Co-funded by the European Union under FP7-ICT-2009-6

APARSEN
Alliance Permanent Access to the Records of Science in Europe Network

Alliance Permanent Access to the Records of Science in Europe Network

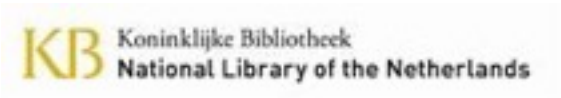
The APARSEN Interoperability Framework for Persistent Identifiers systems and added value services

WP 22 - Identifiers and citability

Maurizio Lunghi, lunghi@rinascimento-digitale.it
Scientific Director at the
Fondazione Rinascimento Digitale

APARSEN
Alliance Permanent Access to the Records of Science in Europe Network

Network of Excellence



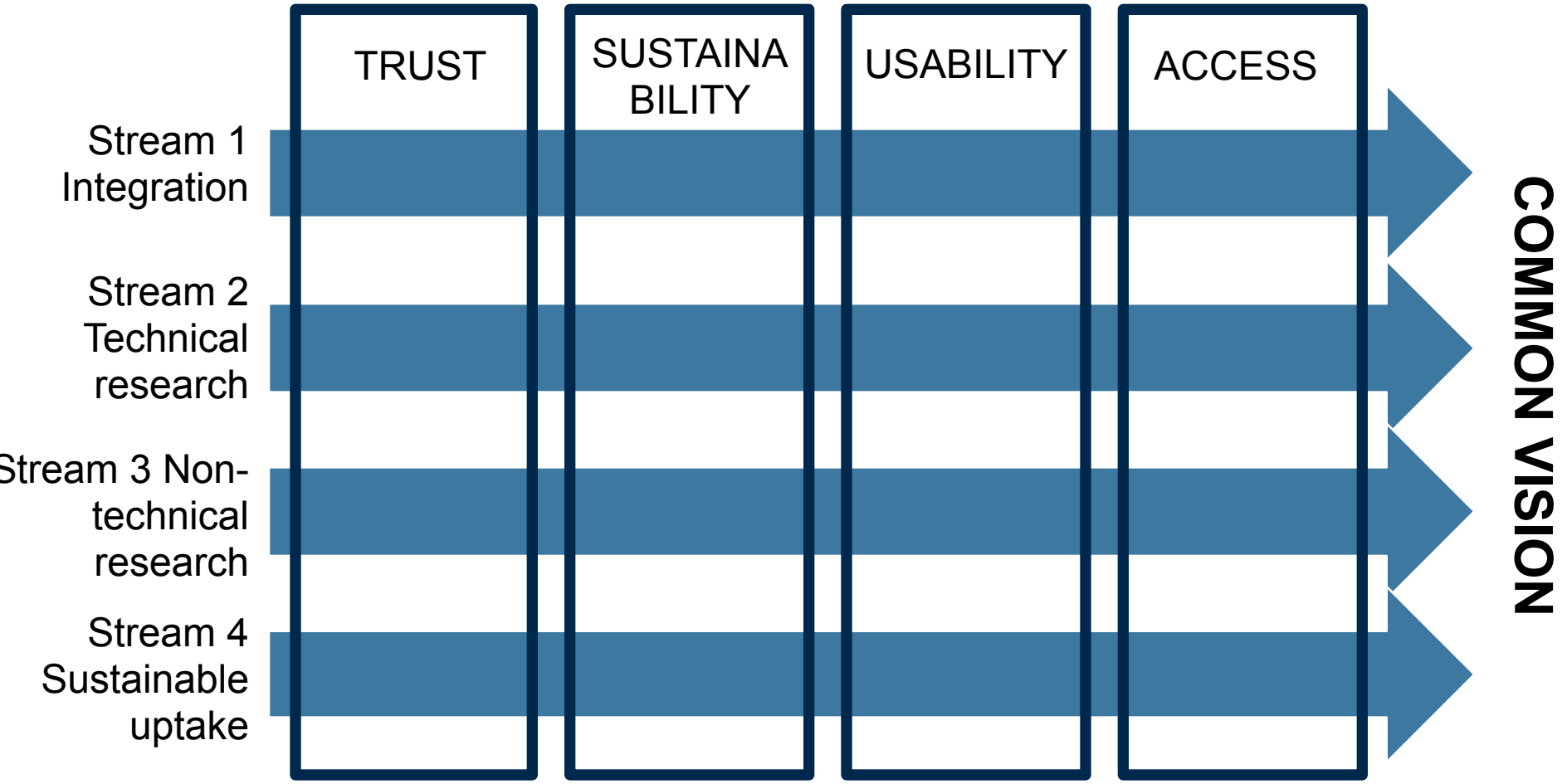
APARSEN
Alliance Permanent Access to the Records of Science in Europe Network

