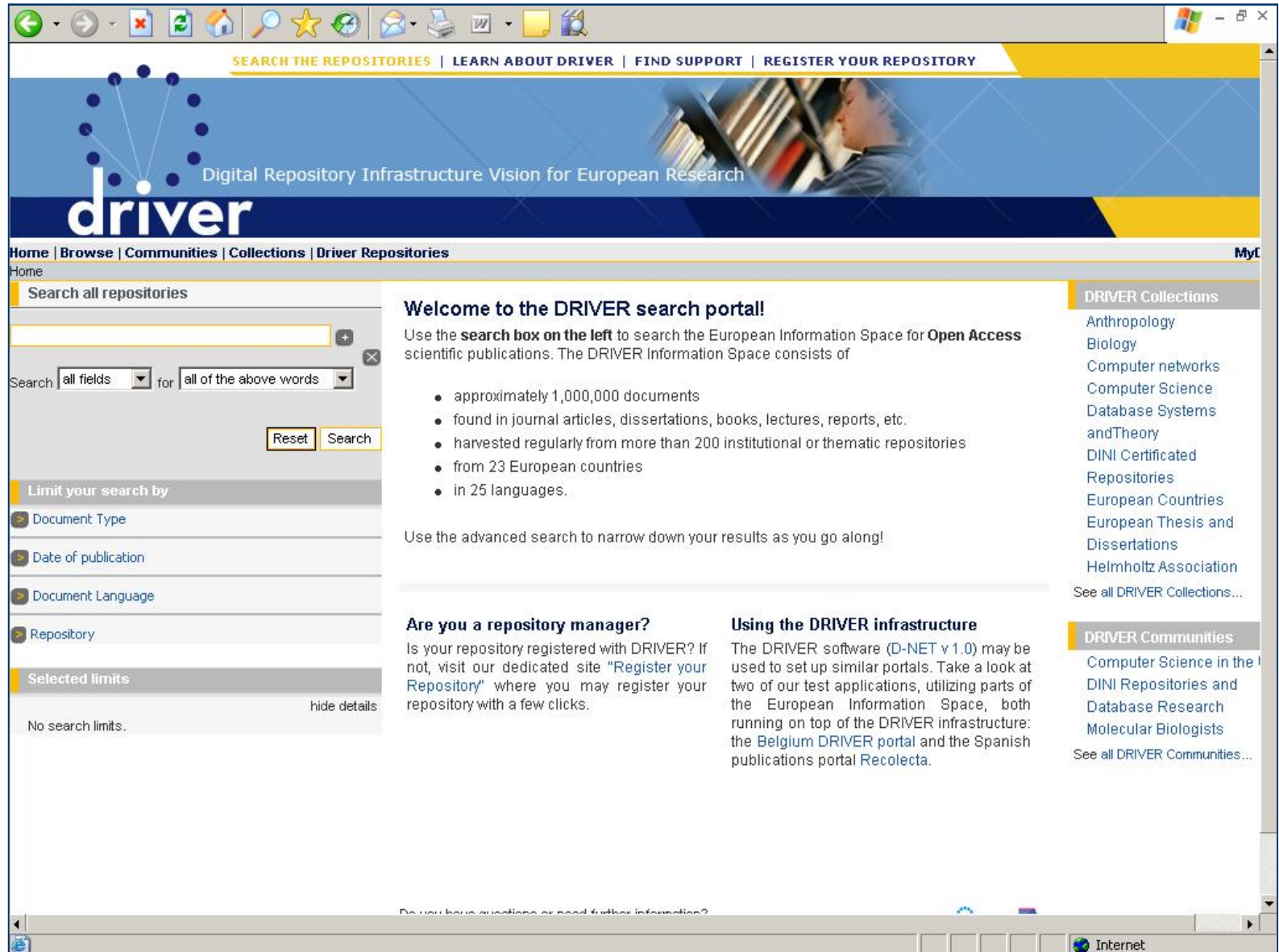
- Repository (metadata)
- Unqualified Dublin Core (wrapped in MPEG21- DIDL)
- OAI-PMH (service)
- Local Access Service
- Resource storage

