

# The 3 Pillars of Standardisation

Driving interoperability in research information

7 Konferenz für Sozial- und Wirtschaftsdaten  
Berlin, February 8, 2017

Presentation by:

Ed Simons



# Speaker

## Currently:

- President of euroCRIS, the *International Organisation for Research Information*.
- Senior Project Manager Research Data Services, *Radboud University*, The Netherlands.

## Formerly:

- Head of the Central IT Support Unit at *Radboud University*.
- Information Manager of the *International Federation of Catholic Universities (IFCU)*.
- Project Leader *International Association of University Presidents (IAUP)*
- International Project Leader, IT Development Projects for *NUFFIC, the Dutch Organisation for University Development Cooperation*.

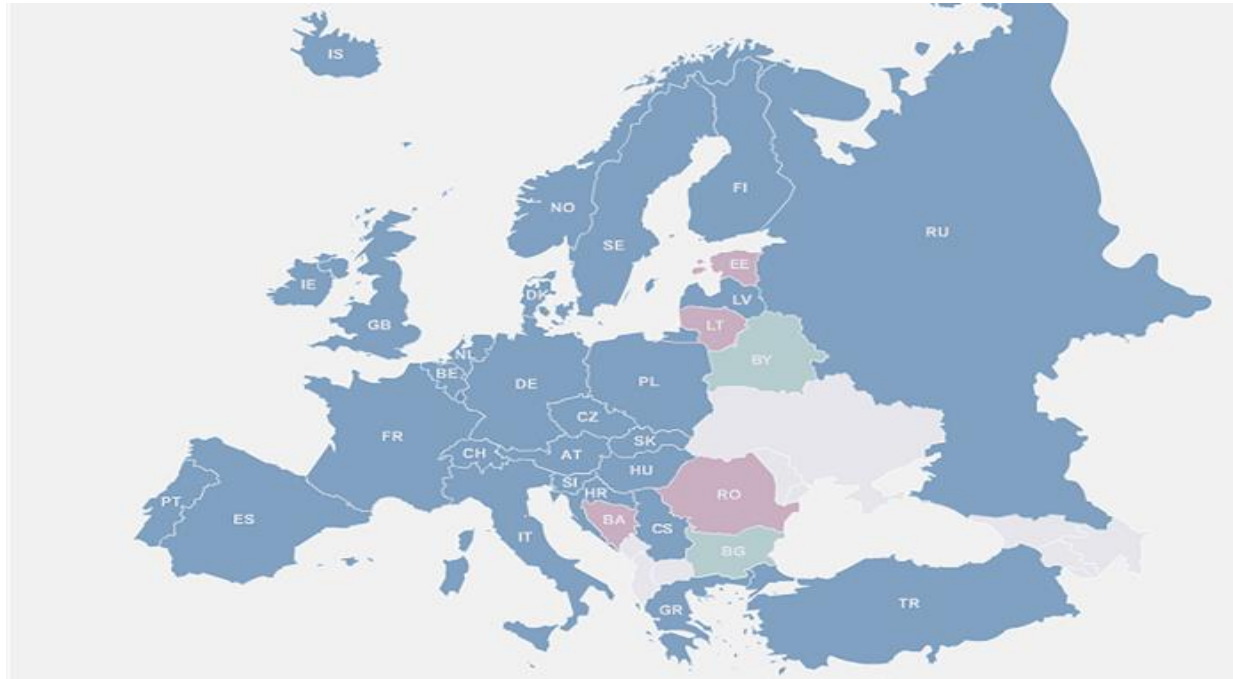
Ed Simons

# What is euroCRIS?

- An international not-for-profit association of experts and users of research information and research information systems (CRIS), with offices in The Hague, The Netherlands.
- *Main activities:*
  - Development and governance of the CERIF data model and promotion of its use.
  - Promotion of cooperation between stakeholders (organisations) in the RI Domain. To fulfill this function, euroCRIS regularly organises international events: Membership Meetings (twice a year) and Conferences (every 2 years).
- *Upcoming events:*
  - Membership Meeting, Dublin, May 2017.
  - Strategic Membership Meeting, Bratislava, October 2017
  - CRIS2018 International Conference, Umea, Sweden, June 2018.

