



# The DRIVER Project



**Paolo Manghi**

*ISTI - National Research Council, Italy*

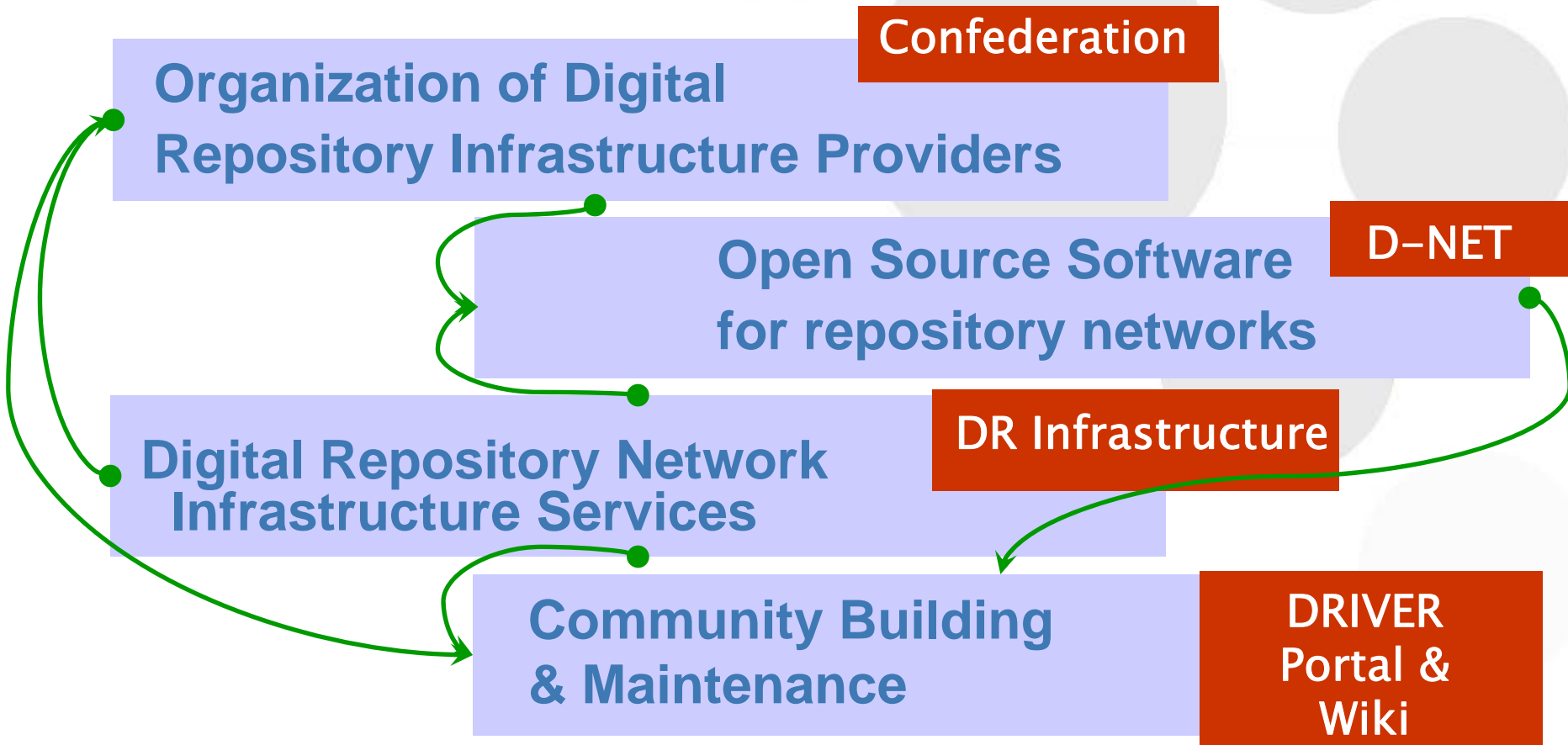


**driver**

# General Information

- Duration: 24 months
  - Budget: 2.7 m EUR
  - Timeplan: 12/'07 -11/'09
  - Main Deliverables:
    - ⊙ Digital Repository Infrastructure
    - ⊙ Compound Object extension prototype
    - ⊙ European Digital Repository Confederation
  - Funded by the European Commission, “Research Infrastructure” Unit, FP 7
  - Consortium Partners
    - ⊙ Univ.of Athens (GR)
    - ⊙ Univ. of Bielefeld (GE)
    - ⊙ CNR-ISTI (IT)
    - ⊙ STICHTING SURF (NL)
    - ⊙ Univ. of Nottingham (UK)
- cont...
- ⊙ Univ. of Bath (UK)
  - ⊙ Univ. of Warszawski (PO)
  - ⊙ Univ. of Gent (BE)
  - ⊙ Univ. of Goettingen (GE)
  - ⊙ Danish Technical University (DK)
  - ⊙ Universidade do Minho (PT)
  - ⊙ Narodna in univerzitetna knjižnica (SLO)

# DRIVER II activity areas and outcomes





# The DRIVER Infrastructure – D-NET



driver

# Scenario

- Proliferation of Institutional Repositories
  - ⦿ Gathering and exposing research outcome (publications) local to the Institution
- Typical user communities requirements
  - ⦿ Discover information, e.g. cross-repository search
  - ⦿ Elaborate information, e.g. cross-comparison, object composition (e.g. complex objects)
- *Communities demand for Aggregative Digital Library Systems, to be provided by the community organization*

# Traditional solution: Aggregative Digital Library Systems (ADLS)

## ○ ADLSs



*Aggregation system*: maintaining and populating an *Information Space* by aggregating content from a collection of OAI-PMH Repositories

