

PV 2009
Ensuring Long-Term Preservation and Adding Value to
Scientific and Technical Data
1-3 December 2009

The CASPAR Finding Aids

Henri Avancini, **Carlo Meghini** &
Loredana Versienti
CNR ISTI, Pisa



Outline

- The CASPAR Project
- Requirements
- Conceptual model
- Architecture
- FIND in CASPAR
- Conclusions

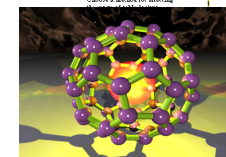
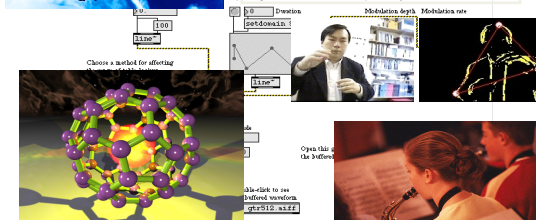
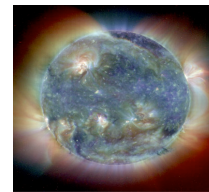


The CASPAR Project

- The CASPAR project is mainly based on standard **ISO: 14721:2003 OAIS**
- In this perspective, its Architecture is defined for
 - Managing key concepts of the **OAIS reference model**
 - Supporting main functionality identified in the **OAIS functional model**
- Moreover, the CASPAR project aims to define, and implement, interfaces and functionally **independent components**

