



# Distributed metadata catalogues

## Theory vs Reality

Ioannis Kanellopoulos, Michel Millot, Lars Bernard

European Commission  
Joint Research Centre

Institute for Environment and Sustainability

Kristian Senkler, Uwe Voges, Udo Einspanier, Jens Voigt

con terra GmbH



# JRC - Access to Distributed Catalogues Project

## Why?

- Experience with Distributed metadata Catalogues
  - Search and discovery of geographic resources
    - Allow users to indicate what type of geographic information or services they look for, using the criteria referred to in the INSPIRE Directive proposal
  - Assess state of the art in metadata catalogue technologies and “test drive” use of standards (ISO 19115/19119/19139, OGC CS 2)
  - Support the development of the EU-Geoportal (Art. 21)
  - Provide feedback & test platform for Metadata and Network Services drafting teams



## Activities related to CS

- JRC – ESA workshop on catalogue services (Oct 2005)
- ESA – FAO – JRC discovery services workshop (April 2006)
- Workshops planned to support the drafting of the Network Services IR:
  - *“Service Chaining and Workflow management – state of art” (April 19/20 2006 Vienna)*
  - *“Geodata Download – state of art” (June 20 2006 Innsbruck)*
  - ***“Resource Discovery and View – validation” (linked with NS DT progress)***



# Functional requirements

## What?

- Access distributed metadata catalogues,
  - distributed catalogues themselves (cascading catalogues).
  - interfaces compliant with the OGC catalogue service specification
  - Standard protocol bindings are both HTTP (SOAP/XML, POST/GET), according to OGC CSW 2.0 and SRW (Search/Retrieve Web Service), as a web service implementation (SOAP/XML) of Z39.50.
  - Real time search
  - Harvesting



- January – June 2006
- System installed and currently in testing phase
  - implementation of an ISO profile of the discovery metadata (based on art. 18 of INSPIRE proposal for directive and current version of the metadata DT)



# Project Partners Who?

- BRGM Catalogue of Geoscience data and services
- Geomis.Portal
- IDEE Catalogue Spain
- Norway geodata catalogue
- Netherlands Catalogue
- IDEC Catalogue
- Geocatalog (NRW)
- EuroMapFinder (Eurogeographics)
- ESA
- JRC Catalog