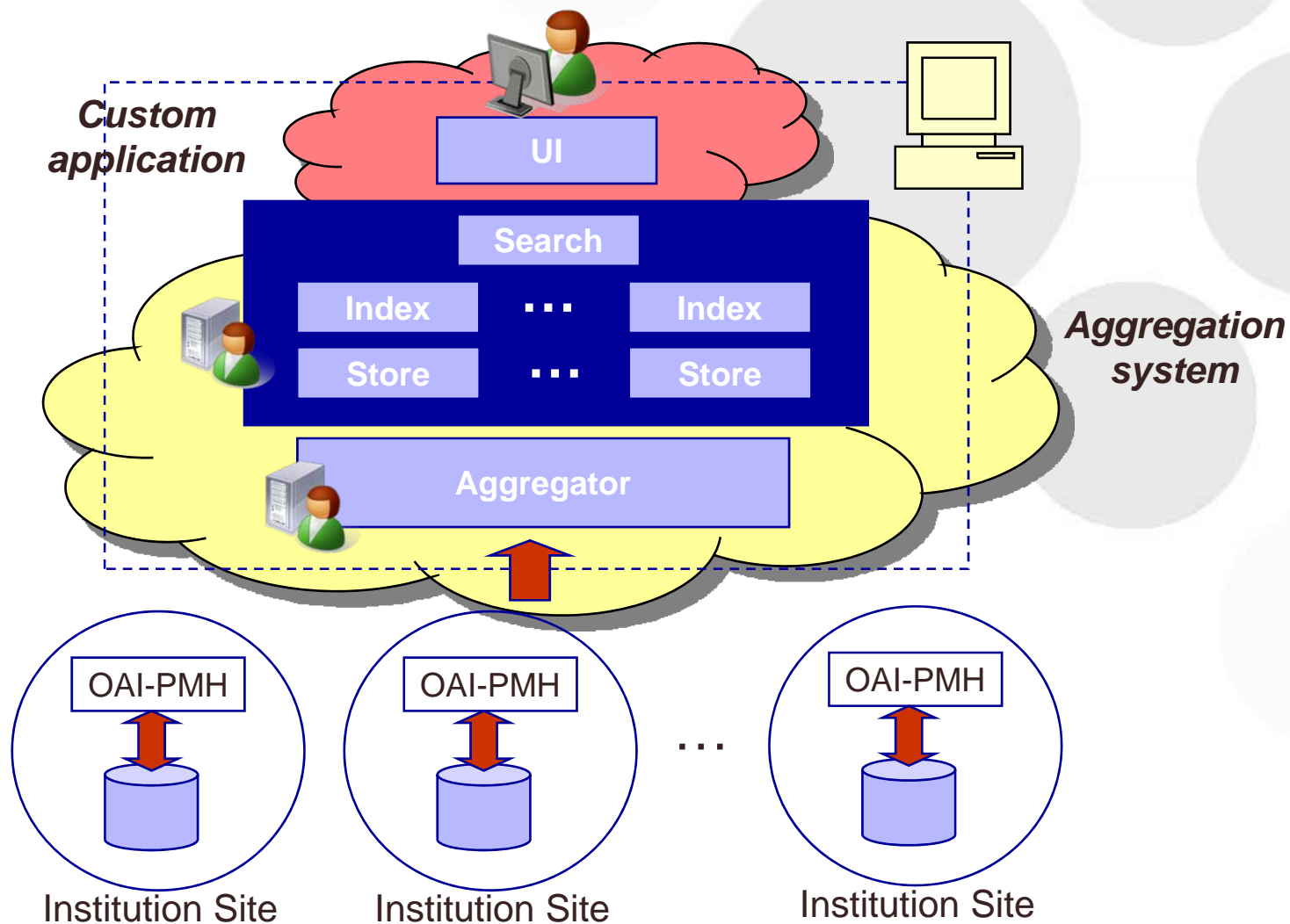


*Custom application*: providing community-specific functionalities via Web User Interfaces

## ○ Well known examples

- ⊙ BASE (Germany)
- ⊙ DAREnet (Netherlands)
- ⊙ OAIster (USA)
- ⊙ Others...

# ADLS architecture and technology



# The DRIVER project goals



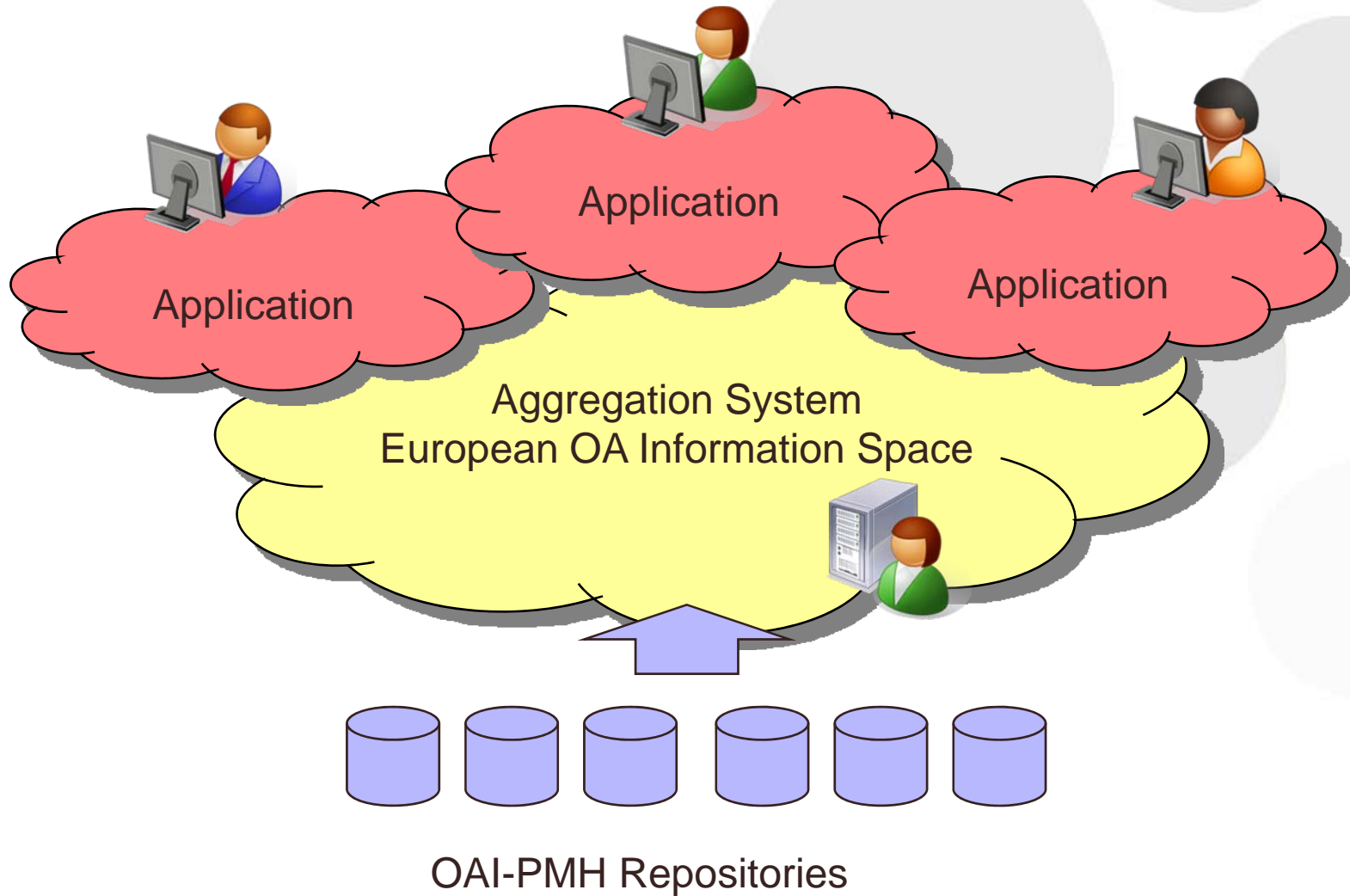
Realization of the European Information Space for Open Access research publications