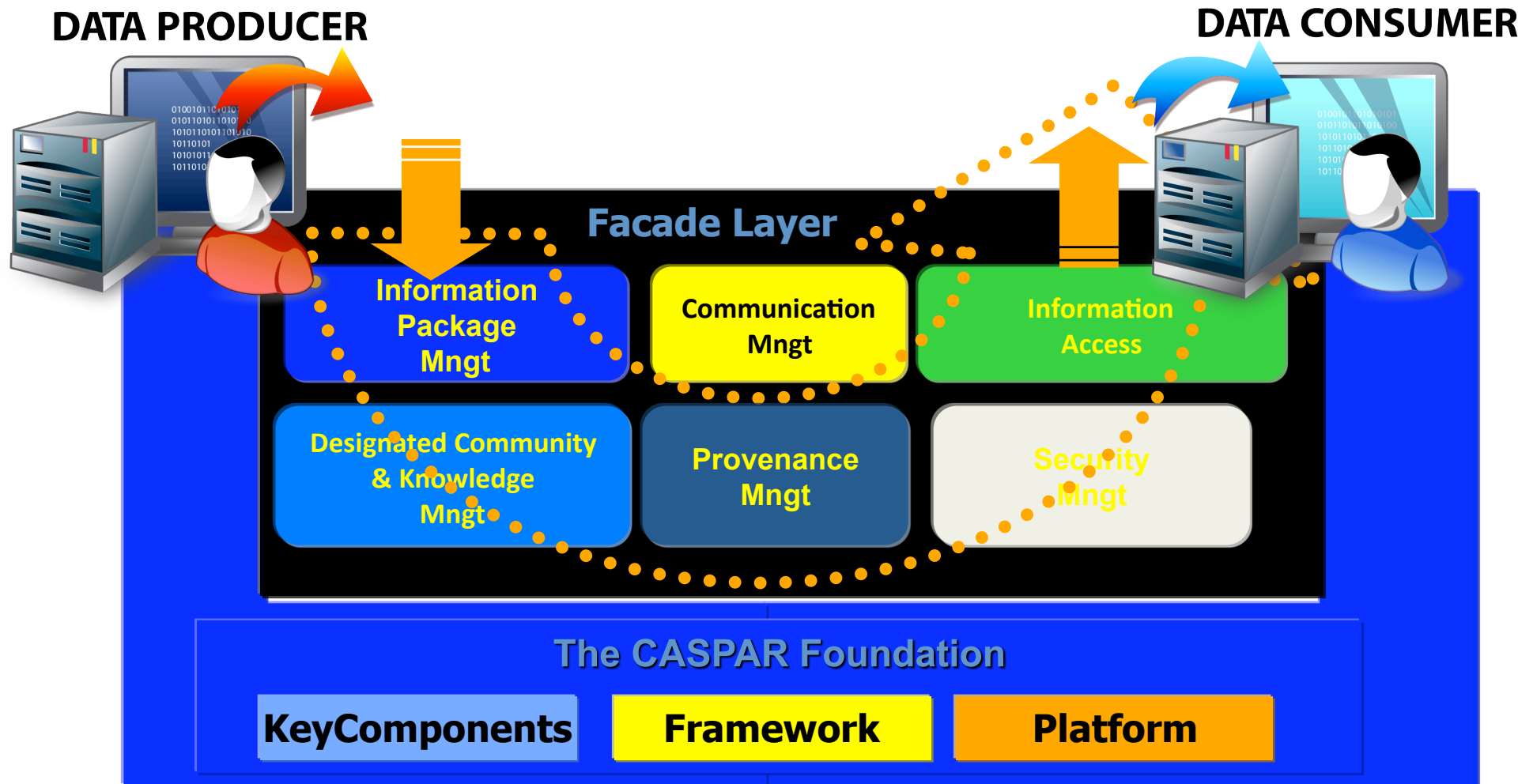


The Consortium

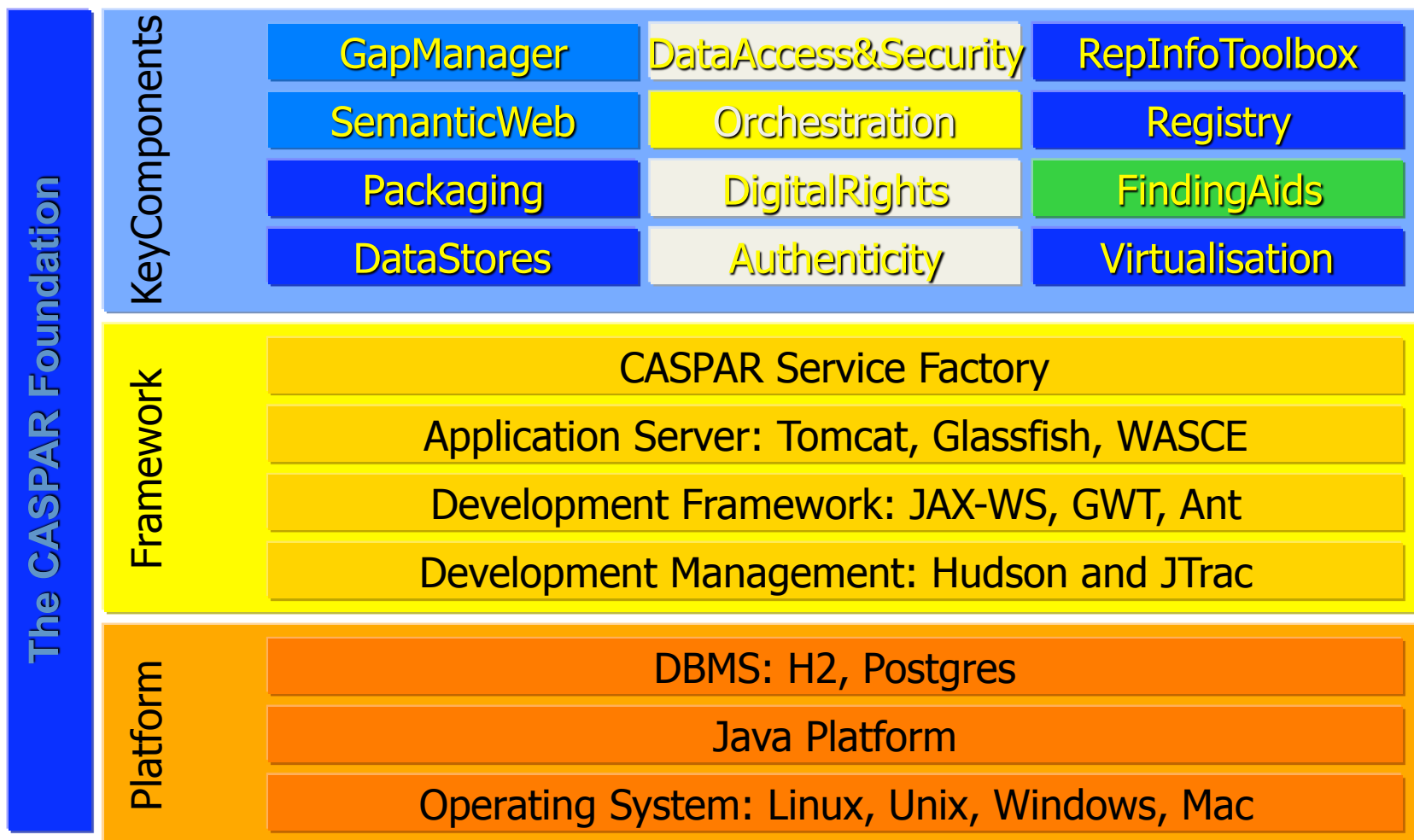


ISTITUTO DI SCIENZA E TECNOLOGIE DELL'INFORMAZIONE "A. FAEDO"