National Organisation

- OAI-PMH (harvester)
- Repository (metadata)
- National Access Service

International Organisation

- OAI-PMH (harvester)
- Repository (metadata)
- International Access Service



SEARCH THE REPOSITORIES | LEARN ABOUT DRIVER | FIND SUPPORT | REGISTER YOUR REPOSITORY

Digital Repository Infrastructure Vision for European Research

driver

Home | Browse | Communities | Collections | Driver Repositories

Home

Search all repositories

Search +

Search all fields for all of the above words

Reset Search

Limit your search by

- Document Type
- Date of publication
- Document Language
- Repository

Selected limits

No search limits. [hide details](#)

Welcome to the DRIVER search portal!

Use the **search box on the left** to search the European Information Space for **Open Access** scientific publications. The DRIVER Information Space consists of

- approximately 1,000,000 documents
- found in journal articles, dissertations, books, lectures, reports, etc.
- harvested regularly from more than 200 institutional or thematic repositories
- from 23 European countries
- in 25 languages.

Use the advanced search to narrow down your results as you go along!

Are you a repository manager?

Is your repository registered with DRIVER? If not, visit our dedicated site "Register your Repository" where you may register your repository with a few clicks.

Using the DRIVER infrastructure

The DRIVER software (D-NET v 1.0) may be used to set up similar portals. Take a look at two of our test applications, utilizing parts of the European Information Space, both running on top of the DRIVER infrastructure: the Belgium DRIVER portal and the Spanish publications portal Recolecta.

DRIVER Collections

- Anthropology
- Biology
- Computer networks
- Computer Science
- Database Systems and Theory
- DINI Certified Repositories
- European Countries
- European Thesis and Dissertations
- Helmholtz Association

[See all DRIVER Collections...](#)

DRIVER Communities

- Computer Science in the
- DINI Repositories and Database Research
- Molecular Biologists

[See all DRIVER Communities...](#)

Do you have questions or need further information?