Provide demanding research communities with free and easy access to this Information Space

- ⦿ Through User portals
- ⦿ Through applications and services

# The DRIVER project goals

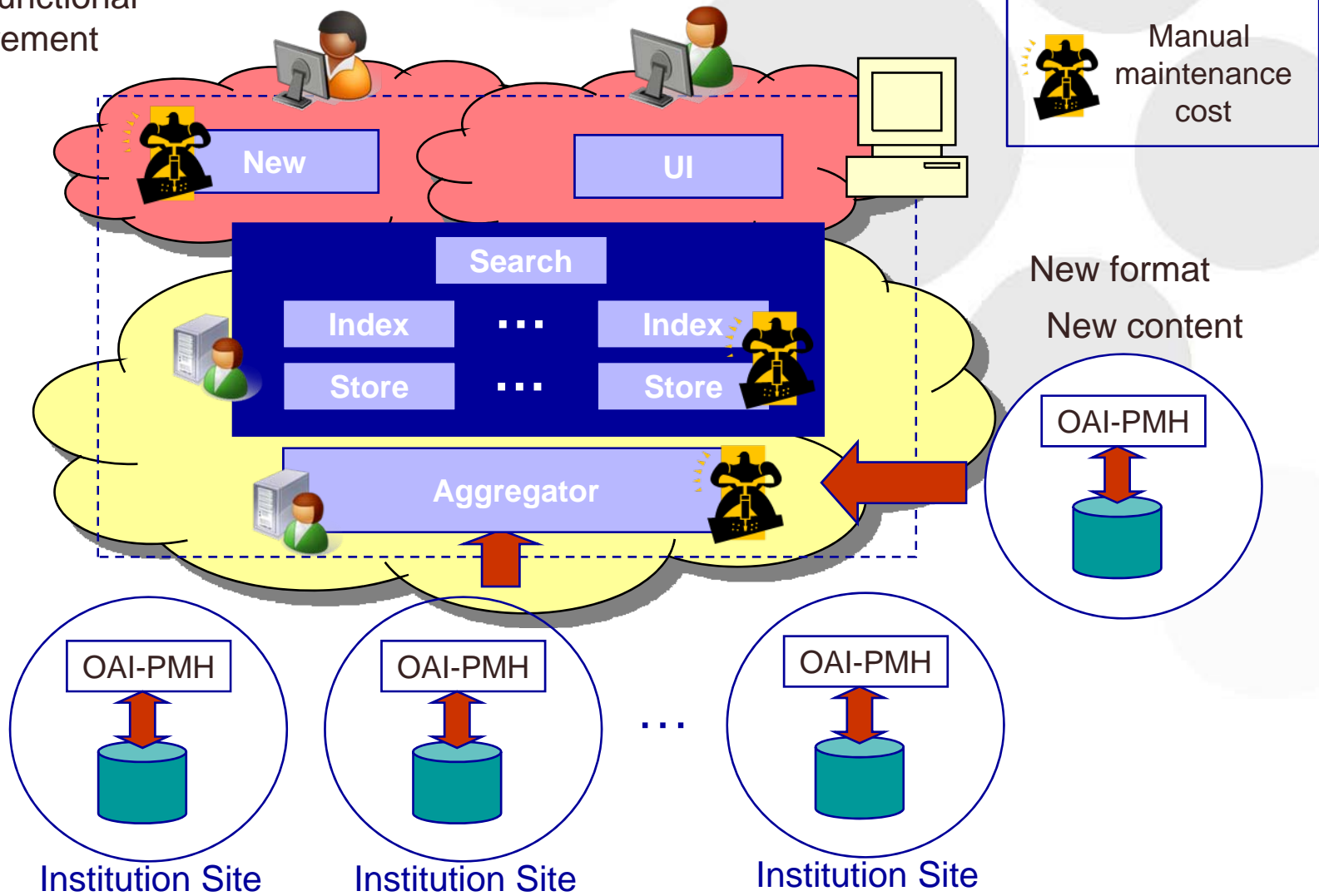


# DL community requirements

- Nowadays communities are increasingly characterized by
  - ◎ Evolving requirements, i.e. content and functionality should be adaptable to new needs, opportunities, and scenarios
  - ◎ Short-term duration, e.g. communities as work groups may require customized applications for the time of the collaboration
  - ◎ Low funds available for technology design/development
- Traditional ADLS solutions are not sustainable

# Traditional ADLSs drawbacks

New application  
or new functional  
requirement



# The DRIVER project solution

## *Infrastructural approach*

- An infrastructure is a running environment
  - ⊙ Maintained by one *responsible organization (RO)* for the benefit of *participating organizations (POs)*
  - ⊙ Where *resources* – e.g. content, functionality, hardware – can be shared by POs and reused by other POs so as to serve their application needs
- In DRIVER functional *resources* are implemented as Web Services
  - ⊙ Running at different sites on the Internet
  - ⊙ Providing functionality in “isolation”; e.g. Index management, storage management, aggregation tools, search

