DRIVER Infrastructure
The Demo

driver
Digital Repository Infrastructure Vision for European Research

**"European Information Space: Infrastructures, Services, and Applications Workshop"**

Paolo Manghi

ISTI-CNR

paolo.manghi@isti.cnr.it

DRIVER Infrastructure

- ## Targets VREs addressing "data integration"
  - Examples in Europe: DARE (NL), BASE (DE), PUMA (IT), DART Europe, etc.

- ## In order to
  - Overcome high maintenance costs, low degrees of scalability, and data sharing

UI

Search

Index   ...   Index

Information Space

Aggregator

External data sources
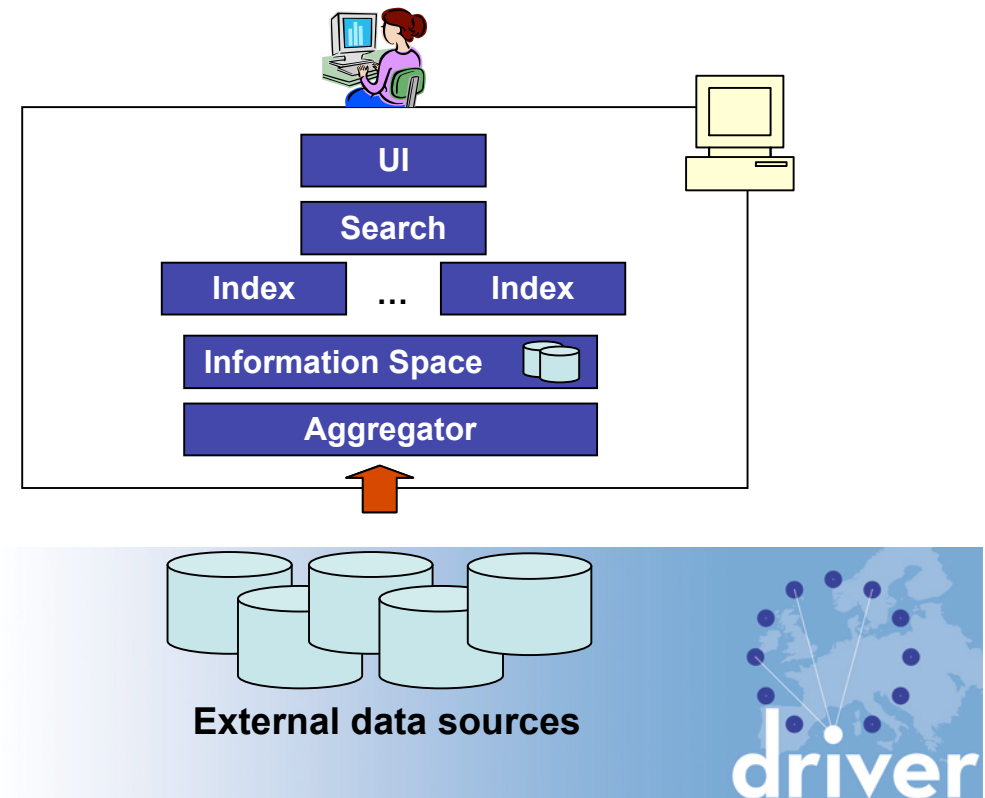
driver

DRIVER Resource KITs

- DRIVER middleware currently outsources two application paradigms
    - Data aggregation KIT
    - End-user functionality KIT
- Open to other application paradigms

DRIVER Data Aggregation KIT

- ## Harvesting tools

  - Easy "harvesting" of heterogeneous external data resources

- ## Aggregation tools

  - Data manipulation: e.g. enrichment, modification, transformations

  - Definition of *Information Spaces:* new data, of higher quality

- ## Data storage and access outsourcing

  - Automatic management of data storage and data access (sharing, replication, indexing, etc)

driver

DRIVER End-User functionality KIT

- ## User Interfaces
  - Automatic adaptation to different Information Spaces features

- ## User profiling tools
  - Acquiring and exploiting user interest info

- ## User Community management
  - Enabling and easing the process of data selection activities

- ## User Recommendation mechanisms

driver

DRIVER European VRE

- *European VRE for Open Access publications*
  - Harvests content from European Repositories
    - Now: 51 Repositories for a total of 250.000 Open Access documents
    - Soon: up to 73 Reps by the end of the year
  - Maintains one European Information Space
    - *DRIVER Metadata Format* (DMF)

- Other VREs are to join DRIVER in the near future
  - Adding new content
  - Sharing existing content

driver

DRIVER infrastructure: the DEMO

Aggregation: populating the DRIVER Information Space (DMF)

Running queries over the DRIVER Information Space (DMF)