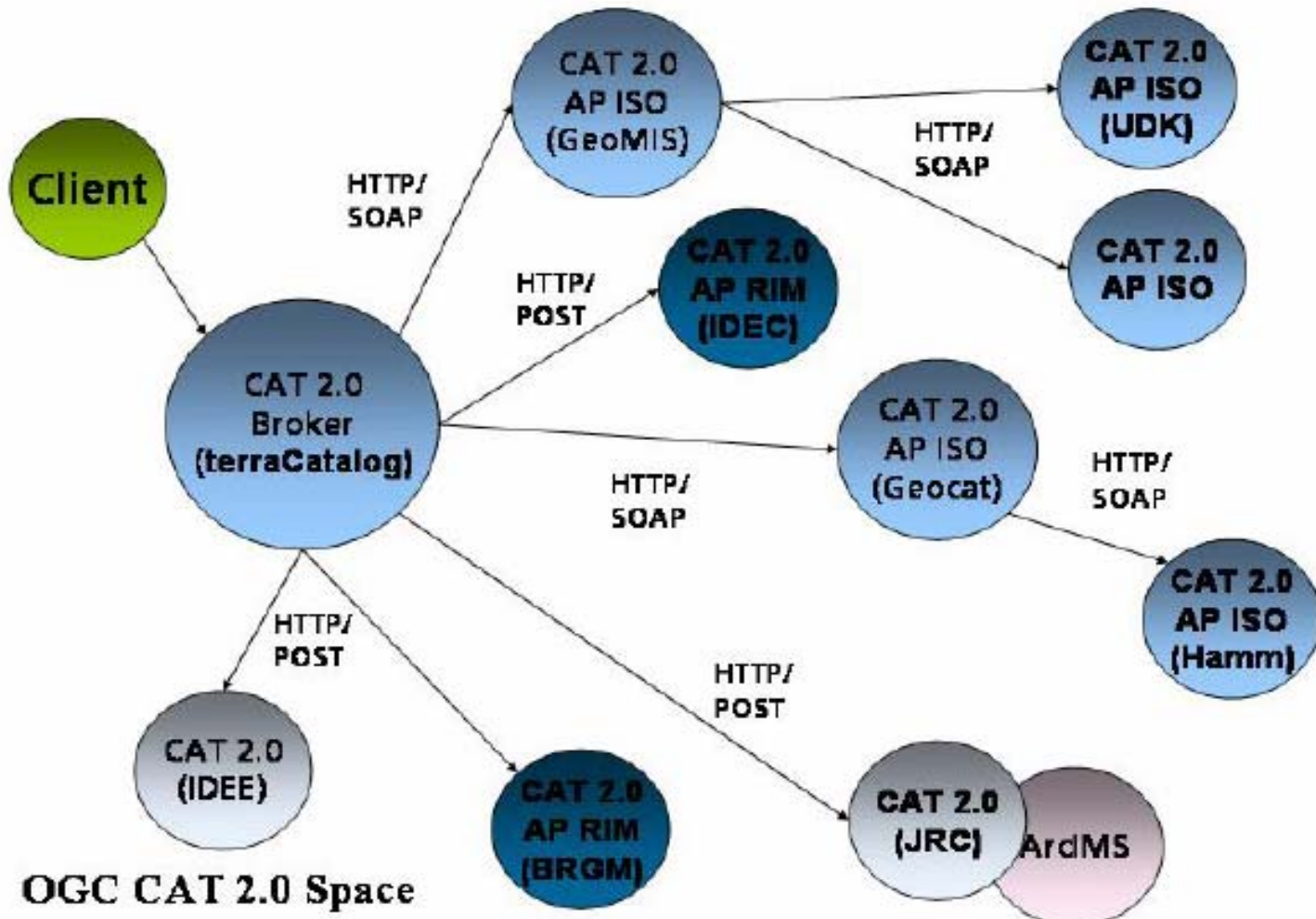


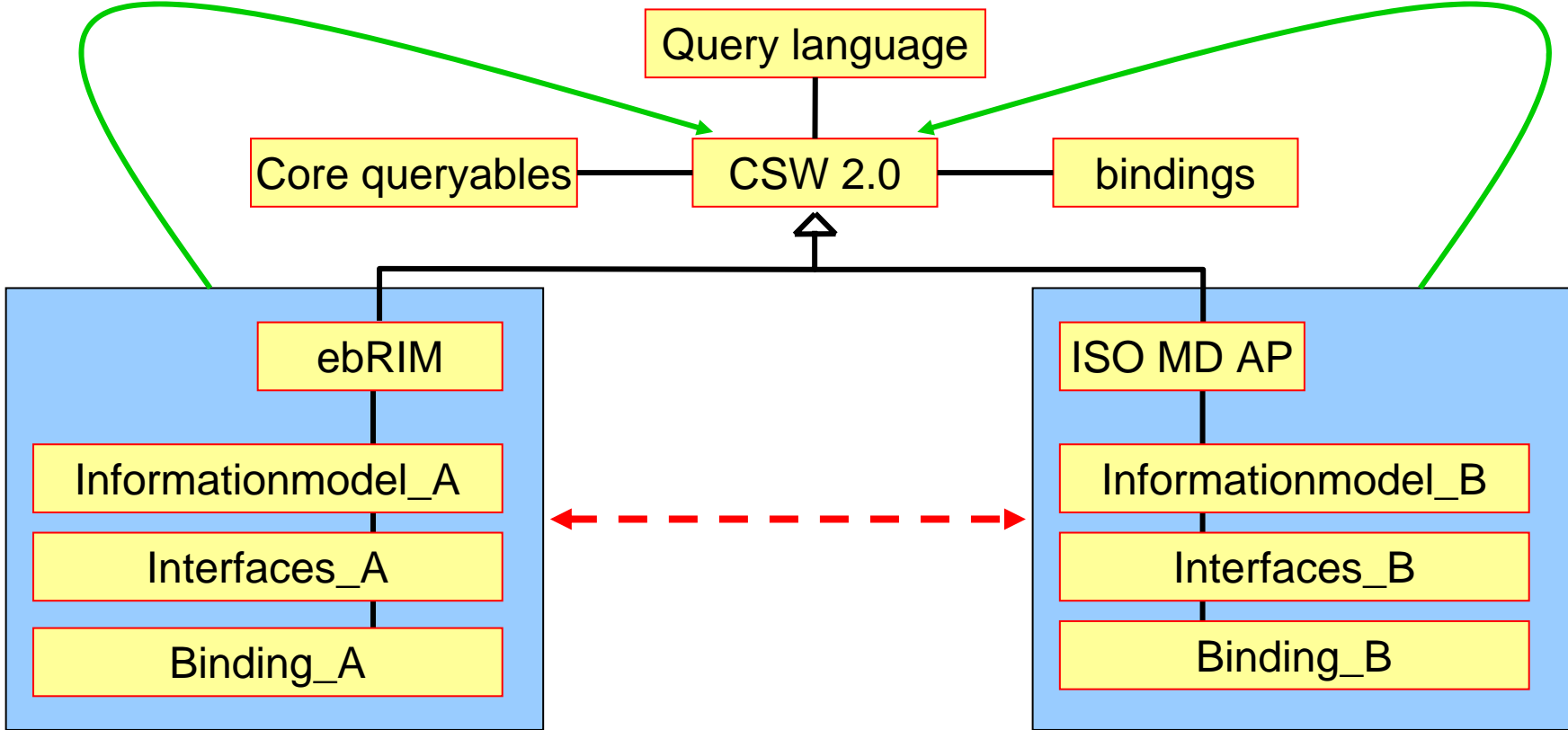
# Project Partners Who?

- BRGM Catalogue of Geoscience data and services
- Geomis.Portal
- IDEE Catalogue Spain
- Norway geodata catalogue
- Netherlands Catalogue
- IDEC Catalogue
- Geocatalog (NRW)
- EuroMapFinder (Eurogeographics)
- ESA
- JRC Catalog