Internet

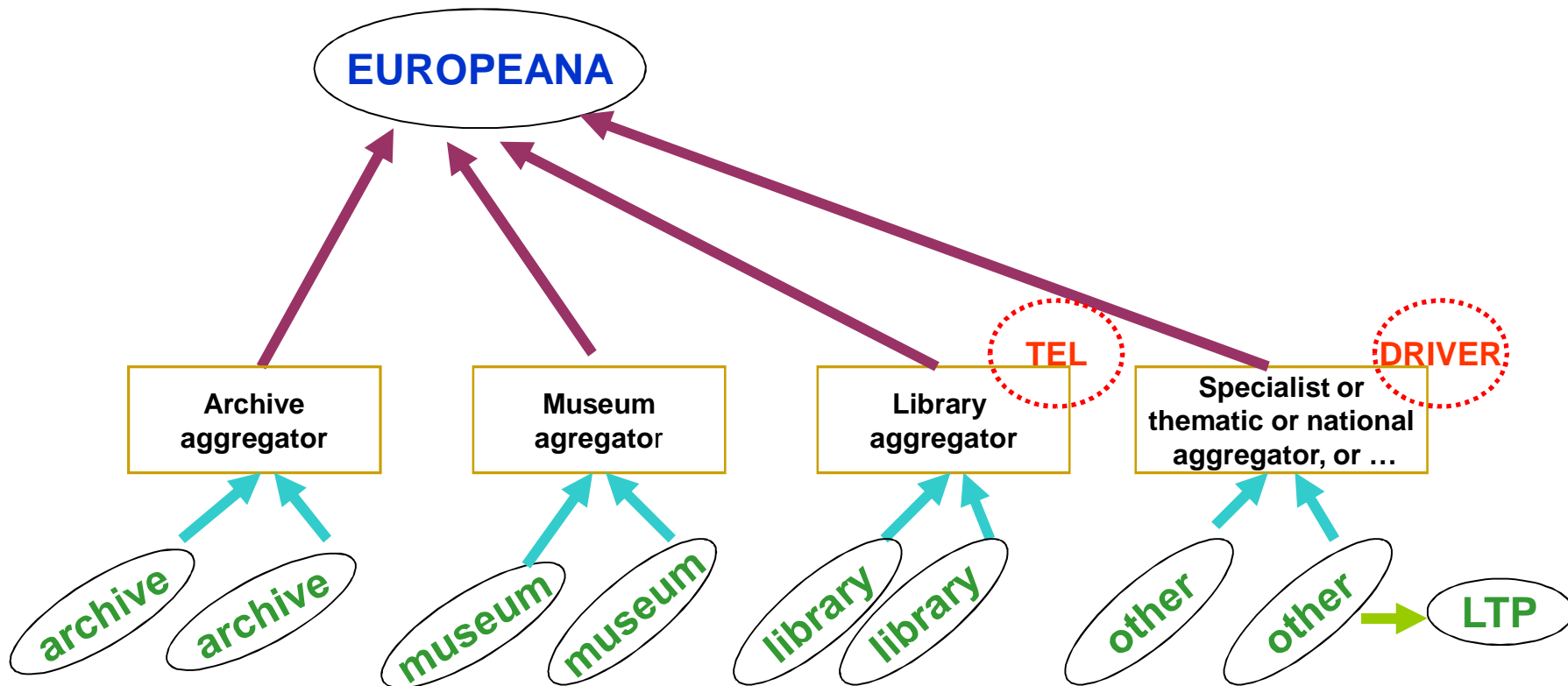
Institutional repositories

- Usually libraries in academic or scientific institutions
- Acquainted with metadata for access
- Already a 'long' tradition in digital environment