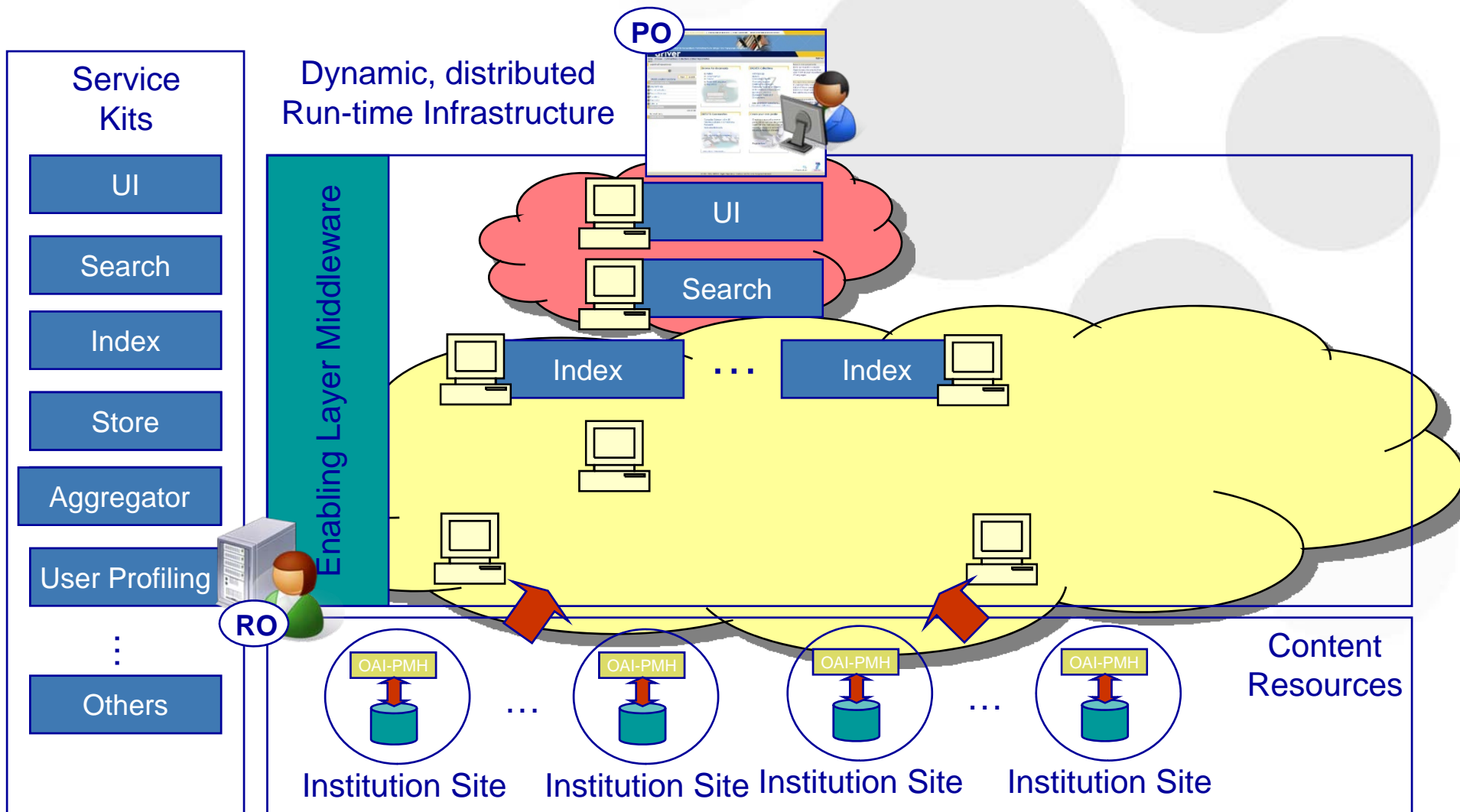
# DRIVER Infrastructure

## *Service-Oriented Architecture*

- The RO manages the *Enabling Services*, a middleware for the management of the infrastructure
  - Information Service: registration, de-registration and discovery of running services
  - Services are “dynamic”: join/leave & discovery of available resources
  
- POs can
  - Provide (register) and share resources in the form of running services
  - Build applications combining (deploying, discovering and reusing) available resources, i.e. set of running services, to provide the expected functionality