International  
Organisation  
for Research  
Information

# euroCRIS: Membership



## Members outside of Europe:

Australia – Brazil – Canada – China - Colombia – Iran – Israël - Malaysia – Nigeria – Pakistan – South Korea - USA

200+ members from 45 countries (mainly Europe)

International  
Organisation  
for Research  
Information

# euroCRIS: Strategic Partners



In preparation:



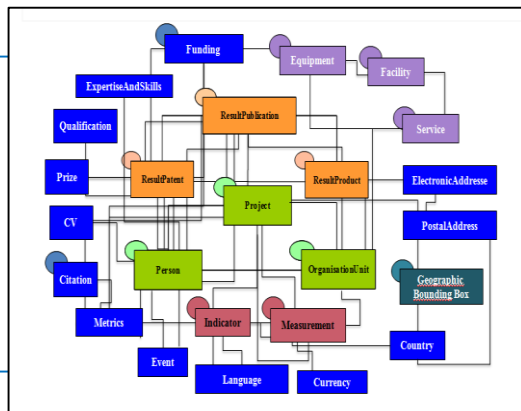
International  
Organisation  
for Research  
Information

euroCRIS  
Current Research Information Systems

# CERIF: euroCRIS's main “product”

**An international standard relational data model for storage and interoperability of research information**

**Official EU Recommendation to Member States**



**Reference model for development of Research Information Systems (CRIS)**

**Standard exchange format (CERIF-XML)**  
for interoperability between systems

Common  
European  
Research  
Information  
Format

# Standard(i)s(ation): why needed?

To allow for an optimal exchange (flow of) information between systems.

In other words: to realise optimal

*interoperability*

between information resources.

Interoperability  
of information  
requires  
standardisation

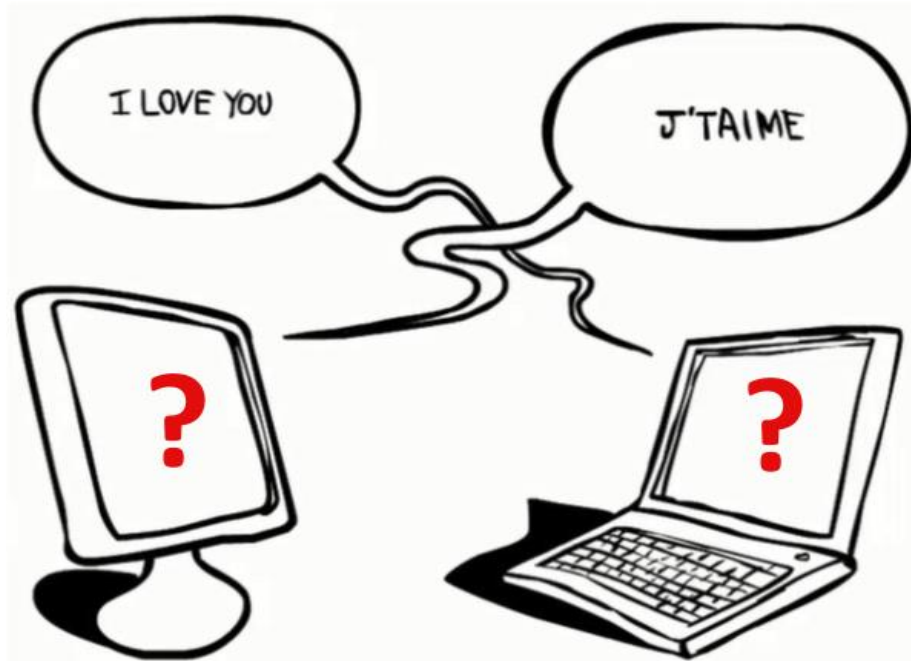
# Why is interoperability needed?

- *Many stakeholders*: researchers, funders, managers, reviewers, libraries, research administrations, academies of science, association of universities, disciplinary networks, etc...
- *Various applications/systems*: CRIS, publication repositories, data repositories, project management systems, etc...
- *Various use cases*: benchmarking/performance evaluation, grant submission, reporting, management information, profiling of institutes, CV-exposure, etc...