- Common standards for describing publications
- Dedicated IT or developers

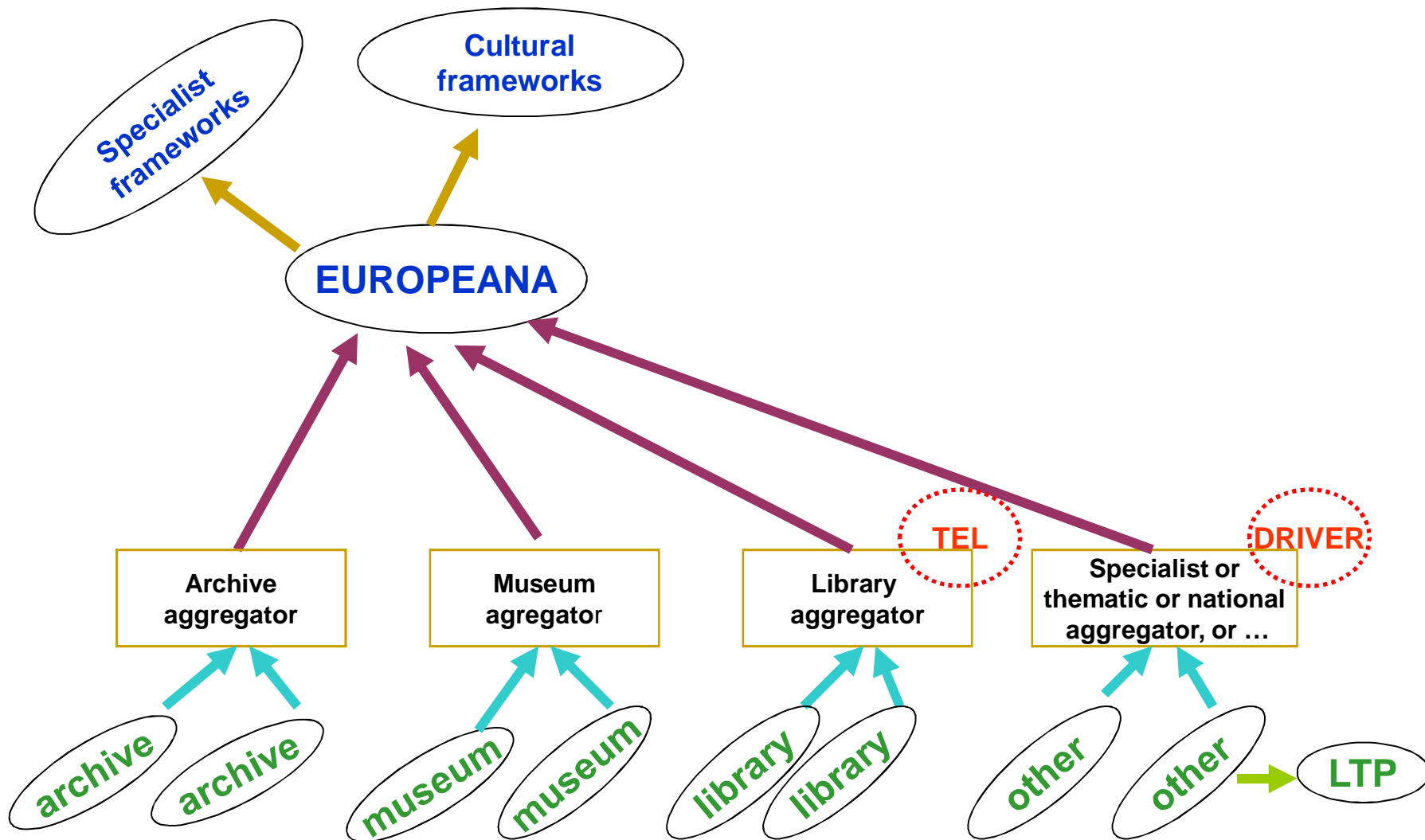
The organisational challenge is bigger than
the technical challenge