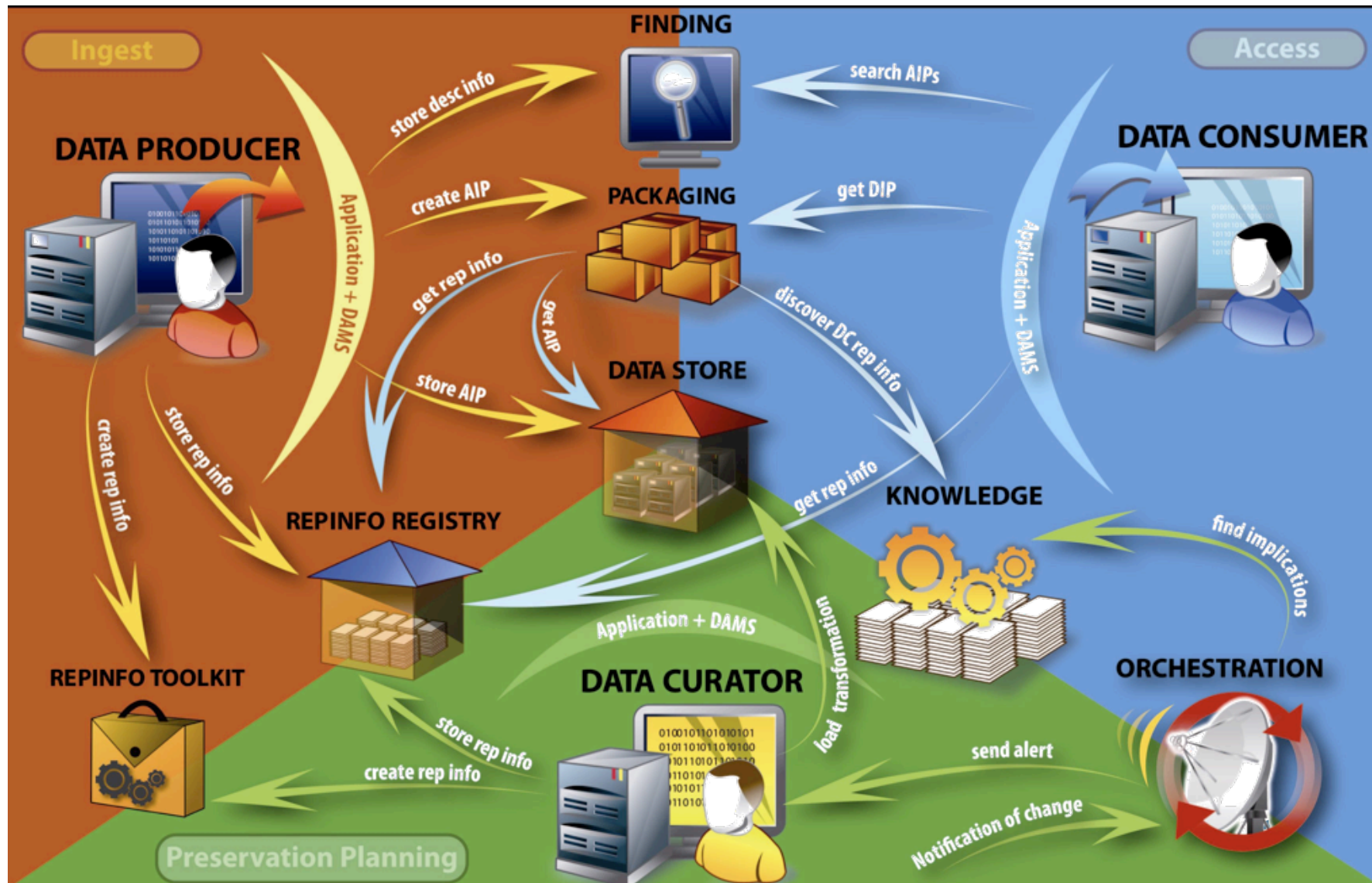
The CASPAR Solution



The CASPAR Architecture



The CASPAR Workflow



Outline

- The CASPAR Project
- **Requirements**
- Conceptual model
- Architecture
- FIND in CASPAR
- Conclusions



Requirements

- Maximize usability
 - included the archives which would like to enhance the finding aids that are already in place.
- Independence from data languages
 - Data Definition Language
 - Data Manipulation Language (including Query Language)
- Expressivity of the language for representing Description Information
- Adherence to standards for wide adoption and long lifetime



Outline

- The CASPAR Project
- Requirements
- **Conceptual model**
- Architecture
- FIND in CASPAR
- Conclusions



The FIND conceptual schema

- The CASPAR Finding Aids is a CASPAR key component that provides the Data Management functionality of the OAIS Reference Model (discovery of AIPs).
- The FA is based on two basic components:
 - Finding Registry, and
 - Finding Manager.



Finding Manager

- *A Finding Manager* supports the management of Description Information, and is bound to a language for defining and for querying DescInfo.
 - A Finding Manager may talk (relational + SQL)
 - another one (RDF + SPARQL)
 - another one (XML + Xquery)
- Every Finding Manager registers with at least a Finding Registry in order to be discovered by applications.



Finding Manager

A Finding Manager supports two main functionalities:

- Management of DescInfo:
 - At the schema level:
 - Create
 - delete
 - browse DescInfo schema elements (i.e., tables or classes or DDTs).
 - At the object level:
 - Create
 - Delete
 - Update
 - browse DescInfo objects (i.e., tuples or objects or documents).