Interoperability  
in the  
Research  
Information  
Ecosystem



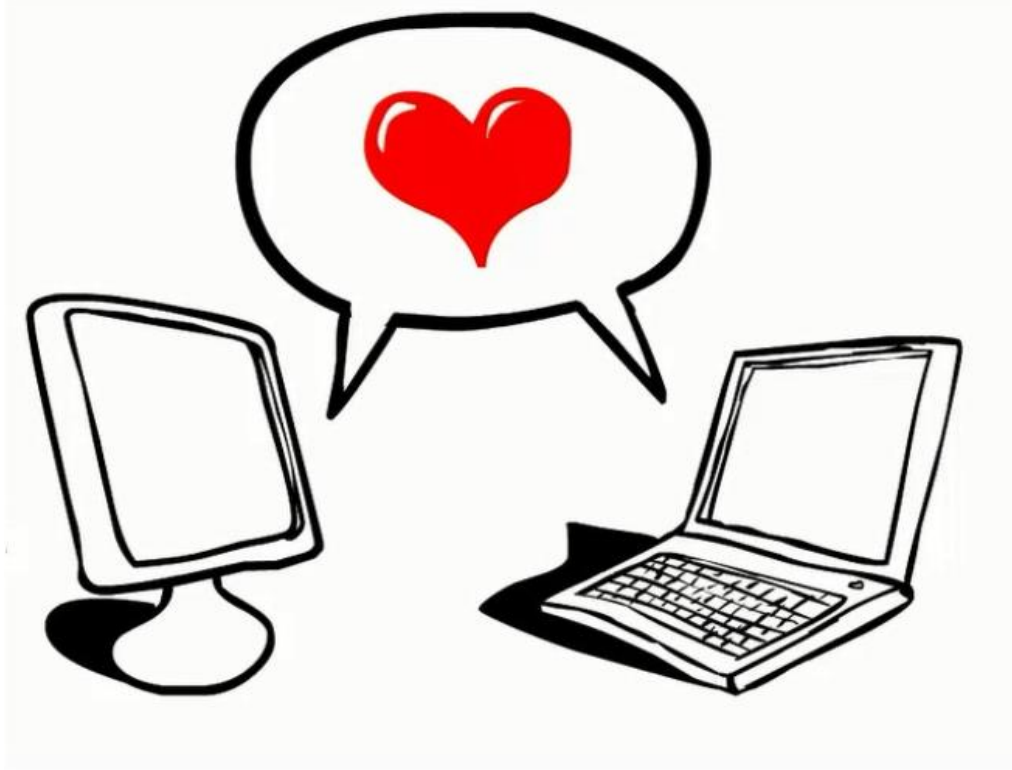
# Aspects hampering interoperability



No shared  
standard  
vocabulary

Lack of standard language: shared vocabularies (terms) and semantics (meaning of terms).

# Aspects hampering interoperability



No shared  
standard  
vocabulary

Solution: use of standard, shared, research information concepts and semantics.

# Aspects hampering interoperability

No standard  
use cases



Who is the best performing vegetable grower?

# Aspects hampering interoperability

No standard  
use cases



Difficult to answer because of lack of standardised use case definition.

# Aspects hampering interoperability

One morning at the Research Administration Office...

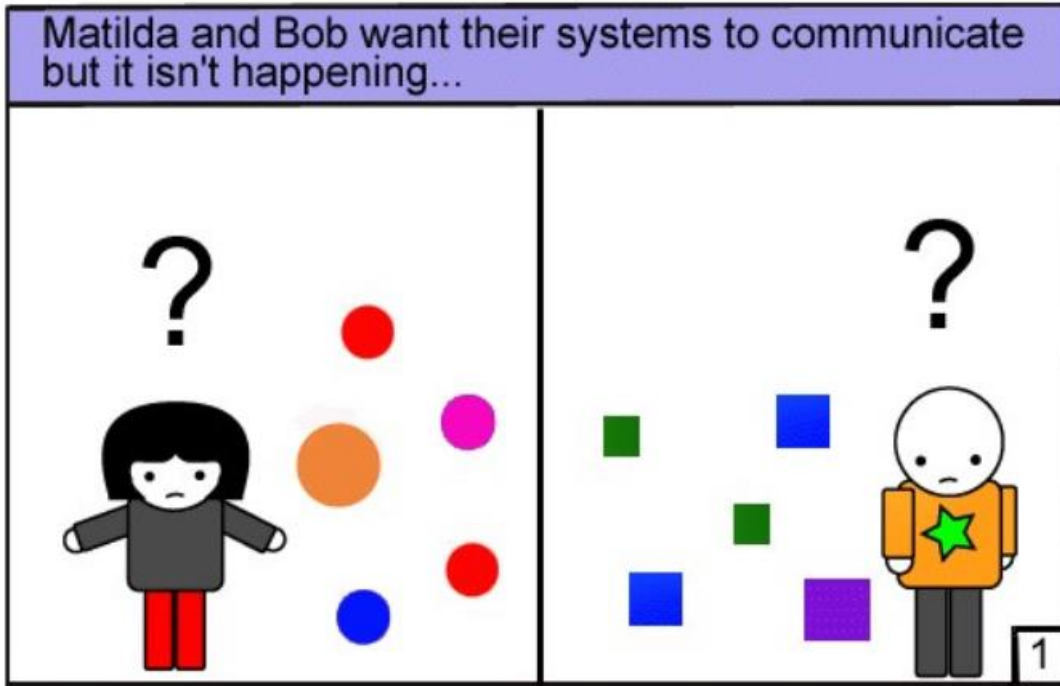


Lack of standard unique and persistent identifiers (in and between systems)  
(applies not only to persons but also to other objects in the RI ecosystem)