- Libraries, Archives, Museum, Moving Heritage organisations, etc
- Printed material, Film, Pictures, Drawings, Music, Spoken word, Letters, Dairies, Minutes, Records
- Scientific heritage, Cultural heritage, Governmental-Parliamentary-Administrative heritage, Personal heritage, ..
- Lot of different metadata formats
- Lot of legacy and closed systems
- Lack of digitised content
- Lack of descriptive metadata for access.
- Few knowledge about interoperability, or at least no skilled people to perform conversions
- Small or no IT departments.

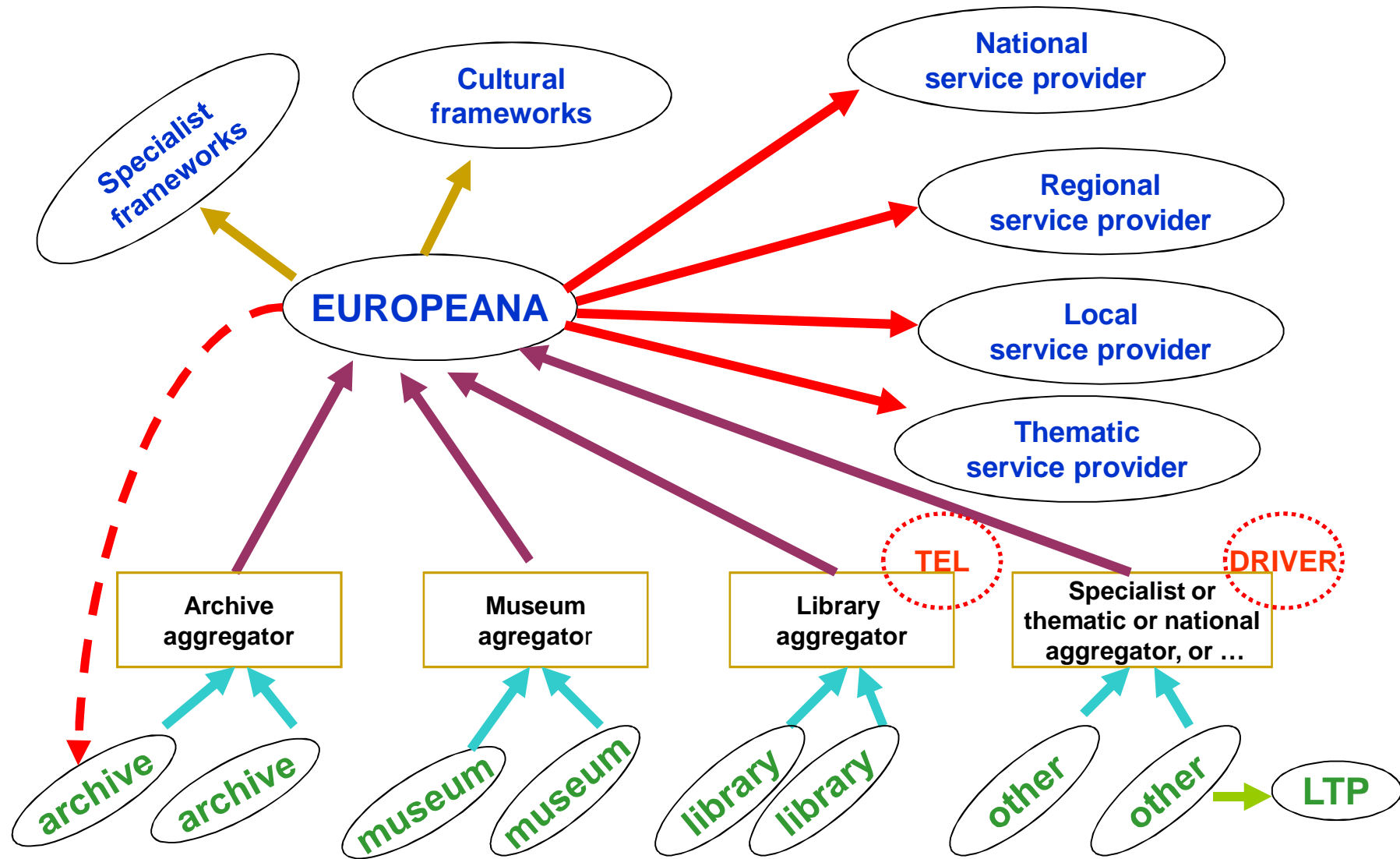
USER



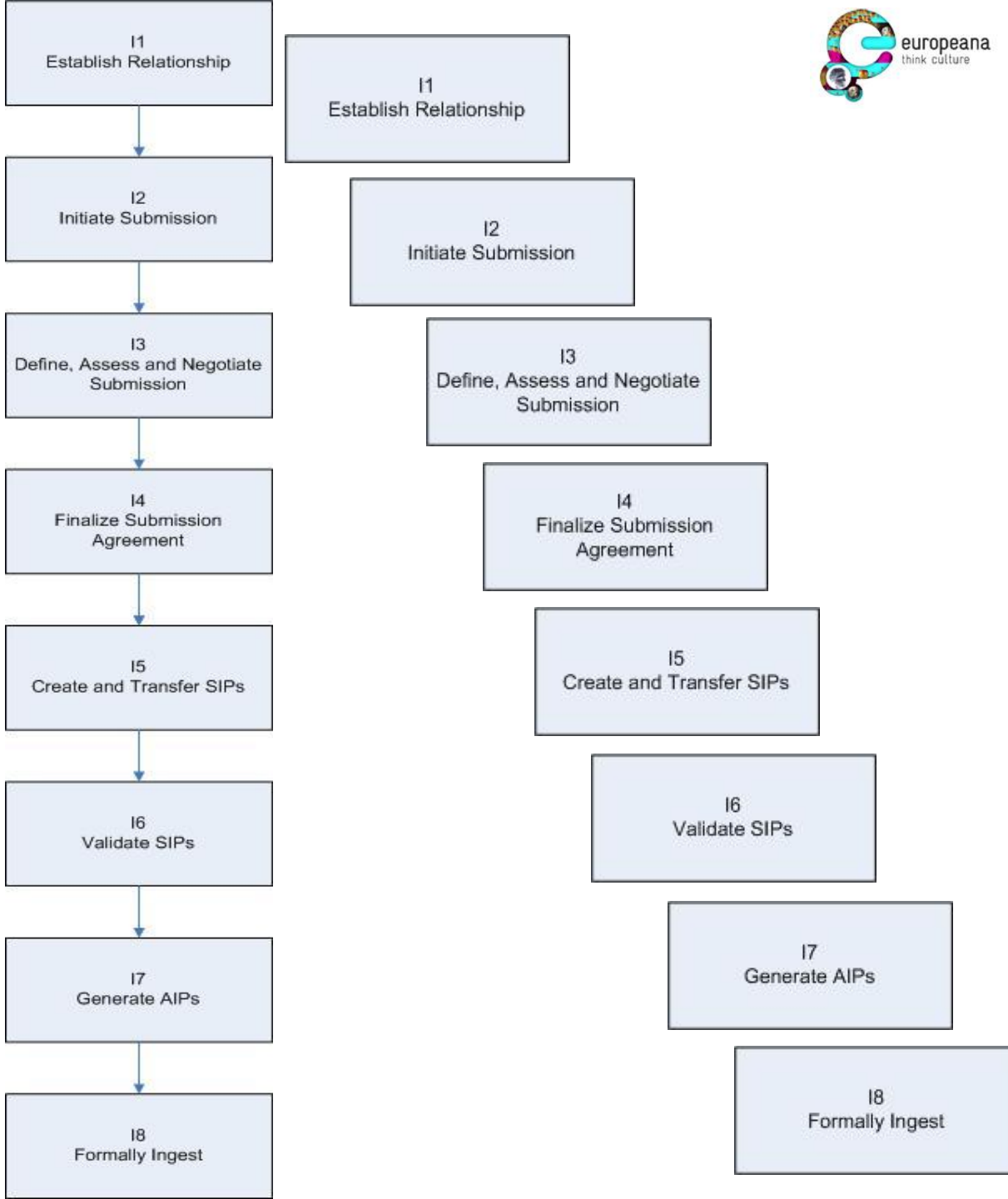
USER

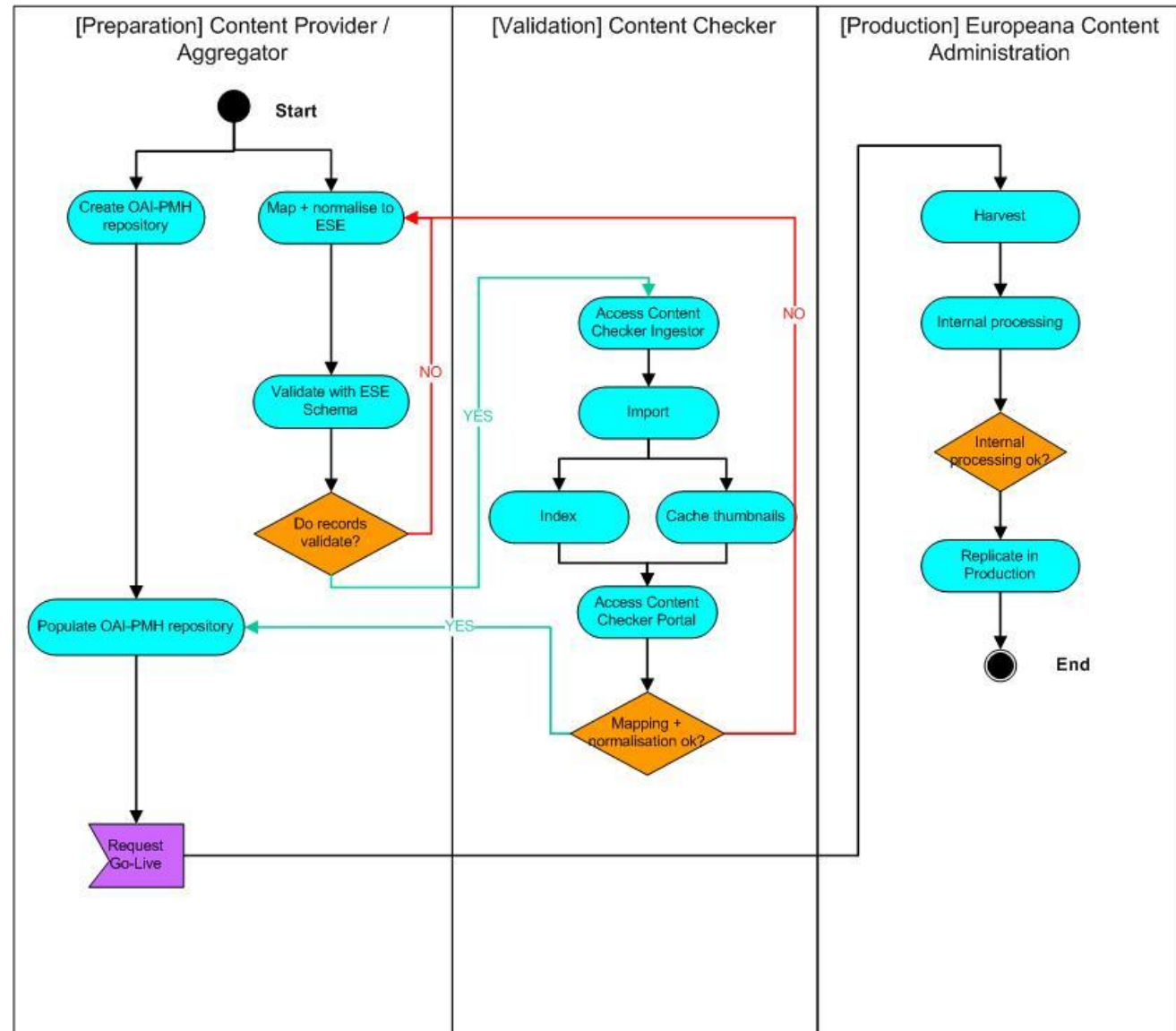
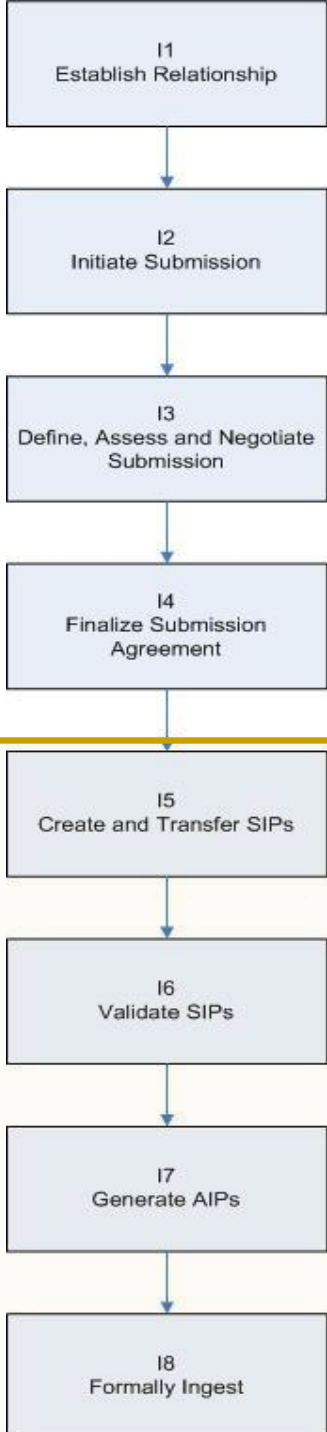


USER



- Content (resources) will stay at the institutions
- Metadata is maintained by the institutions