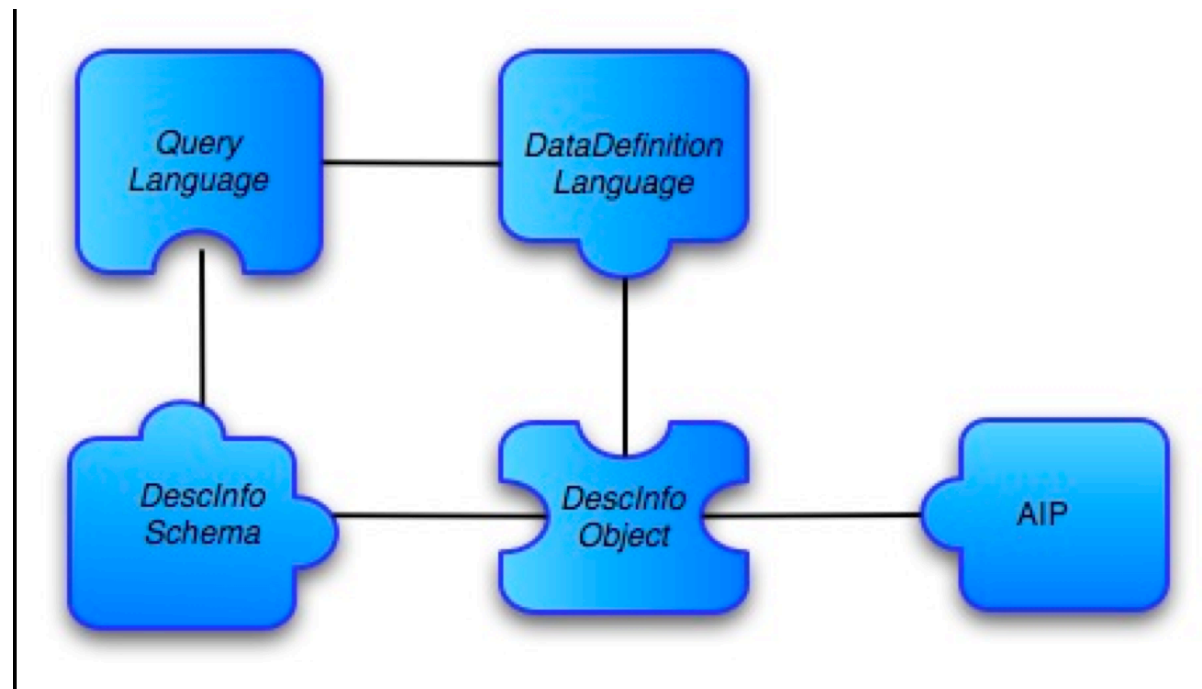


Finding Manager

- Management of the association between DescInfo objects and AIP identifiers, including usage of these associations for AIP discovery:
 - Create
 - Delete
 - Query
 - Browse (AIP-id, DescInfo-id) pairs.
 - Discovery of AIPs via queries on DescInfo objects.



Finding Manager concepts



Finding Manager

- A Finding Manager registers with a Finding Registry by providing a description of itself to the Registry.
- This description contains required information, such as:
 - (Data definition & query) language spoken by the Finding Manager.
 - Handle for invoking the Finding Manager.
 - Additionally, information concerning properties of the Finding Manager that applications consider useful for discovery purposes.