Virtual Centre of Excellence

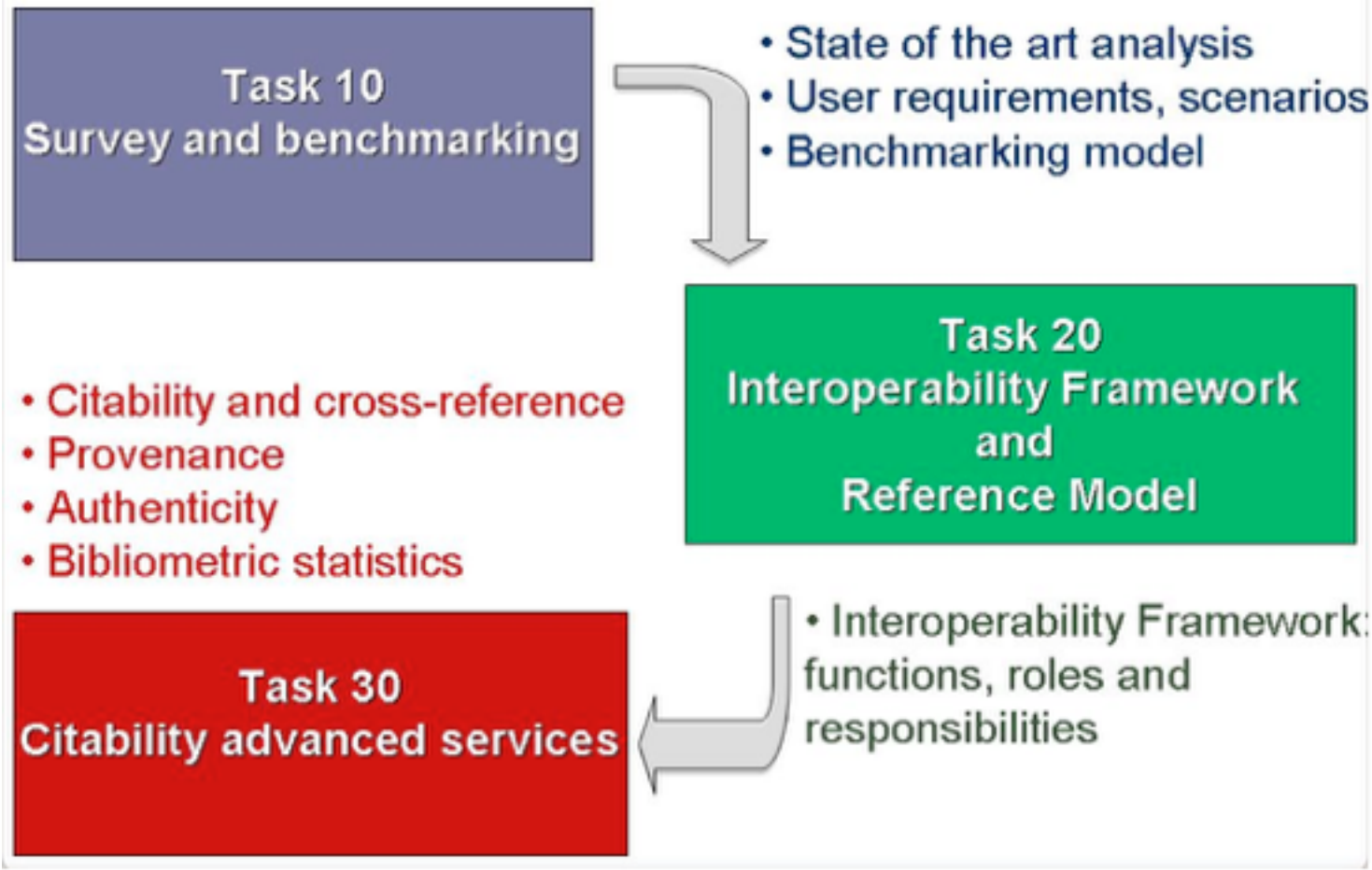
- **Defragmentation** of efforts, through positioning individual developments in a common framework + common terminology + **common vision development**
- Collected information of capabilities in terms of expertise, services, tools, training material in digital preservation http://www.alliancepermanentaccess.org/index.php/community/common-vision/
- These can be put together by a **VCoE** to provide solutions and services for digital preservation problems leading to a **Virtual Centre of Excellence** founded on a common vision of digital preservation
- **Costs** for preservation and cost models with revenues for **sustainability** of digital repositories and services

