



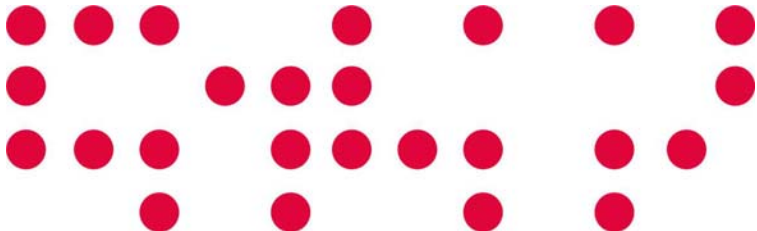
# Work Item D: Fostering federated media content and/or metadata repositories

Jose Maria Fontanillo  
(RedIRIS)

November 5th 2009



1. Work item D Objectives
2. Open questions
3. Conceptual model
4. Metadata
5. Metadata exchange
6. Federation models
7. The ARCA example
8. Conclusions





# Work item D Objectives

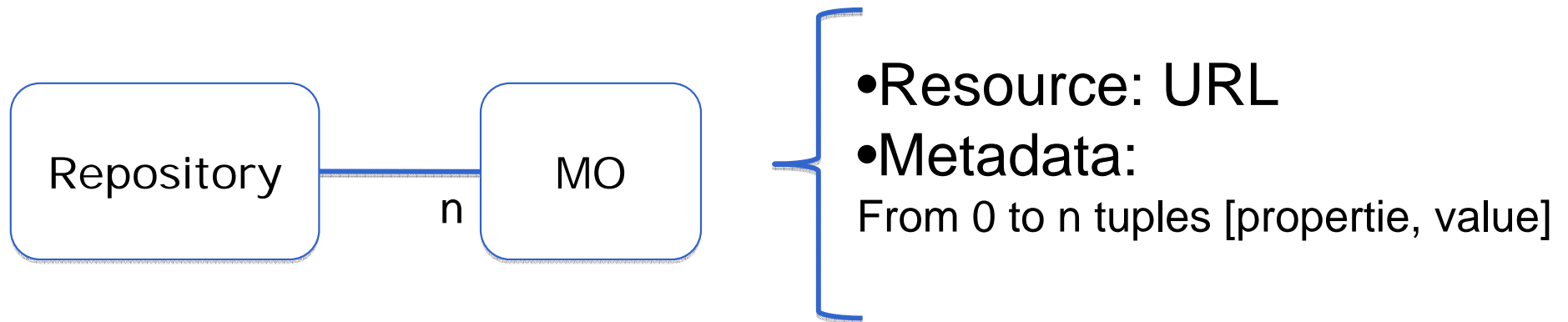
- Facilitate the creation of federated content aggregating from different organisations.
- Collect recommendations on the metadata levels, structures and formats to ensure the better searchability.
- Describe the problem space of using the NREN AAI federations with media content storage and repository services.
- Liaise with TF-EMC2 to find a path to solve the problems and/or suggest suitable workarounds
- Liaise with TF-Storage on content storage issues.





# •Federating content aggregating from different organizations.

- Does it make sense?
- Do we have repositories yet?
- Is there any initiative to federate it?
- Which services could be developed in the federation?
- Which is the target of those services? Repositories?, end users? Both?
- Harvest metadata or full content?
- Types of content?
- Metadata with catalogation items, keywords, both?
  - Could we agree some catalogin?
- Which could be the lowest common denominator for metadata interchange?
- Open content only?