Finding Registry

A *Finding Registry* supports the publication and discovery of Finding Managers,

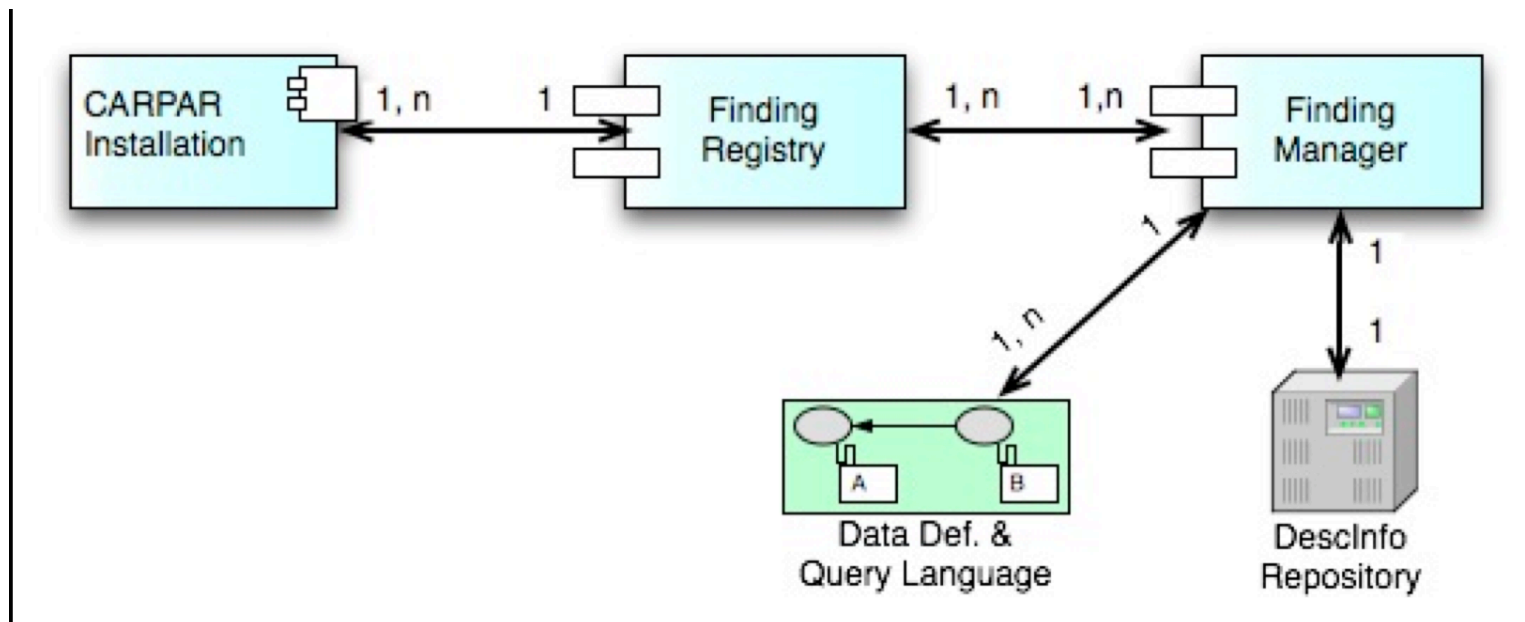
- in the same way a UDDI server supports the publication and discovery of Web Services.

Functionally, a Finding Registry supports two main functionalities:

1. Management of Finding Managers, i.e:
 - Registration
 - Deregistration
 - Discovery
 - Browse
 - Access
2. Indexing and retrieval of all the Description Information objects owned by the Finding Managers registered with the Finding Registry.



The global picture



Outline

- The CASPAR Project
- Requirements
- Conceptual model
- **Architecture**
- FIND in CASPAR
- Conclusions