- Institutions or aggregators normalize metadata
- Aggregators harvest metadata from institutions
- Europeana harvests metadata from aggregators
- Compared to DRIVER aggregators have bigger responsibility for harmonization of metadata
- Aggregators could be national, regional, sectional, thematic, ...
- Bigger institutions can do without an aggregator but have to normalize their own metadata and maintain an OAI-PMH service
- Aggregators could be new organisations, but preferably existing institutions.





- From prototype (searchable catalogue) to operational V1.0 (interactive, semantically rich, more meta-levelled, more networked)
- OAI-PMH is used for harvesting
- E(uropeana) S(emantic) E(lements) is the standard metadata model
- ESE = Dublin Core Metadata Elements Set + some extra fields to meet the needs for the portal.
- Metadata registry
- OAIS (reference model) for organising the ingest process
- No prescribed formats, no prescribed encoded schema's for txt or other type.
- ESE is not enough for semantic connectivity
- Research is done on SKOS and RDF, OAI-ORE, multilingualism etc.
- Needed for semantic connectivity is the richness of the original metadata, but normalisation then is an even bigger challenge

- Practise is still diverse
- Some OAI-PMH harvesting
- Still a lot of batch delivery with conversion on the Europeana side.
- Europeana Local makes institutions 'Europeana-ready' and creates conversion tools/services or gateways.
- A lot of experience from The Europeana Library (in 4 years from few national libraries harvestable to nearly every European library harvestable).
- How not leaving small institutions behind?