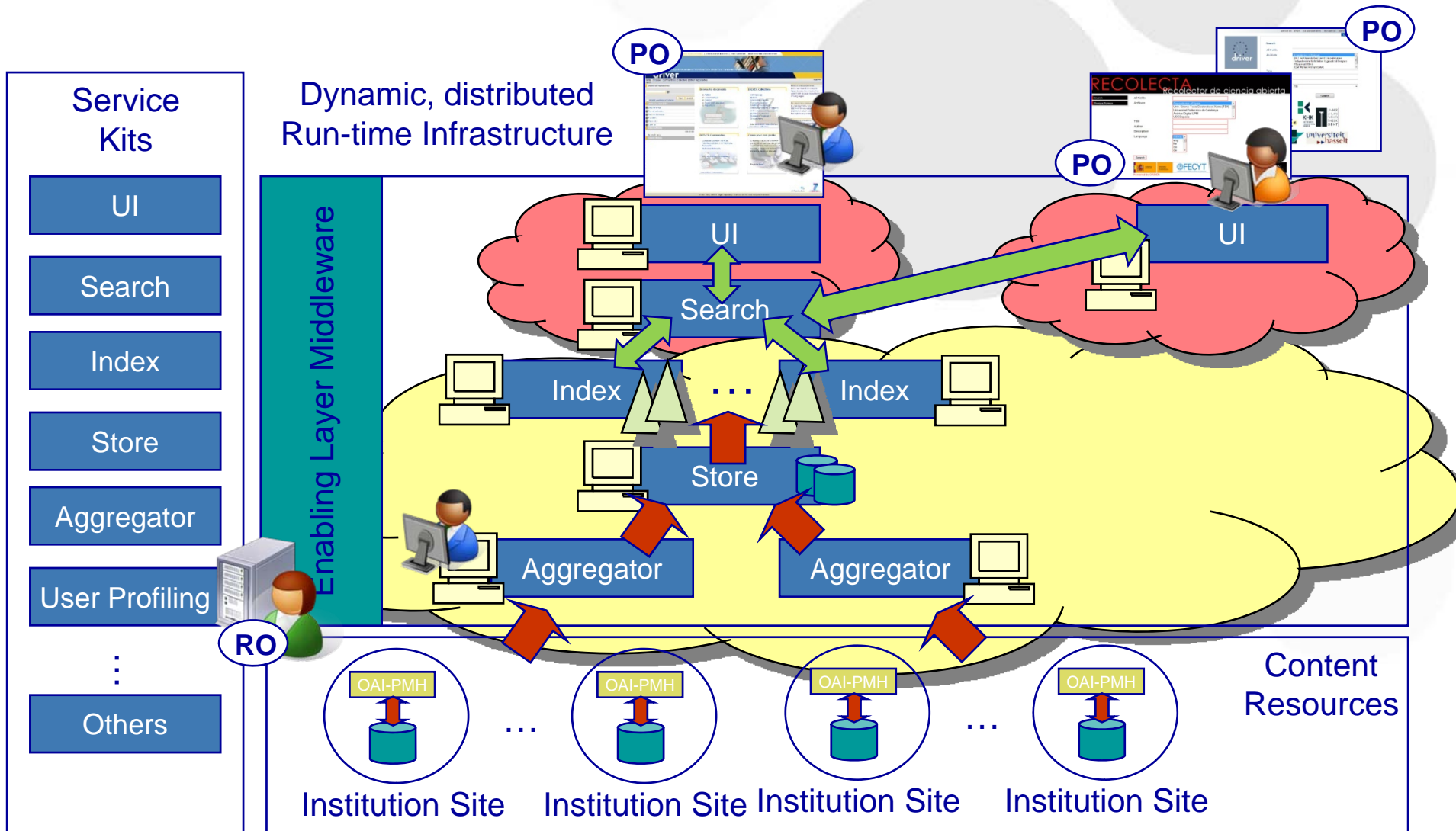
# ADLSs in DRIVER

*Infrastructure distribution and deployment*



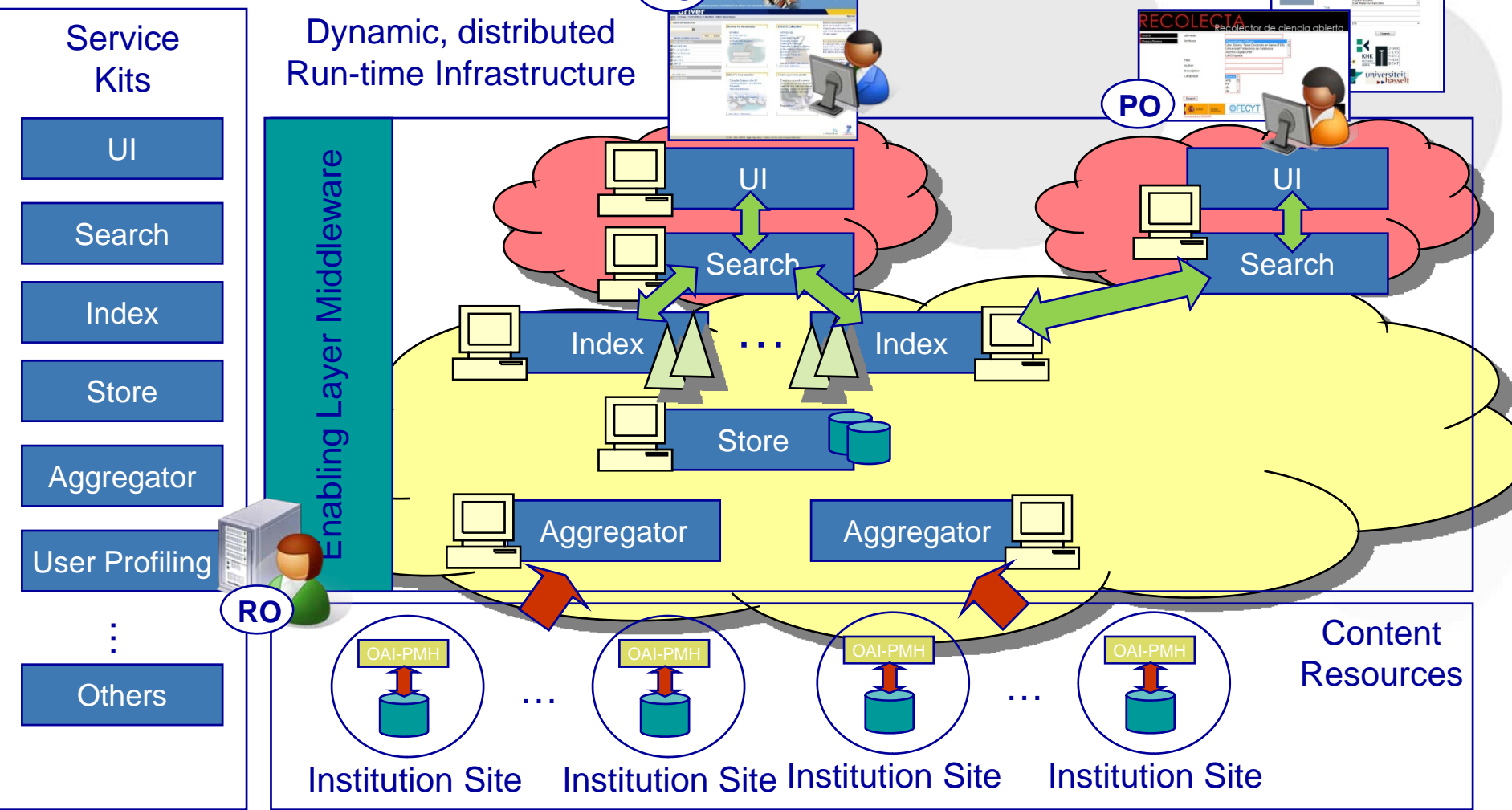
# Infrastructure

## Use and reuse



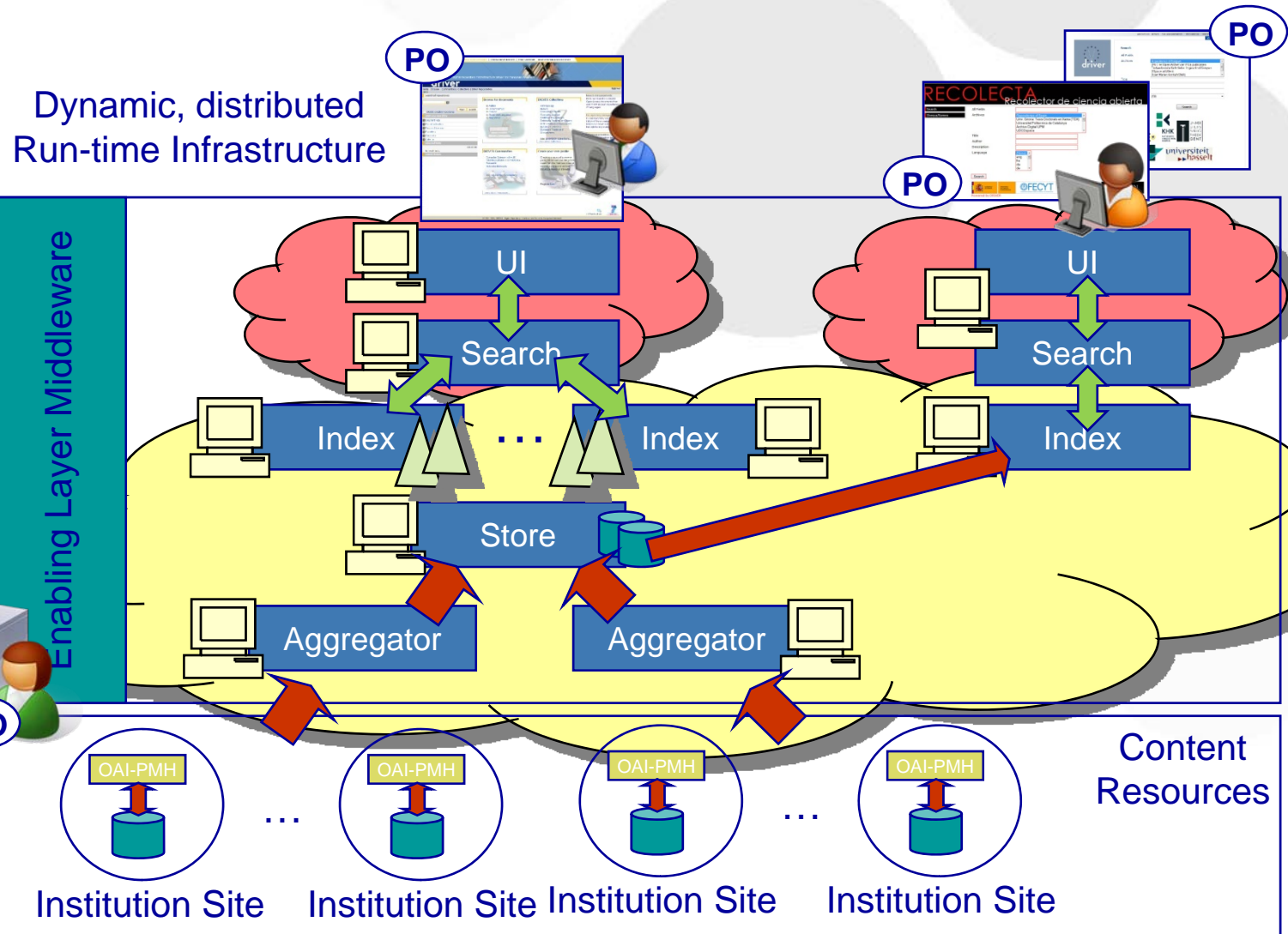