Solution: implementation of international standard identifiers

No unique  
identification  
of data  
elements

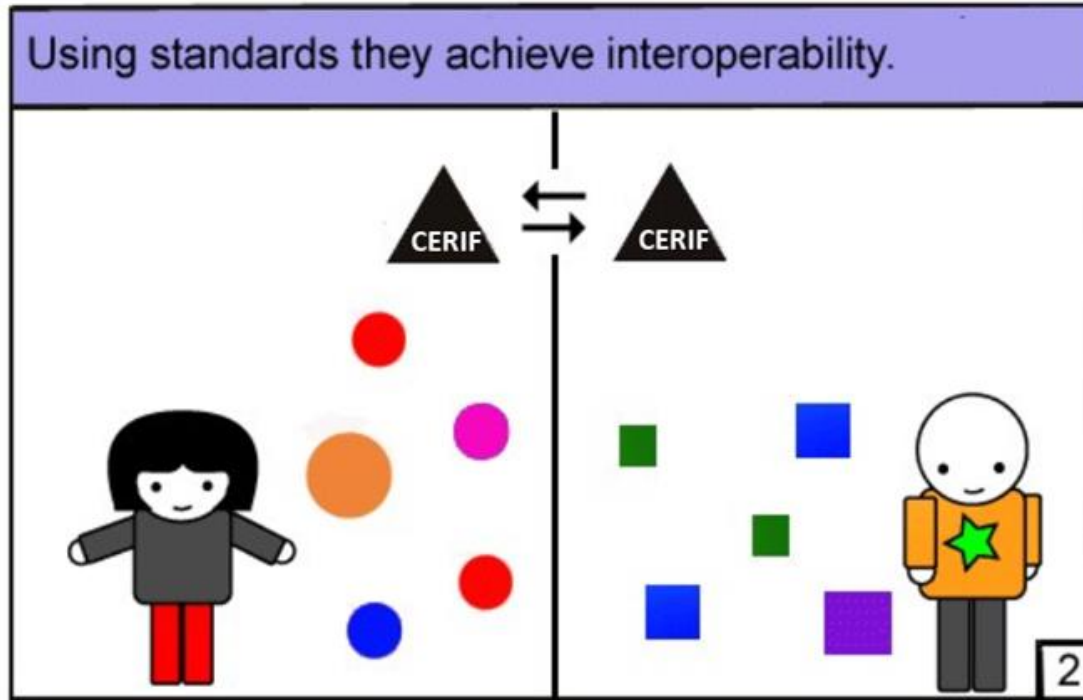
# Aspects hampering interoperability



Lack of shared standard exchange format.