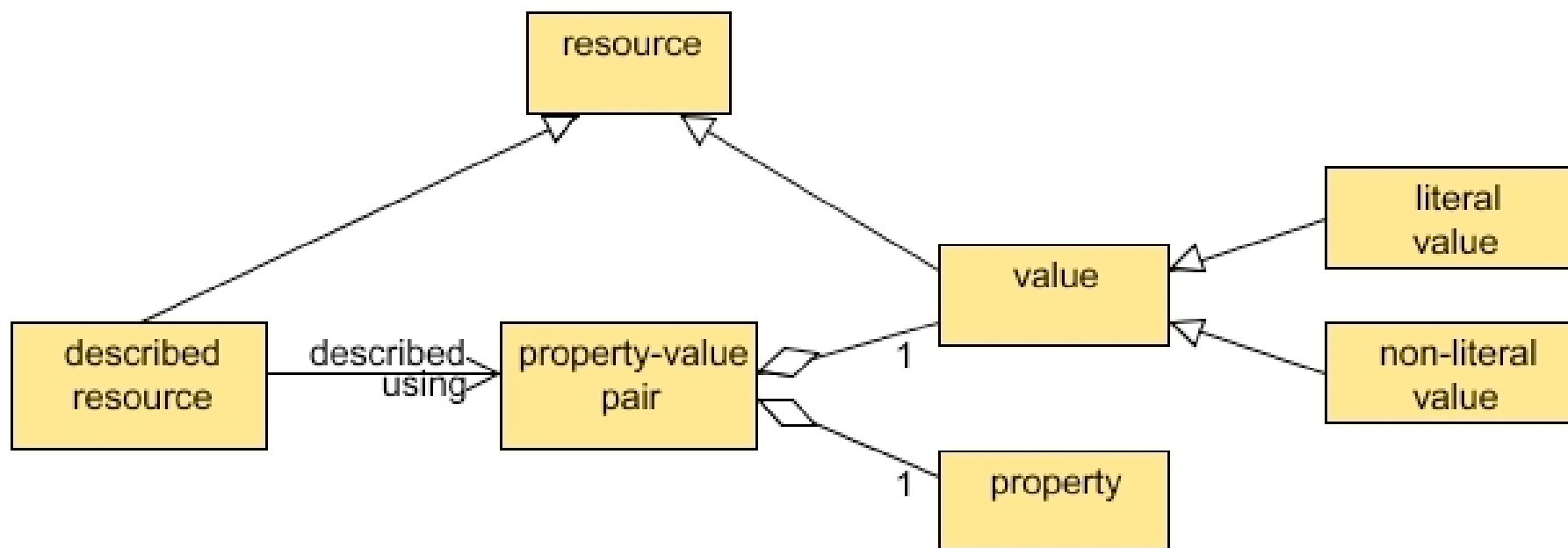
A MO can be part of other MO to build complex MO



- Each repository represents metadata of its MOs in a different format
- The ways I found to represent metadata: custom xml schema, dublin core, custom schema based in SQL. Any more?
- We need to agree on a common semantic and format although different metadata formats can be used internally.