# Infrastructure

## Reuse (2)



# Infrastructure

## Reuse (3)



# Infrastructure Openness

Example of the DRIVER-II extension

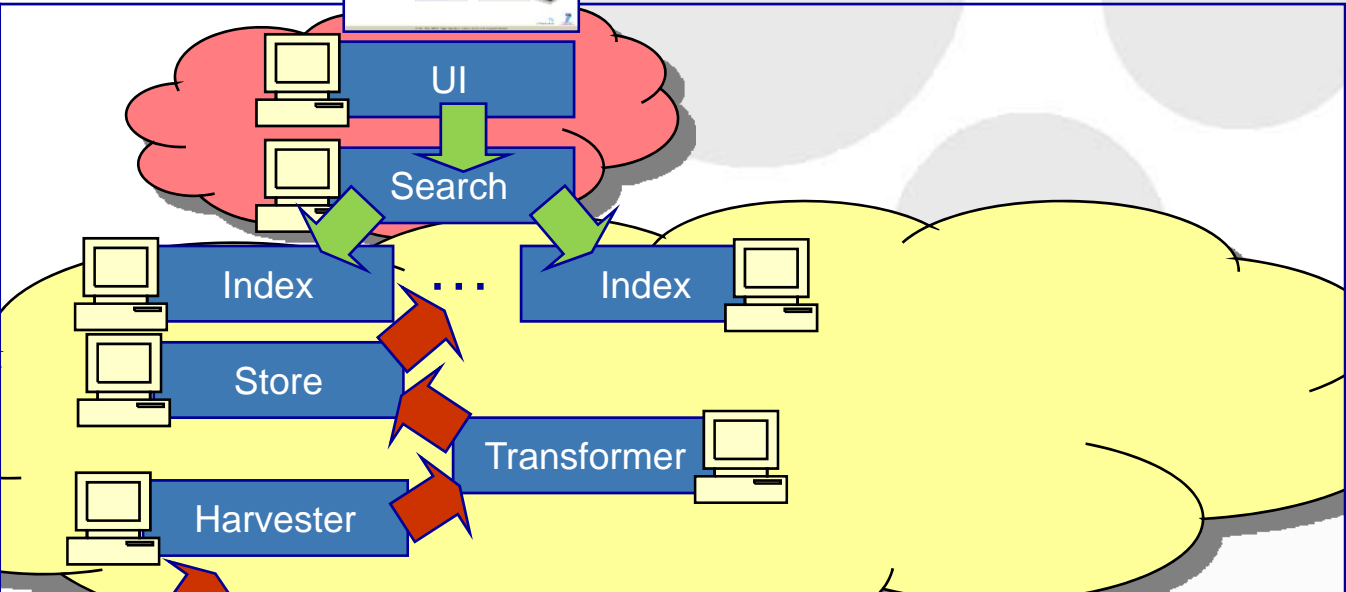
Dynamic, distributed  
Run-time Infrastructure