No standard  
exchange  
format

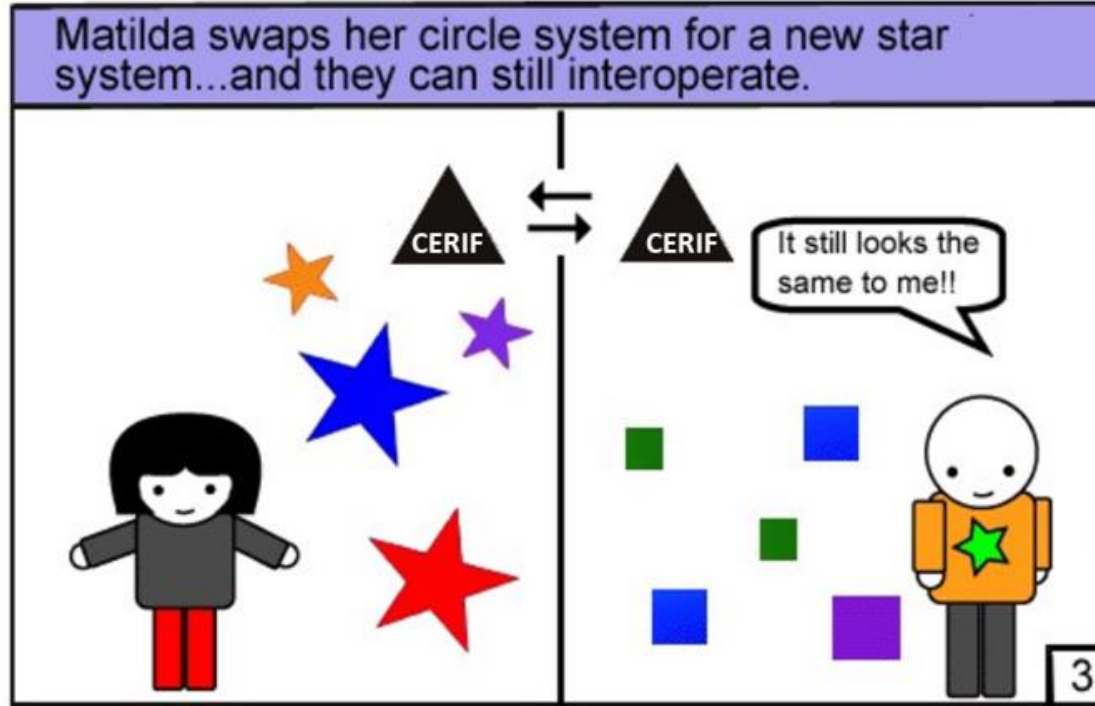
# Aspects hampering interoperability



No standard  
exchange  
format



# Aspects hampering interoperability



No standard  
exchange  
format



# Conclusion: the 3 Pillars of Standardisation

- *Shared (unique) Definitions* of Terms (vocabularies) and Use Cases.
- *Unique Identifiers.*
- *Standard Exchange Format.*

Vocabularies  
Identifiers  
Exchange format

# The 3 Pillars of Standardisation: solutions

- Both standards concerning the various aspects of interoperability as well as organisations governing these standards have matured and are now in place.
- *Definitions*: e.g. CASRAI-vocabularies and profiles
- *Identifiers*, e.g.: ORCID, ISNI, DOI,....
- *Exchange format*, e.g.: CERIF (euroCRIS)

Solutions are  
(becoming)  
available

# Standardisation in Research Information Data: some positive developments

Positive  
developments

- Research Information Communities, e.g. CRIS and Repository communities are more and more approaching each other and working together.
- There is a growing awareness with policy makers and stakeholder organisations both on a national and international level that standardisation is necessary in order to realise optimal research information infrastructures.
- There is a willingness and commitment to cooperate between standards organisations backed by a cooperative attitude of system developers/vendors to implement the standards.

# Communities coming together

- The last few years various Strategic Partnerships have been agreed upon between organisations representing various stakeholder groups within the research information domain (e.g. Between euroCRIS and COAR, EARMMA, VIVO, etc...)
- Last November a basic agreement has been reached between CASRAI, EARMMA, euroCRIS, JISC, OpenAIRE, ORCID and ALLEA (to be confirmed) to work together in the promotion and realisation of a European Research Information Infrastructure.

Positive  
developments

# Communities coming together

- This also results in cooperation on a more practical level, e.g.:
  - The joint definition by OpenAIRE and euroCRIS of a subset of the CERIF-format to harvest information for OpenAIRE from CRIS-systems.
  - The collaboration between VIVO and euroCRIS to create a mapping between the CERIF-model and the VIVO-ontology.
  - The Indicators and Metrics community working together with euroCRIS in the “Snowball Metrics” initiative and the ACUMEN-portfolio initiative.

Positive  
developments

# Growing awareness on the policy level

Conclusion of a report of a working group of the European Parliament (April 2014):



Science and Technology  
Options Assessment

## Measuring scientific performance for improved policy making



Science and Technology Options Assessment  
European Parliamentary Research Service  
European Parliament  
April 2014  
PF.522.383



“We conclude that a European Integrated Research Information Infrastructure is *technically feasible* thanks to recent technological developments *and especially the maturity of the European CERIF standard*, which allows seamless interlinking of datasets and/or research information systems, in different formats and including non-CERIF systems”  
(page 14 of the Report)

Positive  
developments

# Growing awareness on the policy level

Positive  
developments



## Science Europe Position Statement

On Research Information Systems

NOVEMBER 2016

- ▶ Promote the adoption of research information systems with the following characteristics:
  - Economic value-creation for service providers that does not depend on siloing and closing data;
  - No pay-walling of research activity data, given that dataset combination is a basic requirement to understand the research landscape; and
  - Interoperability based on a common exchange format and open standards for the definitions of the entities (such as researchers, organisations, grants, activities, and outputs) and attributes in the research information domain, and for the identification of these entities with unique and persistent identifiers.

# Also nationally things are happening, e.g.

- In Italy ORCiD is being implemented on a national scale, through the CRIS-infrastructure (all universities have CRIS-systems in place).
- In Portugal both ORCiD and CERIF are planned to be implemented, within the framework of the PT-CRIS (national CRIS-project).
- In Greece a national Research Information Infrastructure has been developed based on CERIF.
- In Flanders the universities supply research information to the “national” RI database by means of CERIF.
- In the Netherlands projects are in progress to implement both ORCiD and CERIF on a national level, the use of CASRAI profiles is under consideration.
- In the UK both ORCiD and CASRAI-profiles are being implemented and CERIF considered as exchange format.
- In Brazil the realisation of a national Research Information Infrastructure based on CRIS and CERIF has been planned, with (consultancy) support of euroCRIS.