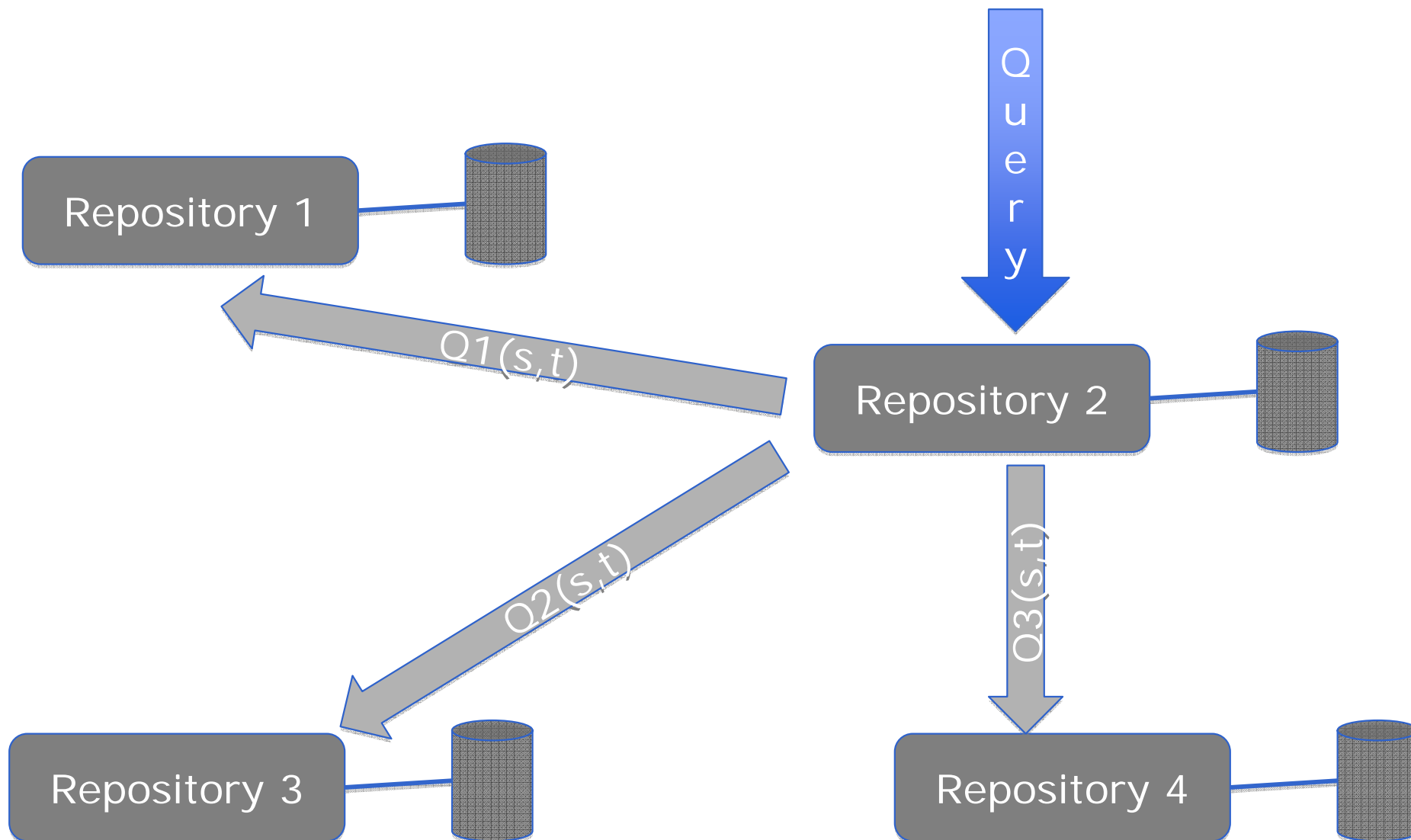
- DC is a model for metadata that can be represented in different ways: xml, rdf, rss, html, etc
- DC has two flavors: simple and qualified
- Simple DC is based on 15 elements: Title, Creator, Subject, Description, Publisher, Contributor, Date, Type, Format, Identifier, Source, Language, Relation, Coverage and Rights





- OAI-PMH is based on a client-server architecture
- 'Harvesters' from repositories (data providers)
- Data providers are required to provide XML metadata in Dublin Core format, and may also provide it in other XML formats
- 6 OAI-PMH requests (verbs): GetRecord, Identify, ListIdentifiers, ListMetadataFormats, ListRecords, ListSets
- The data providers can be persistent, transient or no information about deletions
- The requests are embedded in HTTP
- Complex