User Functionality KIT

Data KIT

Repositories

DRIVER middleware

Harvesting: enriching DRIVER with content from OAI-PMH Repositories

driver

DEMO links

- http://search1.driver-repository.eu/webInterface/

- http://driver10.isti.cnr.it/

- http://129.70.12.42:8007/cgi-bin/ASmanager.pl
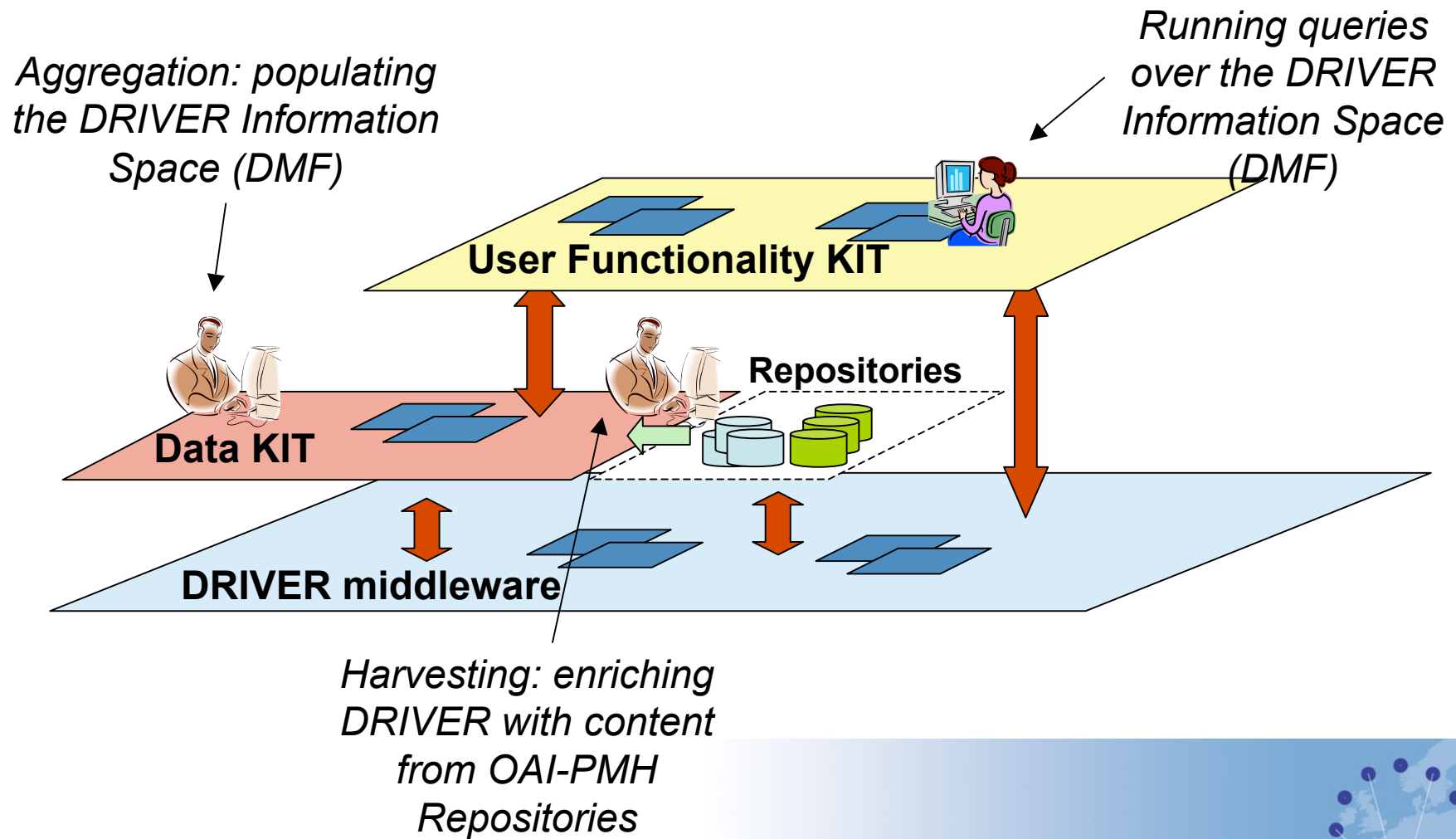
- Open Alpha release on DRIVER Information Space, final release by the end of the year
- Operates over the DRIVER Information Space and DMF documents
- It includes many functionalities, such as user registration and profiling, collections and communities, others will be introduced soon, such as browsing or user private collections
- From here we can search the whole information Space or focus on a subset by using collections
- For example:
  - we run the query: "american election"
  - Select the second document
  - Show the DMF: data cleaning, repository info, country
  - Enables definition of collections not only content-based but location-based
  - Show the document splash page
  - Get back to the search page and leave it there (it will be used for the same query)
- Move to the administrative interface, from which we can manage all infrastructure resources
- For example: monitoring of Nodes/Services and Repositories, thus external resources
  - New instances of the services can be added or removed and the infrastructure will keep the VRE's operative by expliting the resources currently available
- We can also register a new Repository
  - New repositories are registered and put in "pending" mode, waiting for validation from an administrator
  - The validation of a repository to be harvested requires the allocation of storage space, of indexes, and the notification to the administrators that will have to configure the harvesting operation
- For example we can validate the Repository British Library as we know it contains documents related with "american election" to see the immediate effect of the harvesting
- Open Aggregator in other Browser TAB
- These effects can be tracked down through the list of data structures created by the infrastructure
- Run again in the first tab the same query

driver

Home page

Query results

DRIVER Metadata Format

Virtual collections

Administrative interface

Repository management

Node and Service management

Repository registration and validation

Repository harvesting and aggregation

DRIVER Aggregator Manager - OAI Admin Panel

List of Repositories

Repository Information

Identify - ListSets - ListMetadataFormats

Harvesting

Nr.	Format unit	Status	Last Harvesting Date	Harvestingtype	Harvesting schedule
1	oai_dc	inactive	2007-10-28T17:25:46Z	REFRESH	WEEKLY

Aggregating

Start aggregating Test Mapping Re-Feed

Repository:	British Library Research Archive
Repository ID:	14-a8a2cc8b-b42c-442f-a67c-a5a3fde13c34_UmVwb3NpdG9yeVNlcnZpY2VSZXNvdXJjZXMvUmVwb3NpdG9yeVNlcnZpY2VSZXNvdXJjZVR5cGU=
Harvesting	17-a03d4b2a-b31d-409b-af4b-ad741d882f19_SGFydmVzdGluZluc3RhbmNlUlENSZXNvdXJjZXMvSGFydmVzdGluZluc3RhbmNlUlENSZXNvdXJjZVR5cGU=