Positive  
developments



# The 3 pillars of standardisation: challenges

## Challenges

### *Definitions of standard terms and use cases:*

- *Mindset / attitude issue:* avoidance and breakthrough of particularism and silo-ed behaviour of organisations and stakeholders. Both on the national and international level an open-minded and cooperative (“reaching out”) attitude aimed at reaching joint and commonly agreed upon definitions is necessary (but also a huge challenge).

### *Identifiers:*

- *Governance issue:* avoidance of existence of multiple “unique” id’s of a given type for the same object. Especially for person-ID’s (ORCiD, Researcher ID,...) this is a challenge, since a researcher can in principle register an ID multiple times.
- *Mapping issue:* correctly linking the various types of ID’s to each other.

### *Exchange format:*

- *Expressiveness issue:* the format has to be *complete* (covering all aspects), *fine-grained* (detailed) and *flexible* (easily adaptable to various use cases) enough.

# The 3 pillars of standardisation: challenges

## Challenges

### *General:*

Avoiding to uniquely or too much focusing on one aspect – Identifiers – and forgetting or underestimating the other two, because identifiers alone will not bring a full or real solution.

# The growing role of CRIS's



Current  
Research  
Information  
Systems

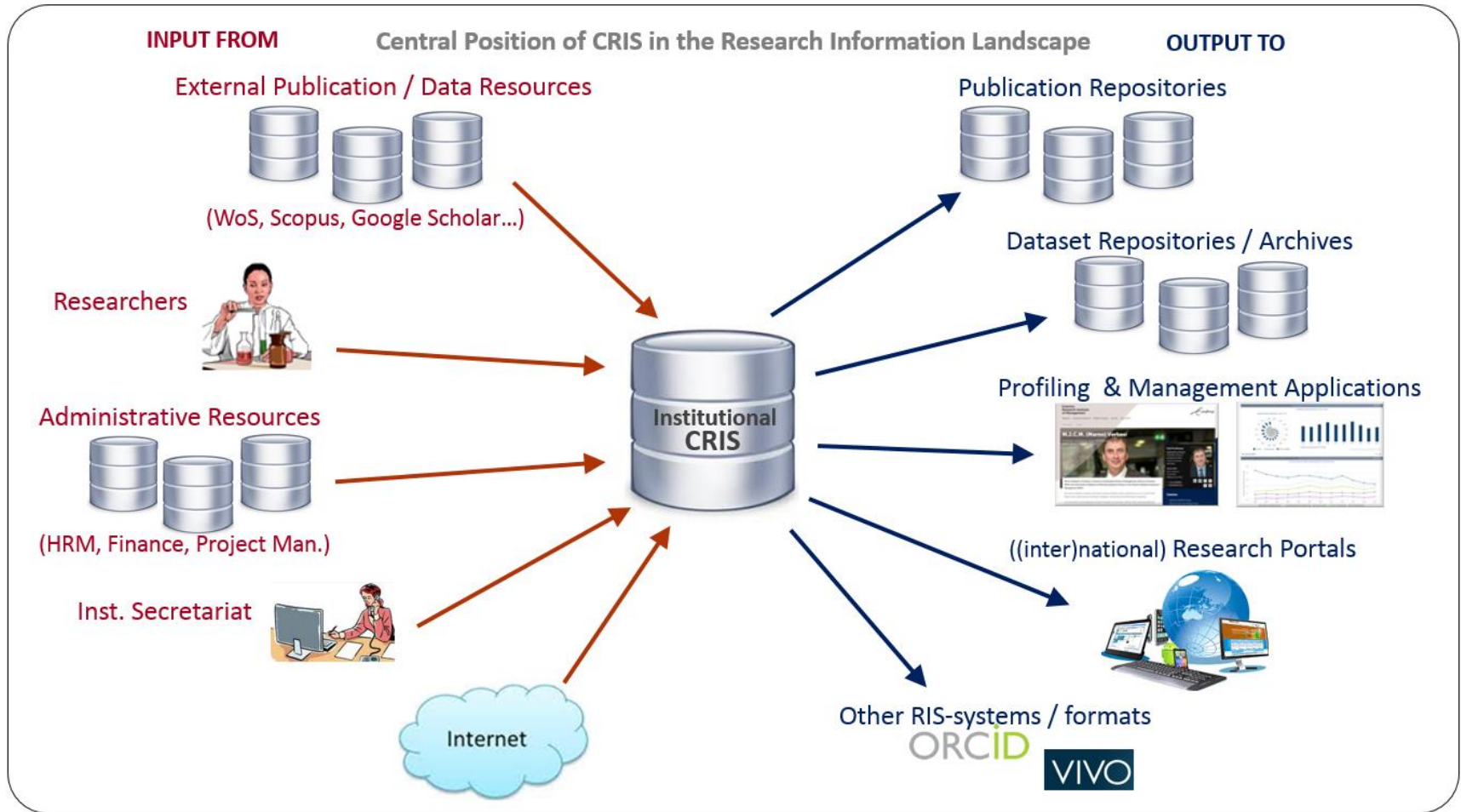
**euroCRIS**  
Current Research Information Systems