APARSEN
Alliance Permanent Access to the Records of Science in Europe Network

Approach of APARSEN



WP22: Identifiers and Citation



WP22: Identifiers and Citation - Objectives

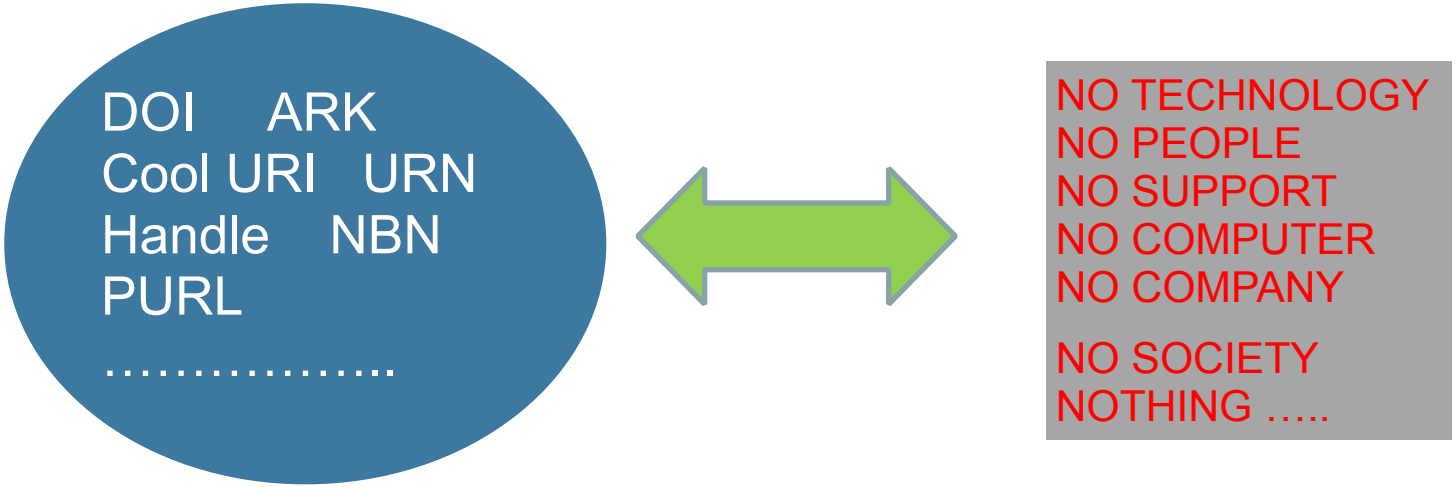
- **GOAL:** to **defragment EU initiatives on PI**
- **GOAL: not new technology** but a way to cooperate and use data from PI domains isolated, a common **format to expose data** and to develop **interoperability** and **services** tailored on user requirements
- **GOAL: community building** and experts involvement, engagement of relevant projects in the domain (PI user scenarios + HLEG + workshops + Webinars)

Persistent Identifier …. what's that ??

A PI is not only a **number**, it's a **service** based on a contract between user community or content holders and service-providers responsible for the implementation and maintenance of the PI-service. In addition to **access** for a resource we need to check other **significant properties**:

stable identification (PI) … multiple too
authenticity + integrity
provenance + author
owner + rights
+
relations with other resources & with actors

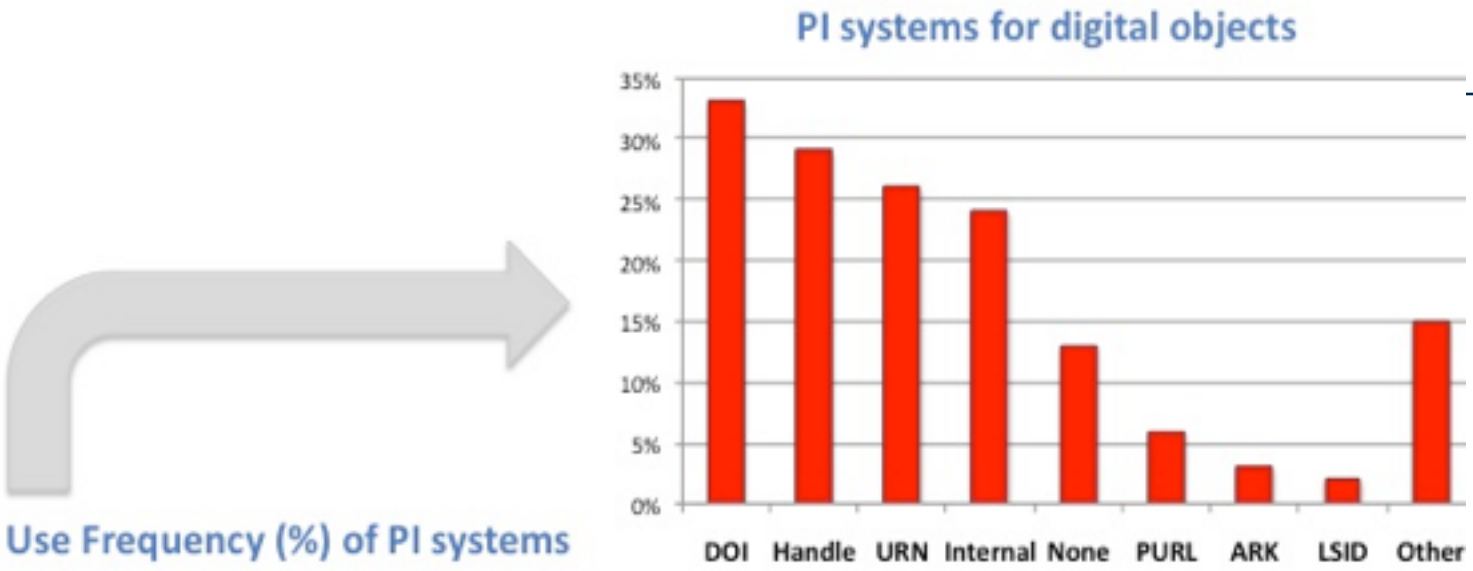
Persistent Identifier …. what's that ??