PO

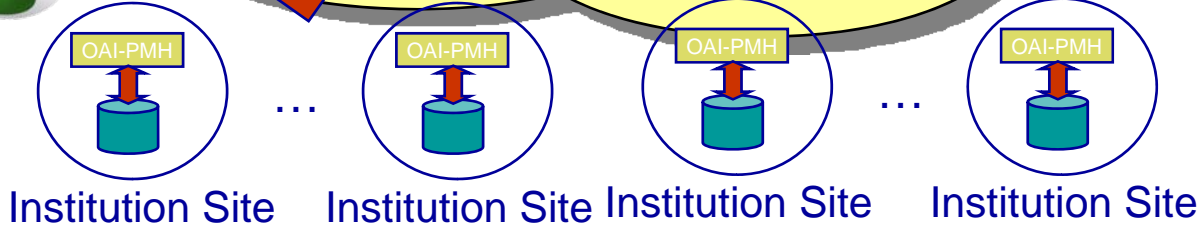


- Service Kits
- UI
  - Search
  - Index
  - Store
  - Aggregator
  - User Profiling
  - ...
  - Others

Enabling Layer Middleware



RO



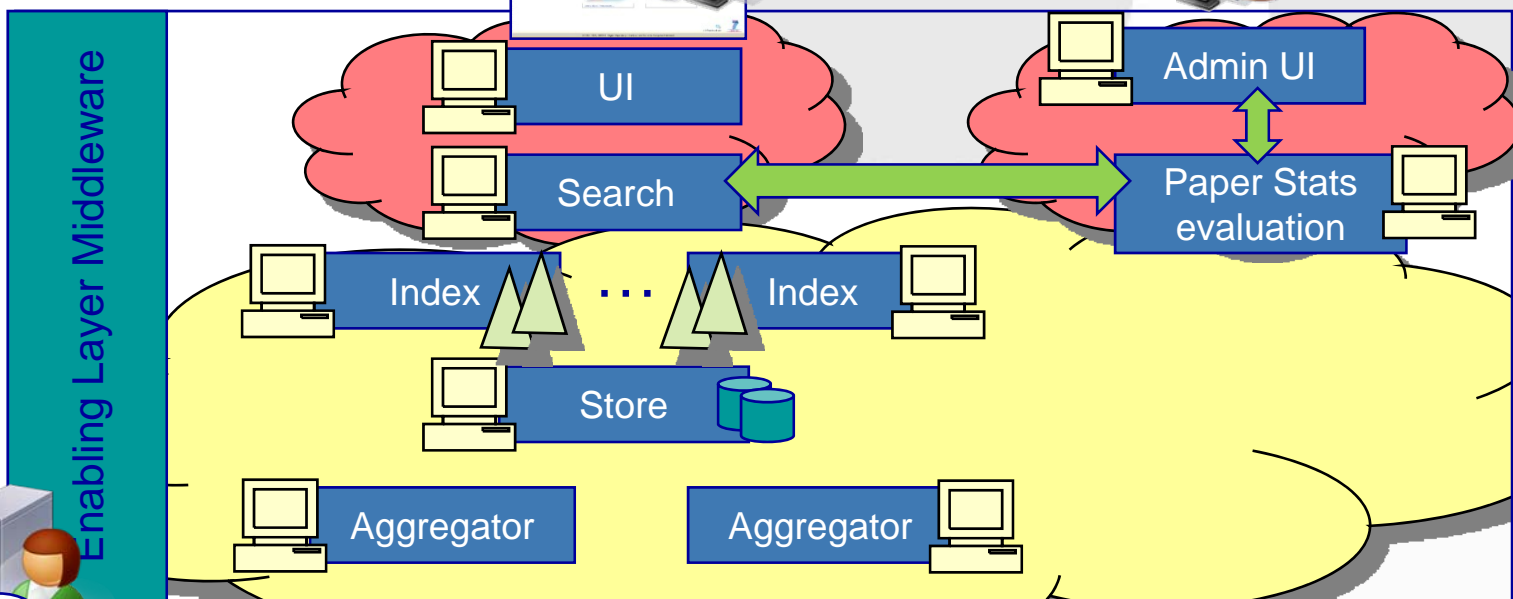
Content Resources

# Infrastructure Openness

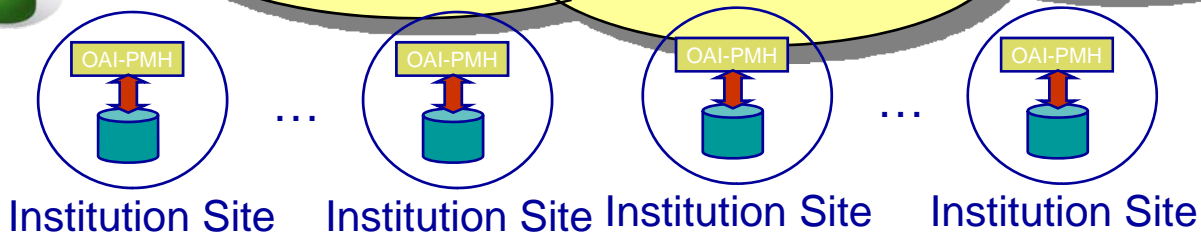
*Hypothetical functionality extension*

- Service Kits**
- UI
  - Search
  - Index
  - Store
  - Aggregator
  - User Profiling
  - ...
  - Others