# Bindings and Application Profiles

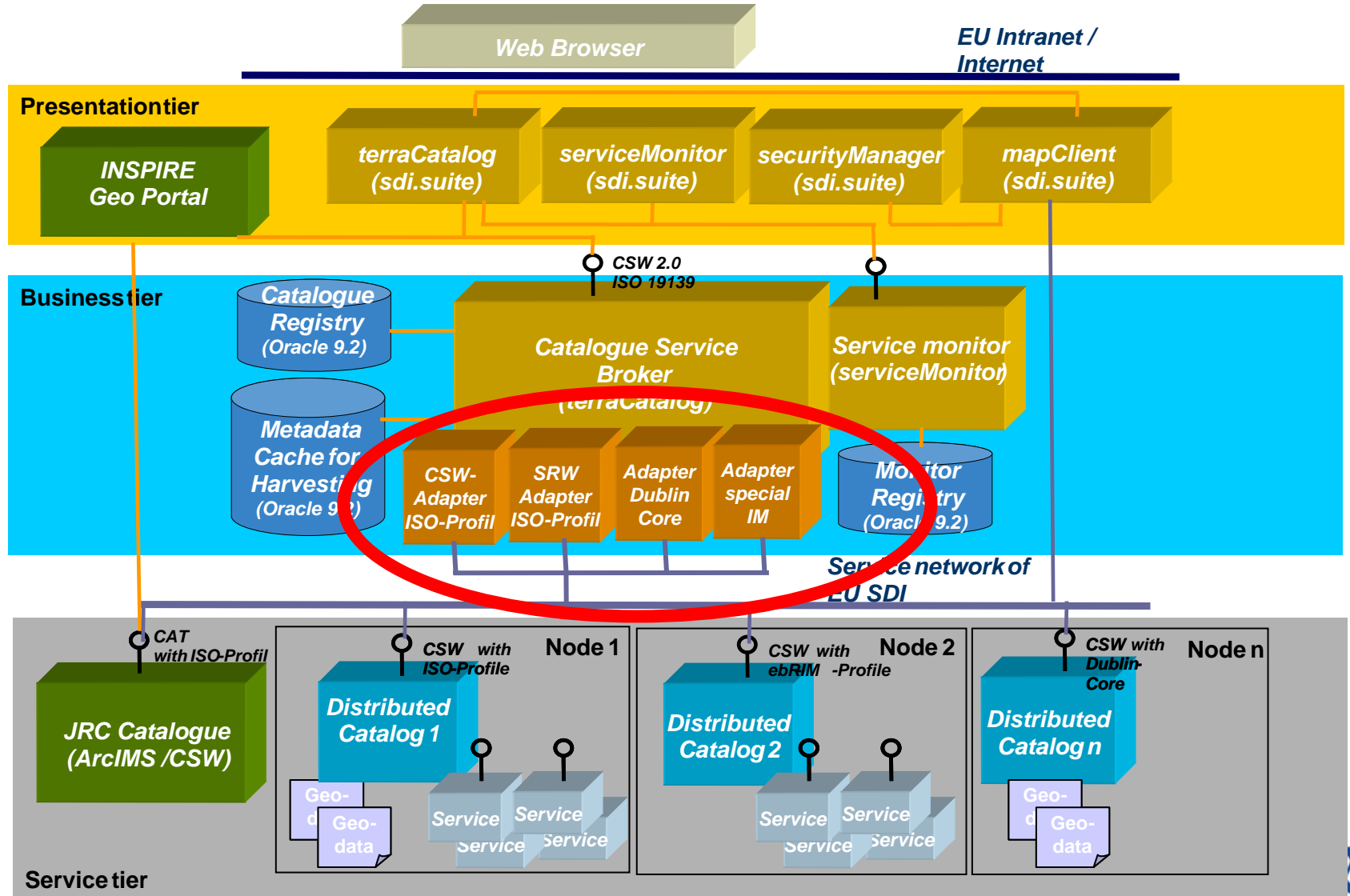






# System Architecture

Joint Research Centre





# Connecting catalogues with broker – the adaptor concept

- Routing Adaptors
- Encoding Transition Adaptors
- Binding Transition Adaptors
  - Simple Binding Transition Adaptor
  - Complex Binding Transition Adaptor
- Profile Transition Adaptors



YES...

Provided the same implementation is used

And with good chances if same AP



- Specific adapters need to be implemented on a case by case basis
- Shortcomings in specification results in too many degrees of freedom in implementations
  - Interoperability depends largely on implementation and not on specification
- Implementations are not compliant at the OGC CORE level



- OGC DC does not clearly define if all common queryables are mandatory.
  - If not, catalogues should deliver it's queryables in its capabilities document;
    - also not standardized
- A capabilities document should include descriptions of the distributed catalogues
  - It should be possible to include names/identifiers in a GetRecords Request, telling which Catalogues to search for
    - In our implementation we did this using an project specific protocol in the SOAP header of the requests on the broker



- Which distributed catalogue delivers how many hits
- From which distributed catalogue does an entry originate
- Shortcomings of the information model mappings of the different CSW 2.0 application profiles
  - semantic of the content of the OGC DC is not well defined, a mapping from DC to ISO or ebRIM is very difficult
  - ebRIM profile does not define a content based information model, the mapping from ebRIM information to DC or ISO is difficult
- Updating is not foreseen in harvesting (through the service)
- Link between data metadata and service metadata



# Recommendations

- TEST SUITE
- One application profile
- One binding
- bridge to UDDI maybe worth looking at
- Investigate alternative approaches?