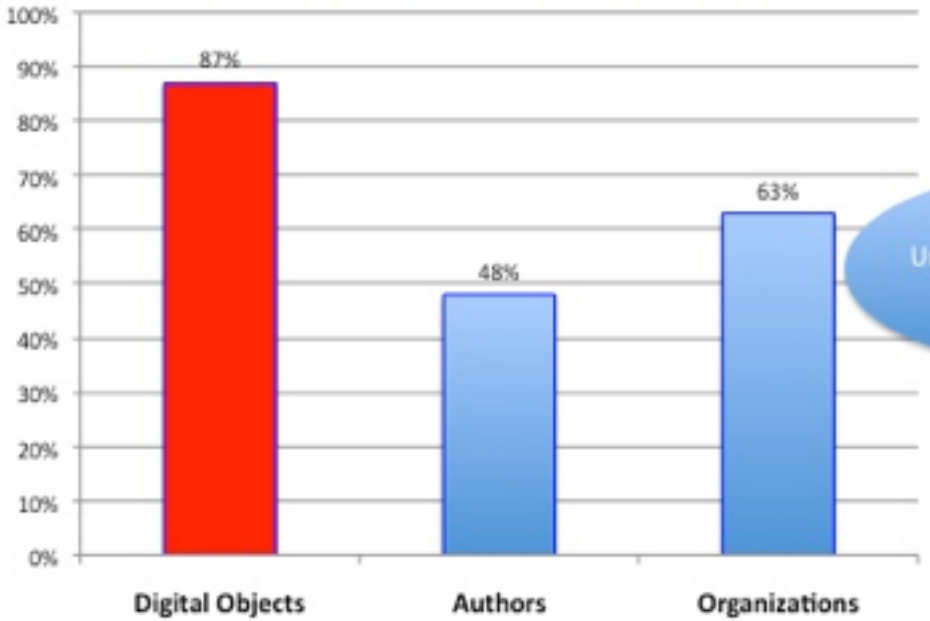


IDENTIFIERS ARE STRINGS WITH A SYNTAX
PERSISTENCY IS NOT A MATTER OF TECHNOLOGY
BUT MORE A MATTER OF POLICY AND
ORGANISATION BEHIND THE IMPLEMENTATION

Co-funded by the European Union under FP7-ICT-2009-6

PI systems for **Digital Objects**





PI SYSTEMS: NEED of TRUST

Requirements	Freq	Percent %
Cross-disciplinary	**83**	**80.5**
Managed by public/ government institution	**74**	**71.8**
Nationally not limited	**57**	**55**
Discipline-specific	10	9.7
Other	9	8.7
Nationally limited	5	4.8
Privately managed	7	6.7

Factors for Trust

Requirements

Cross-boundary systems but…

Factors contributing to the trust	Freq	Percent %
Trusted organization running the system	**74**	**71.84**
Methods of verification	**68**	**66.02**
Supported by stable funders	32	31.07
Validation by publishers	31	30.1
Author self-curation	27	26.21
Other	8	7.77
Validation by educators	7	6.8