# The growing role of CRIS's

- Initially meant as administrative systems to report on research, CRIS's in the course of time have developed into multi-functional information systems covering all aspects of research and useful for all kind of stakeholders, including and more and more also the researchers (for profiling of their research online, for CV-generation and management, etc...).
- Given this development CRIS's have moved into a central position in the Research Information landscape, functioning more and more as the primary resource for research information data.

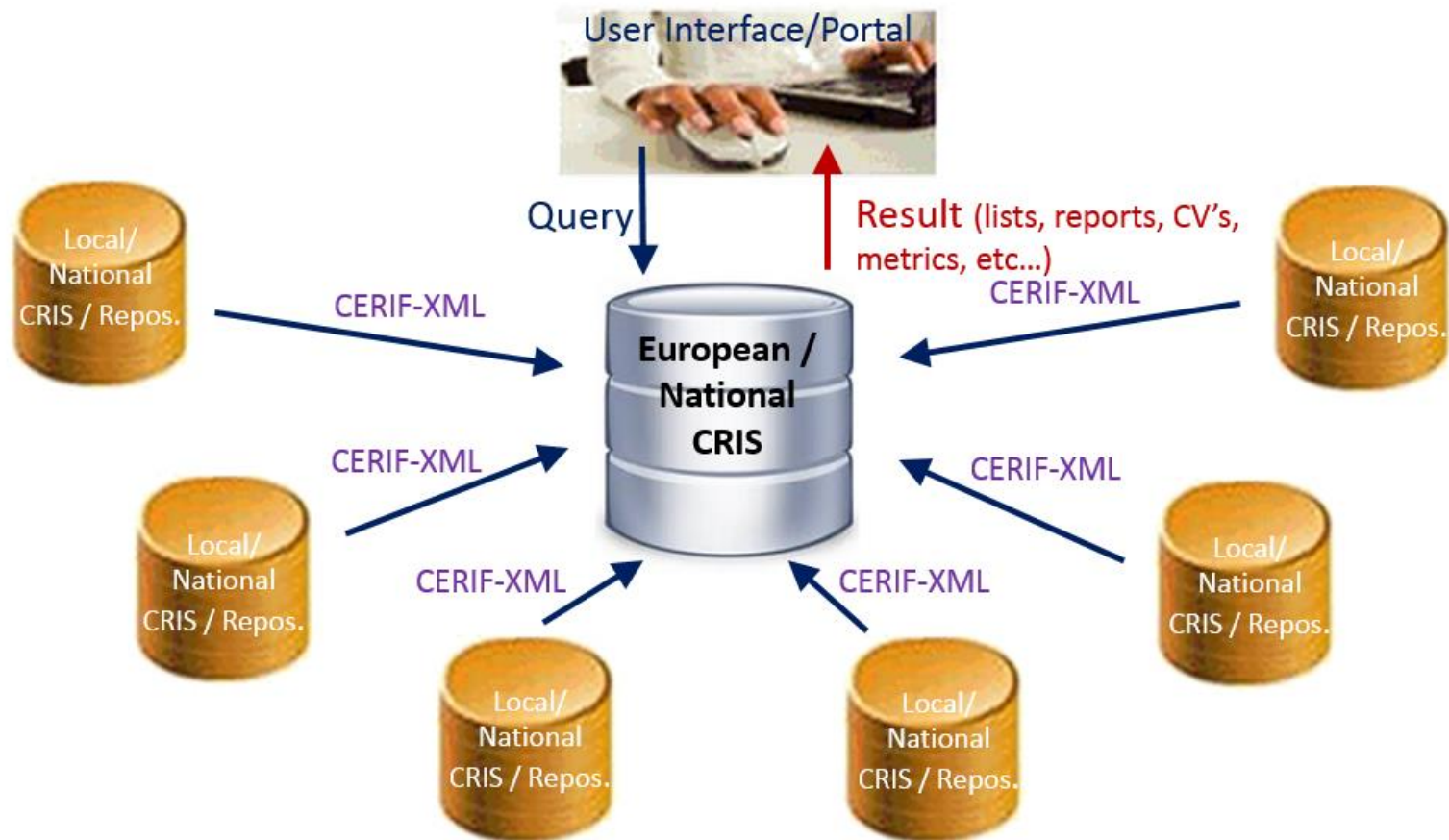
Current  
Research  
Information  
Systems