# Synchronous model with aggregator



GOBIERNO DE ESPAÑA

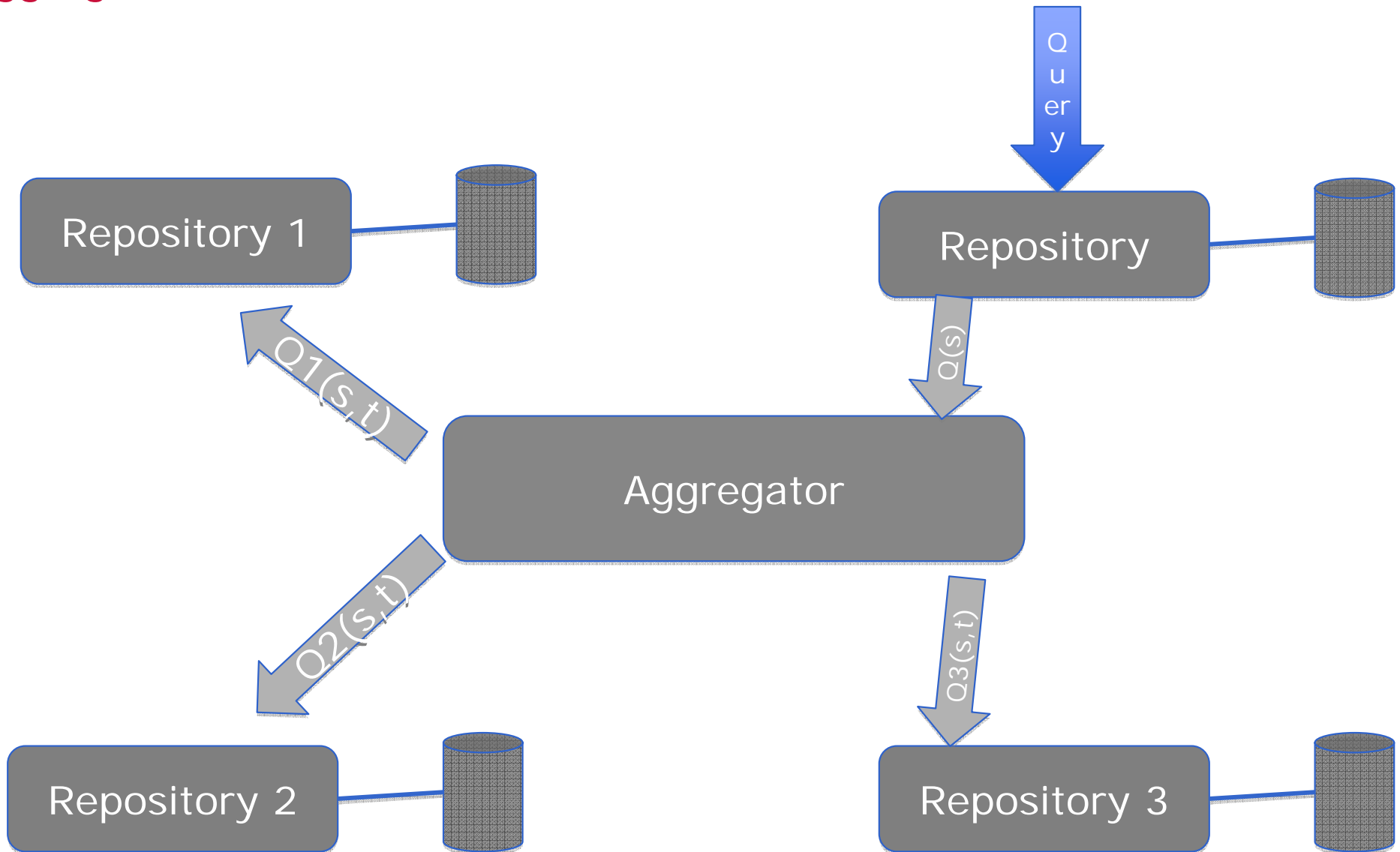
MINISTERIO DE CIENCIA E INNOVACIÓN

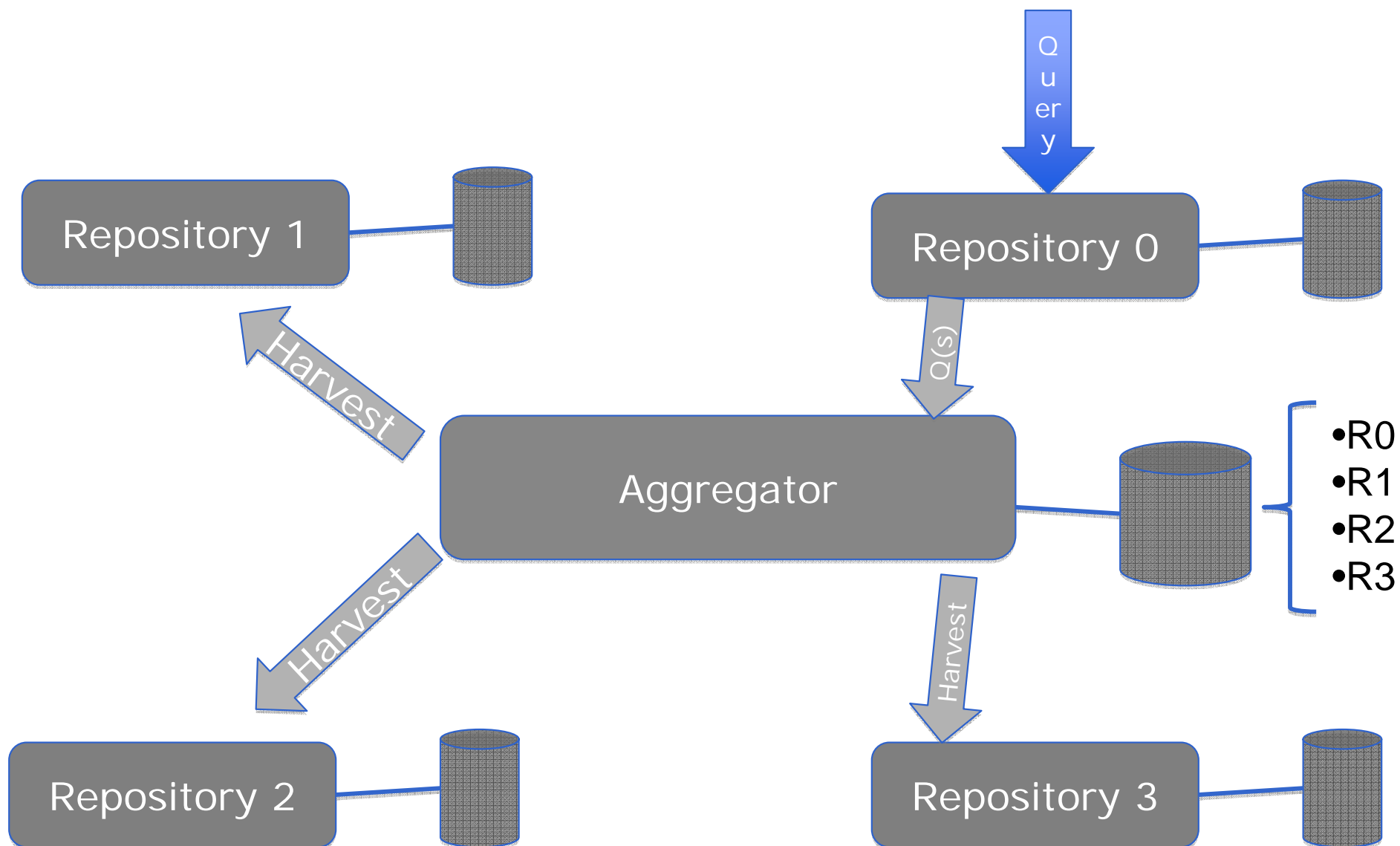
MINISTERIO DE INDUSTRIA, TURISMO Y COMERCIO

SECRETARÍA DE ESTADO DE TELECOMUNICACIONES Y PARA LA SOCIEDAD DE LA INFORMACIÓN



Red IRIS







- Simple approach
- Based in the use of RSS 2.0 (extended with most used namespaces in commercial services) to harvest metadata
- Full metadata databases are replicated in ARCA
- Not incremental, each day the full database is rebuilt from scratch.





- We need to agree in a basic set of tags to be used in metadata with recommendations on the syntax and semantics. Simple DC for example
- A model of federation
- A model of services with targets



GOBIERNO  
DE ESPAÑA

MINISTERIO  
DE CIENCIA  
E INNOVACIÓN

MINISTERIO  
DE INDUSTRIA, TURISMO  
Y COMERCIO

SECRETARÍA DE ESTADO  
DE TELECOMUNICACIONES  
Y PARA LA SOCIEDAD DE  
LA INFORMACIÓN



Red  
IRIS



[www.red.es](http://www.red.es)

Edificio Bronce,  
Plaza Manuel Gómez Moreno s/n  
28020 Madrid. España

Tel.: 91 212 76 20 / 25, Fax: 91 212 76 35 **1**