Co-funded by the European Union under FP7-ICT-2009-6

From USER REQUIREMENTS to INTEROPERABILITY SCENARIOS

PI services Basic features

PI services advanced features

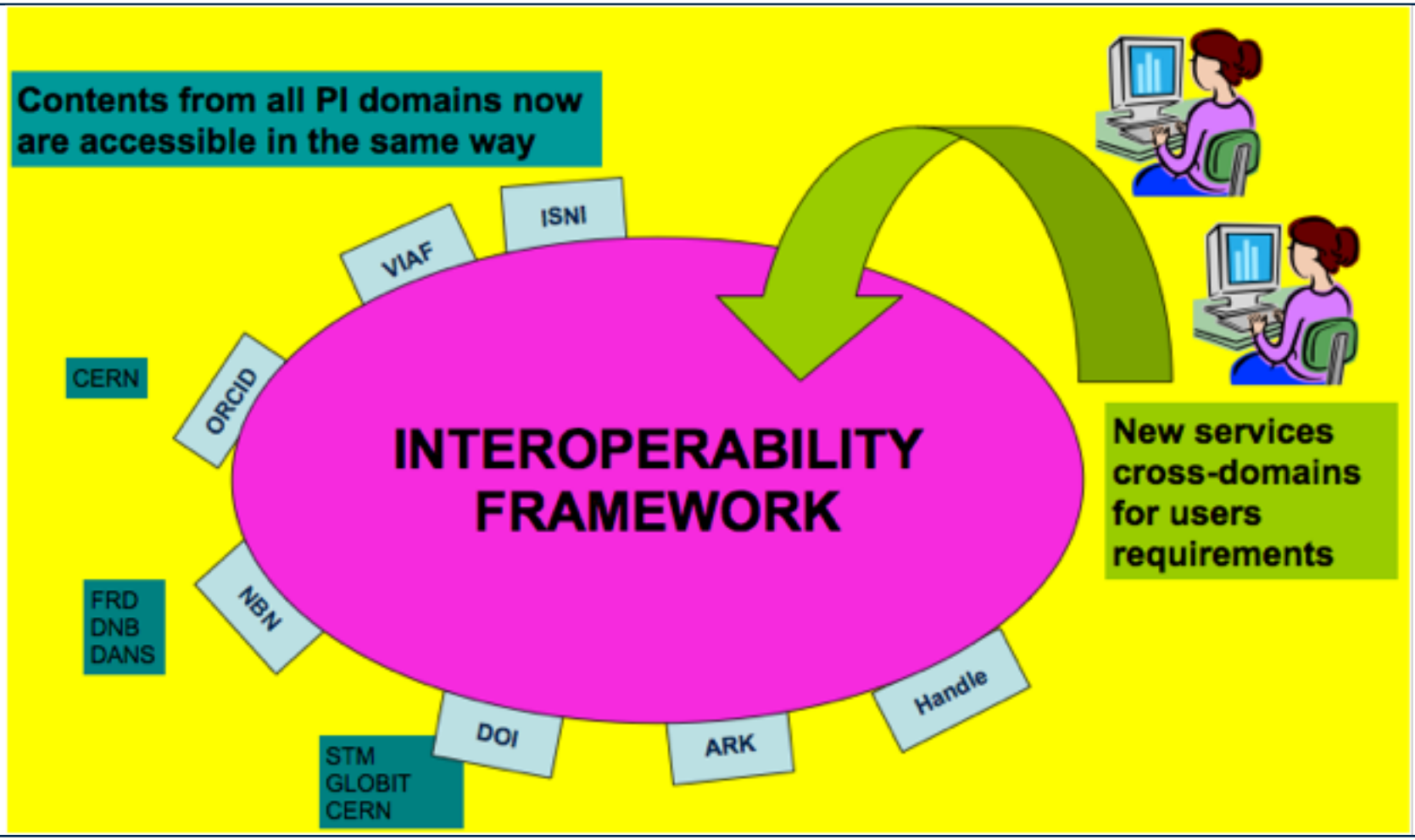
Value	Freq	Percent %
Citability	**76**	**74**
Global resolution service	**62**	**60**
PI resolution service to the resource	**57**	**55**
Digital Object certification	**55**	**53**
PI resolution service to metadata	**50**	**49**
Association of PI to multiple location (URLs)	41	40
Metrics	31	30
Multiple association name	27	26
Link digital object to dynamic dataset	19	18
Others	3	3

Co-funded by the European Union under FP7-ICT-2009-6

WP22 → HLEG on PI

Anila Angjeli	Gabriella Scipione	Samuele Carli	Andrea D'Andrea
Sébastien Peyrard	***Egbert Gramsbergen***	***Paolo Budroni***	***Mark van de Sanden***
Ernesto Damiani	***Jeroen Rombouts***	***Sabine Schrimpf***	***Antoine Isaac***
Giovanni Bergamin	***Hervé L'Hours***	***Jurgen Kett***	***Jan Brase***
Laurents Sesink	***Carlo Meghini***	***Karaca Kocer***	***Heikki Helin***
Maarten Hoogerwerf	***Claudio Cortese***	***Lars Svensson***	***Marcin Werla***
Martin Braaksma	***Yannis Tzitzikas***	***Juha Hakala***	***Aldo Gangemi***
Martin Dow	***Martin Doerr***	***Tobias Weigel***	***Oreste Signore***
David Giaretta	***Mariella Guercio***	***Piero Attanasio***	***Norman Paskin***
Alan Danskin	***Maurizio Messina***	***Andrew Treloar***	***Roberto Delle Donne***
Marcus Enders	***Laure Haak***	***Amir Aryani***	***John Kunze***

WP22: INTEROPERABILITY FRAMEWORK (IF)



Interoperability Framework (IF) basic concept

Our concept of 'interoperability' is quite simple and is not used to indicate the ability of PI systems to interoperate between them in a direct way (DOI will never speak with NBN, it's not required) but it **is conceived in terms of a common way of access to data belonging to heterogeneous PI domains which are identified through different identification schemes. Our goal is to make accessible data from all the PI domains in the same format so that users can use them without worrying about different internal organisation and policy**.

Interoperability Framework (IF): schema

Our proposal of the Interoperability Framework (IF) is based on some key concepts and criteria that have been evaluated by the HLEG on PI

a) **PI systems or domains**
b) Assumptions for the IF entities
c) Trust criteria for PI systems
d) Ontology schema for the IF entities

Persistent Identifier systems

PI technologies help make stable reference to digital resources, even if it is well-known that **persistency** is not only a technical issue, no technology can exist indefinitely or guarantee services without a trusted organization behind and a clearly defined policy.