EUROPEANA + DRIVER and e-frameworks

- Both services now primarily working with metadata
- Dublin Core is core of current metadata model
- Information is extracted into keywords or abstracts

Issues

- Persistent identifiers
- Enhanced publications
- Integrated multi-media search
- Multilingualism
- Semantics interoperability of records and cross-language searching
- Integrated access to and usability of source data (primary data, content interoperability, such as full-text)



Thank you

Paul Doorenbosch
Koninklijke Bibliotheek
paul.doorenbosch@kb.nl

The logo consists of the letters 'KB' in a bold, gold-colored serif font, centered within a white square box with a thin blue border.

Europeana Semantic Element (ESE)



- **Europeana “Schema”** for the Prototype
- **Based on Dublin Core** Metadata Elements Set (DCMES)(ISO)
49 Elements (26 Elements & 23 Refinements)
- Created through discussions in July/August 2008



- europeana:**country**
- europeana:**provider** (dc:source)
- europeana:**language** (dc:language)
- europeana:**type** (dc:type, dc:format)
- europeana:**year** (dc:date)
- europeana:**isShownBy** (dc:relation)
- europeana:**isShownAt** (dc:relation)
- europeana:**object**
- europeana:**uri** (dc:identifier)



- All normalised:
 - Syntax
 - Value
- Let's examine their characteristics

Normalised ESE terms: **Country**



- Definition:
 - Country of content provider.
 - If several countries: Europe
- Format:
 - String, ex: *switzerland, germany,...*
- Reference:
 - TEL controlled list. Supports TEL interface translation mechanism
- Mechanism:
 - Manual
- In portal:
 - Facet browsing of search results



- Definition:
 - Organisation sending the data to Europeana
- Format:
 - String, ex: *Musées lausannois, Nasjonalbiblioteket,...*
- Reference:
 - Europeana controlled list of content providers:
<original_name>
- Mechanism:
 - Manual but potentially can be automated
- In portal:
 - Facet browsing of search results



- Definition:
 - Language of provider's country (ESE:languages of the metadata)
- Format:
 - 2-letters, ex: it, no,fr, en, es,...
- Reference:
 - ISO639-1 language codes
 - Exception: If several languages: "mul"
- Mechanism:
 - Manual but potentially can be automated
- In portal:
 - Facet browsing of search results



- Definition:
 - Type of the original object
- Format:
 - String
- Reference:
 - 4 Europeana types: *IMAGE, TEXT, SOUND, VIDEO*
- Mechanism:
 - Manual: Mapping specified by content provider
- In portal:
 - Categorisation display
 - Facet browsing of search results



- Definition:
 - Date of creation of the original object (analog or born digital)
- Format:
 - 4 digits [YYYY], ex: 1950
- Reference:
 - Europeana year
- Mechanism:
 - Automatic extraction with “YearExtractor” converter
- In portal:
 - Facet browsing of search results
 - Browsing by time (timeline)



- Definition:
 - URL to the digital object
- Format:
 - URL (http://...)
- Mechanism:
 - Automatic or manual
- In portal:
 - Linking



- Definition:
 - URL to the digital object with context
- Format:
 - URL (http://...)
- Mechanism:
 - Automatic or manual
- In portal:
 - Linking



- Definition:
 - URL to the digital object as thumbnail
- Format:
 - URL (http://...)
- Mechanism:
 - Automatic or manual
- In portal:
 - Display

Normalised ESE terms: **URI**



- Definition:
 - Record identifier for Europeana system
 - Format:
 - URI
 - Mechanism:
 - Automatic: special algorithm guaranteeing uniqueness (and integrity) of records
- <http://www.europeana.eu/resolve/record/91101/OBAF44EDF8B98F1322DEEAD4AB989778E6394418>
- In portal:
 - MyEuropeana
 - Full digital object view in Europeana