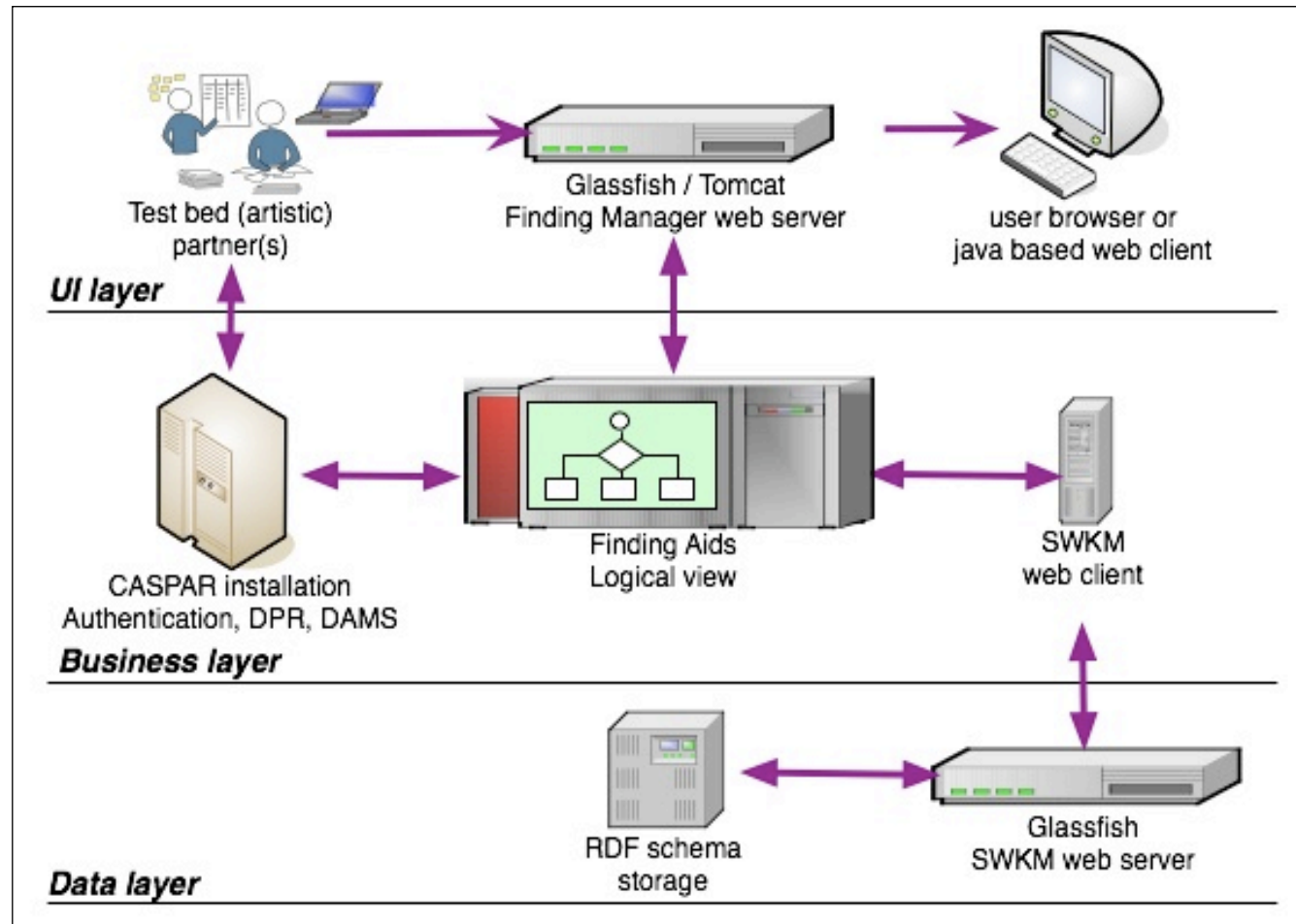


Architecture

- For the needs of the CASPAR project, we implemented a Semantic-Web based Finding Manager:
 - Spoken data language: RDF
 - Query Language: SPARQL
 - Platform: RDF Suite (implemented at FORTH)



Architecture of FIND



Outline

- The CASPAR Project
- Requirements
- Conceptual model
- Architecture
- **FIND in CASPAR**
- Conclusions



FIND in CASPAR

Community	DescInfo Schemas	Schema links
ESA (Scientific)	Scientific schema (ad-hoc RDF schema)	http://rdfs.esrin.esa.int/EGOC.rdfs#
IRCAM (Artistic)	CIDOC CRM FRBR extension	http://cidoc.ics.forth.gr/rdfs/CIDOC4.3.rdfs# http://cidoc.ics.forth.gr/rdfs/caspar/frbr.rdfs#
UNESCO (Cultural)	CIDOC-CRM Extension for ESRI ASCII Grid data objects CIDOC Extension for UNESCO automatically generated from XML	http://www.casparpreserves.eu/testbed/cultural/esrigrd http://www.casparpreserves.eu/testbed/cultural/ewe/epdl
Univ. of Leeds (Artistic)	CIDOC CRM FRBR extension	http://cidoc.ics.forth.gr/rdfs/CIDOC4.3.rdfs# http://cidoc.ics.forth.gr/rdfs/caspar/frbr.rdfs#



Outline

- The CASPAR Project
- Requirements
- Conceptual model
- Architecture
- FIND in CASPAR
- **Conclusions**



Conclusions

- A simple yet powerful Finding Aids
- Semantic web languages are good for preservation too
 - Allow for rich schemas to be used in different places of the OAIS RM
 - Extensible
- Can build on existing standard and technologies



Thank you!

- Questions?