PI systems are meant as:

a) reliable **technology**

b) trusted **organization**

c) precise **policies** for digital preservation

Interoperability Framework (IF): assumptions

MAIN ASSUMPTIONS:

1. In the IF we consider only entities identified by at least one PI
2. Only PI domains that meet some criteria are eligible to be considered in the IF: trusted PI domains
3. We delegate the responsibility to define relations among resources and actors to the trusted PI domains
4. We don't address digital preservation policy but delegate that to the trusted PI domains

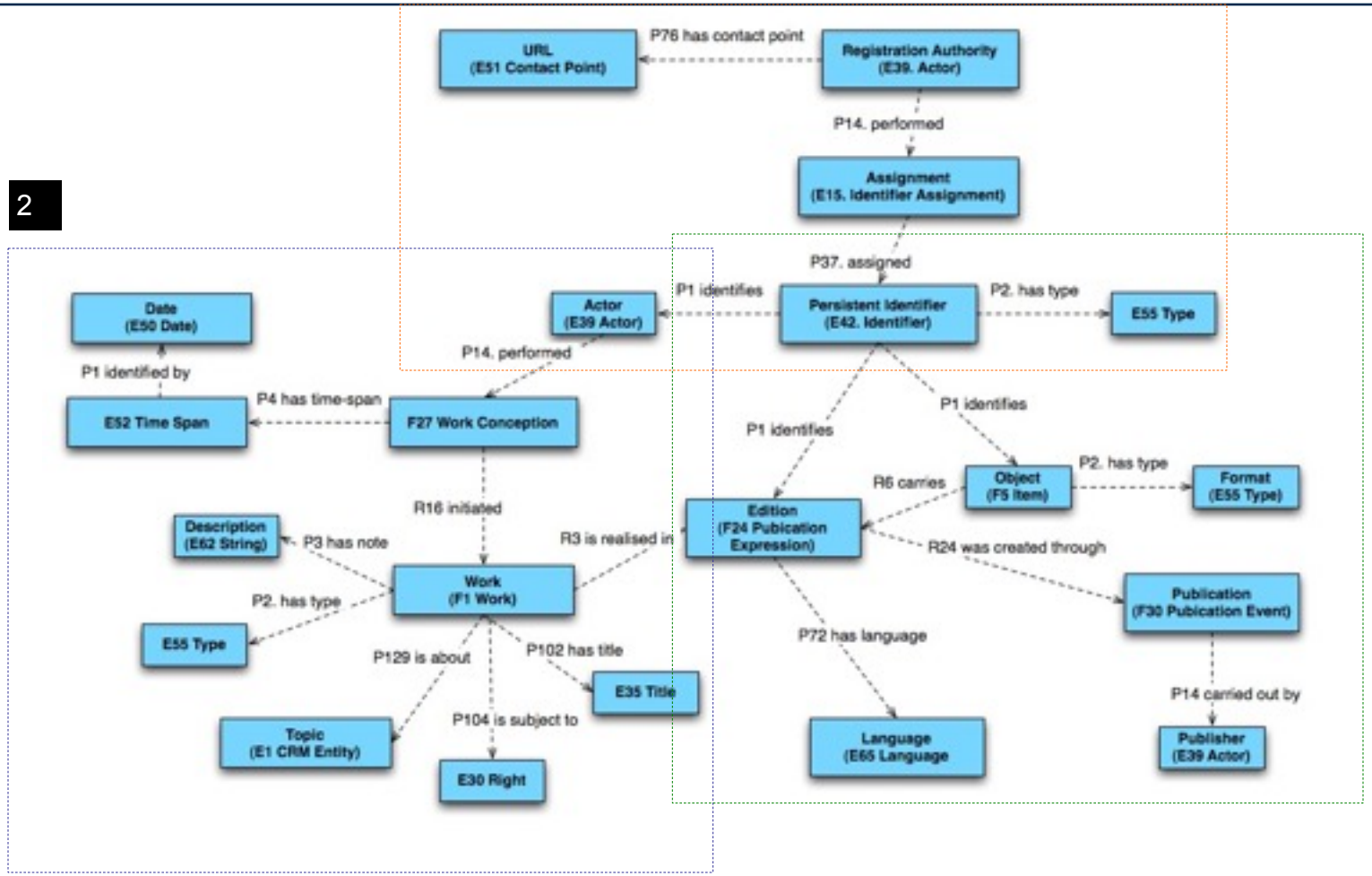
WP22: Trusted PI systems: criteria

1. **Having at least one Registration Agency.**
2. **Having one Resolver accessible on the Internet.**
3. **Uniqueness of the assigned PIs within the PI domain and so also globally unique.**
4. **Guaranteeing the persistence of the assigned PIs.**
5. **User communities of the PID should implement policies for digital preservation (e.g. trusted digital repositories)**
6. **Reliable resolution.**
7. **Uncoupling the PIs from the resolver.**
8. **Managing the relations between the PIs within the domain.**

Scenarios Integration

1

2



3

Example of atomic services – PI-Alternative PI

- **Input***: PI for objects*
- **Output***: PIs list*
- **Description***:* The PI-alternative PIs service retrieve alternative PIs of a given PI exploiting the co-reference relations among PIs exposed in the Interoperability Framework presenting them with the same level of service.

http://93.63.166.138/demonstrator/demo7

Example of cluster service – Scientific carreer assessment

- **Input:** Author PI
- **Output:** list of publications (metadata + PI) filtered
- **Description:** This service retrieves all the objects associated to the author PI satisfying the search criteria from all PI domains and presents them in a common format with the same metadata schema. The final output is the list of publications of the author, which responds to the specified criteria.