# The growing role of CRIS's: institution's perspective



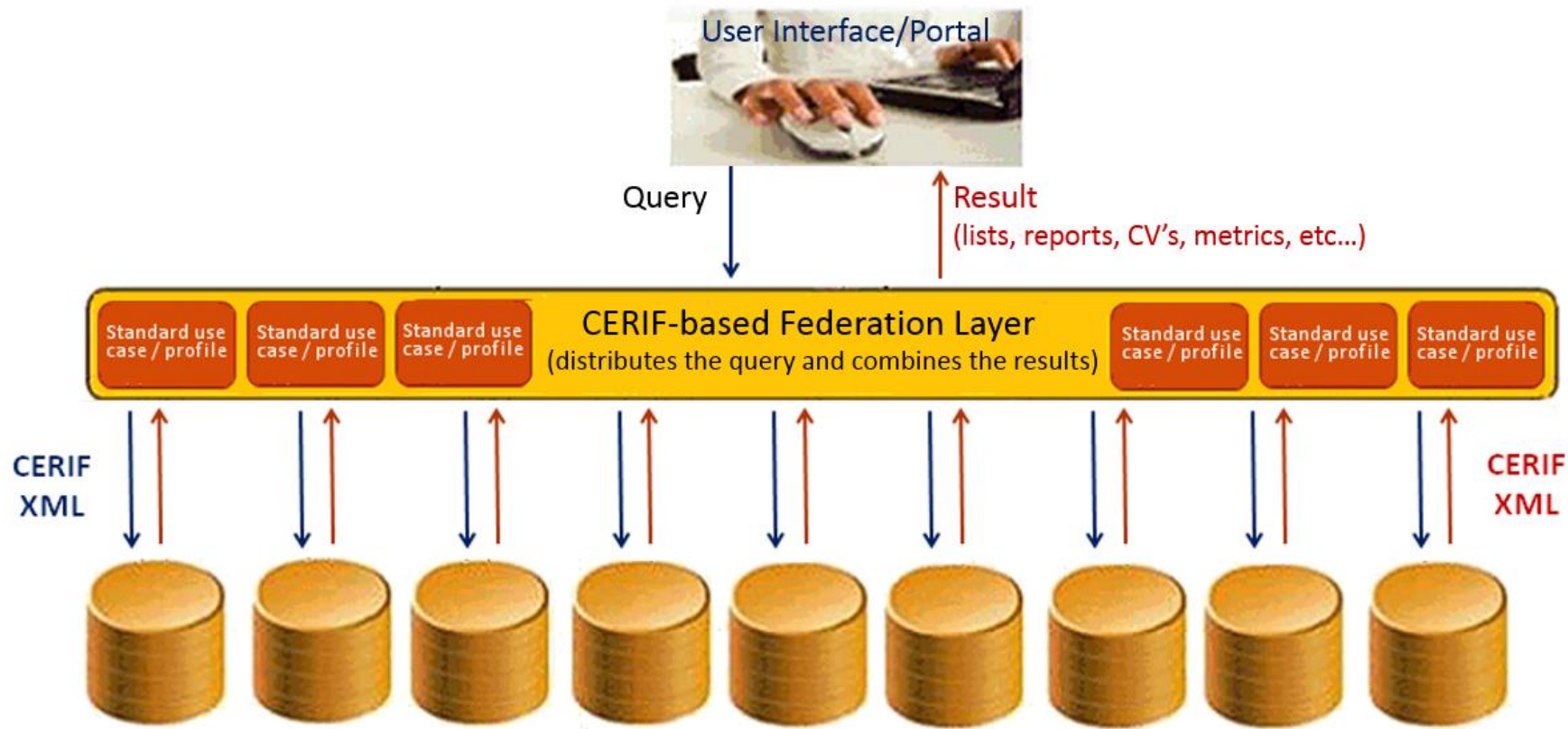
# The growing role of CRIS's: (inter)national perspective

## 1. CENTRAL SYSTEM (CRIS) SOLUTION



# The growing role of CRIS's: (inter)national perspective

## 2. FEDERATED (distributed) SOLUTION



# Peering into the future: relevance for the EOSC

- An optimal European Research Information Infrastructure, e.g. based on (the information in) CRIS's is a necessary, complementary, part of the European Open Science Cloud, indispensable to optimally realise the FAIR-aspect.
- Without such a RI infrastructure, providing the needed metadata about the research data in the Cloud, the EOSC will function suboptimally.

European  
Open  
Science  
Cloud



Thank you very much for your attention!

Questions?