http://93.63.166.138/demonstrator/demo7

INTEROPERABILITY FRAMEWORK (IF)

PI domain for digital objects PI-dg

PI domain for actors PI-ac

PI domain for physical objects PI-ph

PI domain for bodies PI-bd

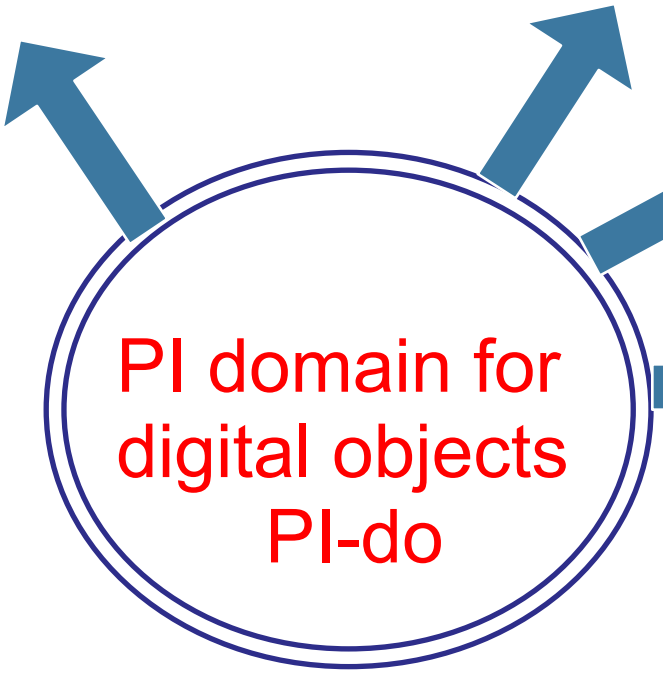
PI domain for …… PI-...

INTEROPERABILITY FRAMEWORK (IF)

Basic info about the resource associated to the original PI-do and its simple description

Provide other PI-do associated to the same resource

Provide PI-do of other resources related to the resource associated to the original PI-do explaining the relation between the resources



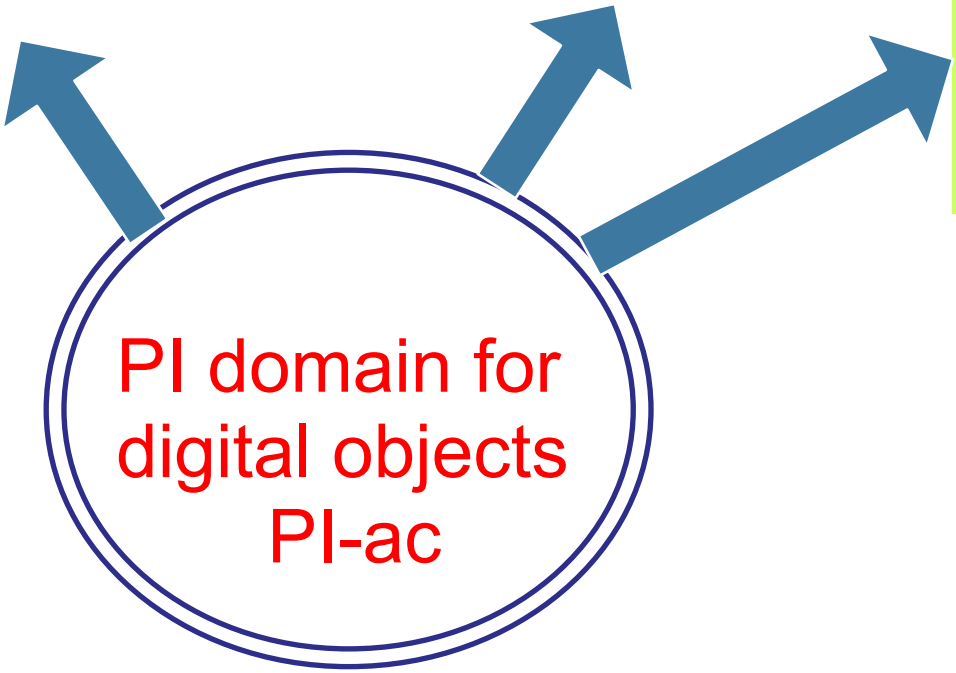
Provide PI-ac of actors associated to the resource related to the PI-do and explaining the role of the actors

INTEROPERABILITY FRAMEWORK (IF)

Basic info about the actor associated to the original PI-ac

Provide other PI-ac associated to the same actor

Provide PI-do of any resource related to the actor associated to the original PI-ac explaining the relation between the resources and the role of the actor



Co-funded by the European Union under FP7-ICT-2009-6

WP22: INTEROPERABILITY FRAMEWORK (IF)

http://93.63.166.138/demonstrator/demo7

4 types of PI systems
1. **PI-do**
2. PI-po
3. PI-bd
4. **PI-ac**

IF is suitable for any type of PI systems
4 assumptions + 8 trust criteria

IF compliant

demonstrator

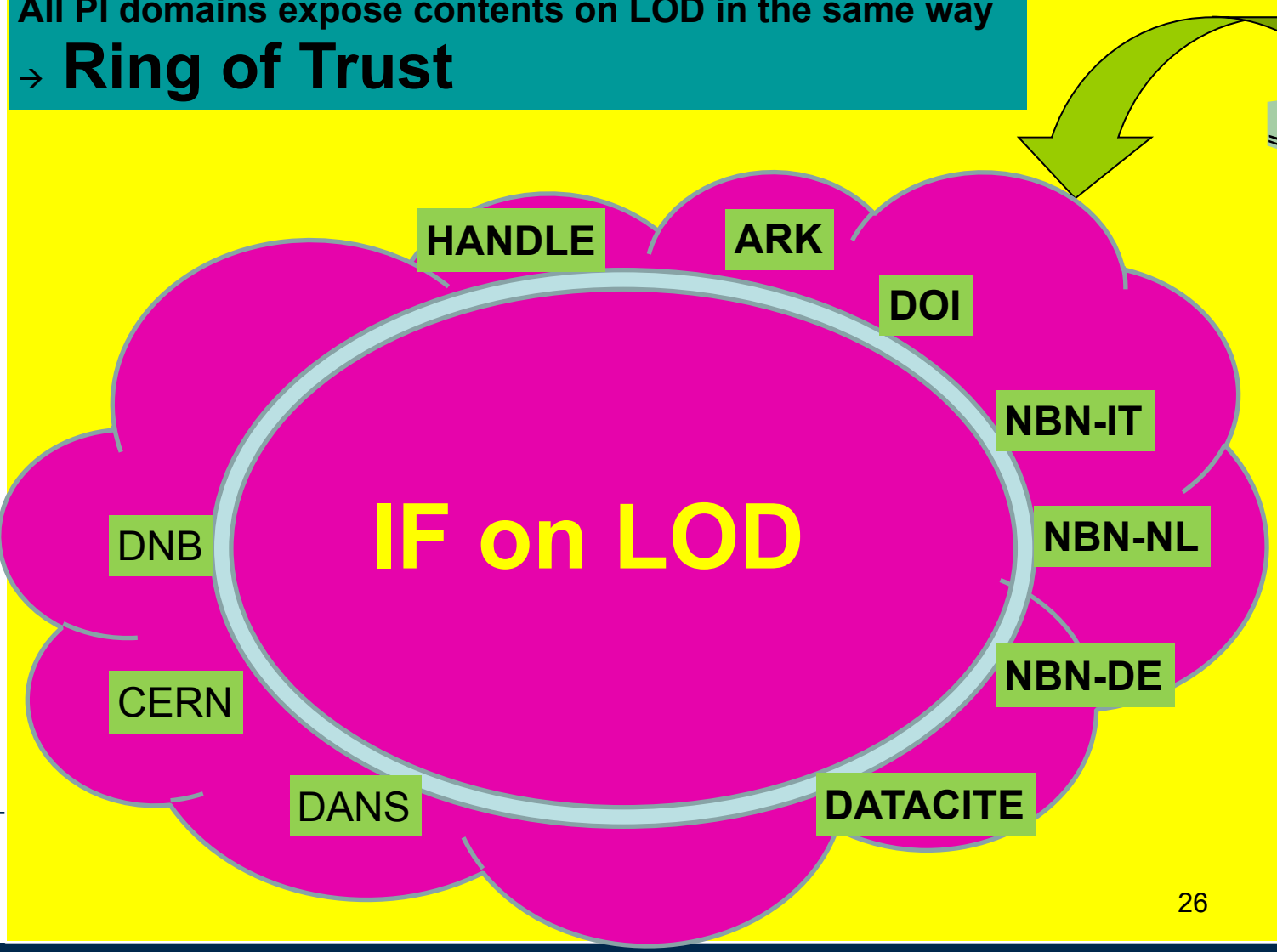
Focus only on PI and related data avoiding metadata describing the content

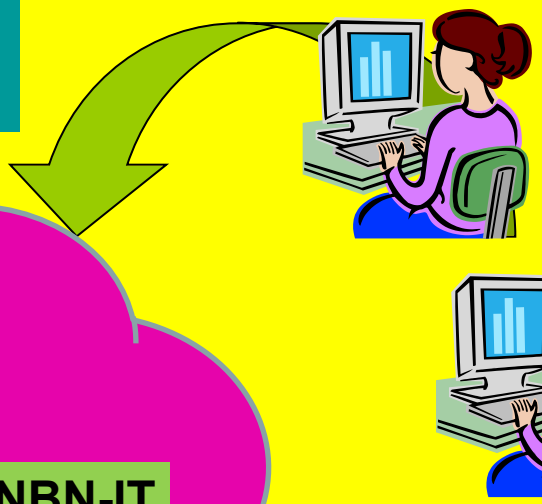
Distributed architecture
3 SPARQL end-points
6 content providers
2 service providers

INTEROPERABILITY FRAMEWORK (IF)

All PI domains expose contents on LOD in the same way

→ Ring of Trust





New services cross-domains for users requirements