Dynamic, distributed Run-time Infrastructure



RO



Content Resources

# Resource Orchestration

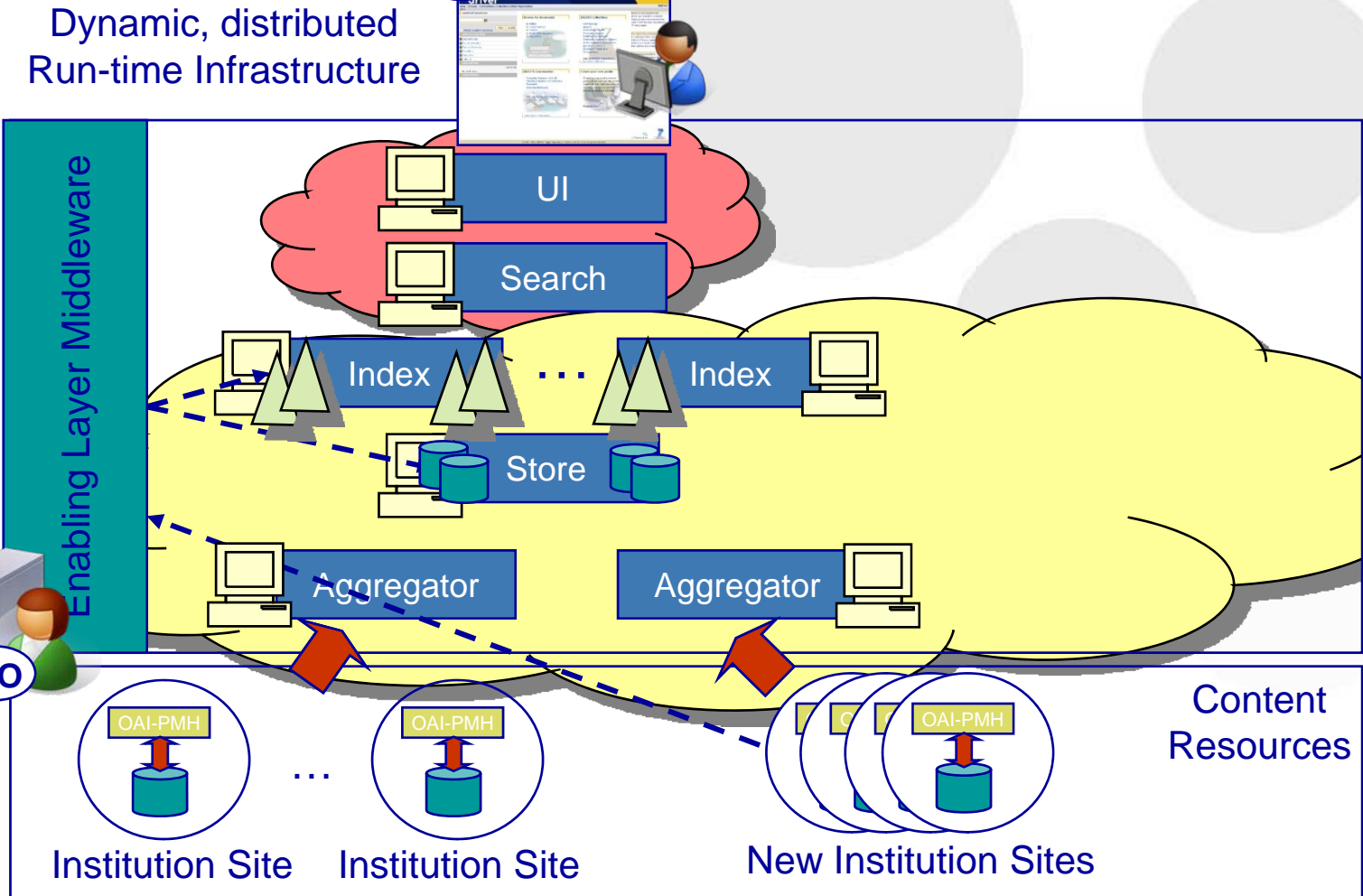
- The Enabling Services give support for service orchestration functionalities
- The RO can configure special *orchestration services (Manager Services)* to automatically administrate and monitor PO applications
  - ⦿ Based on subscription and notification

# Service Orchestration

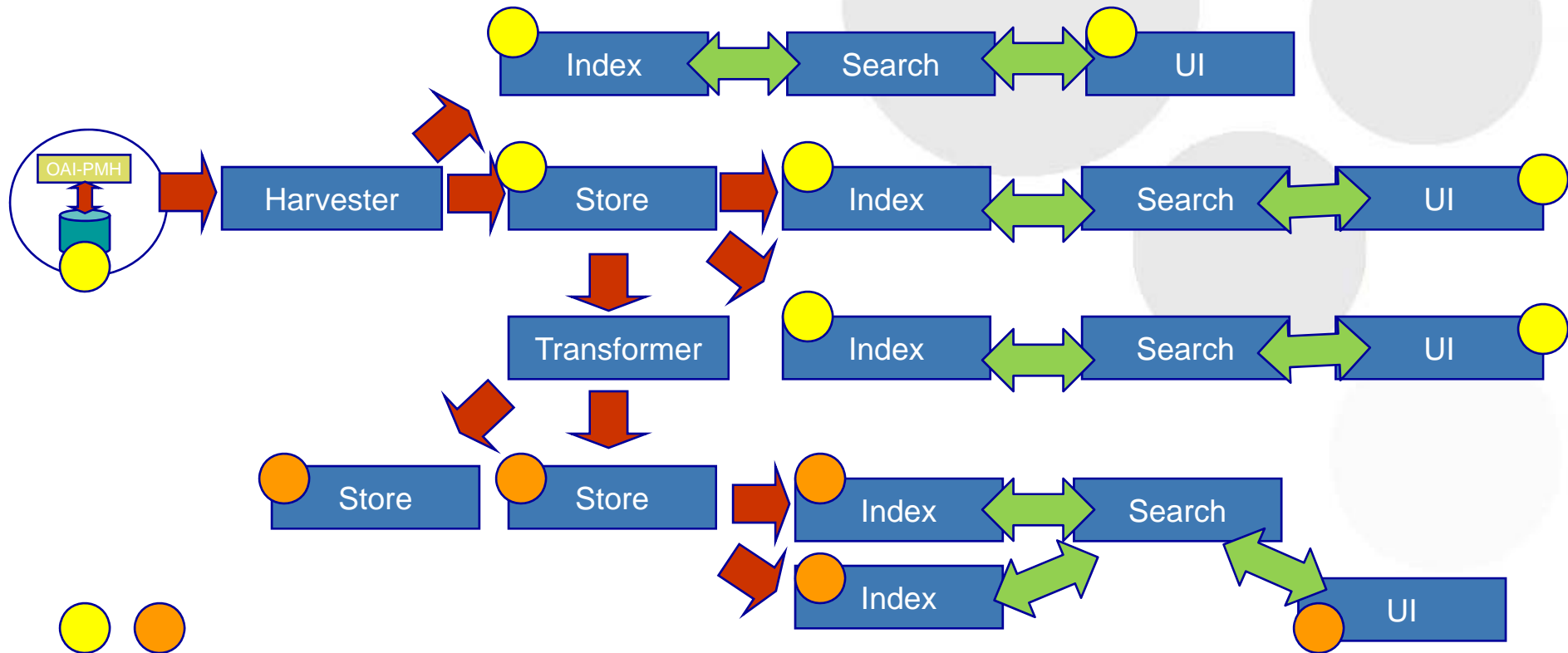
Self-administration: DRIVER orchestration

**Service Kits**

- UI
- Search
- Index
- Store
- Aggregator
- User Profiling
- ...
- Others



# Other orchestration strategies...

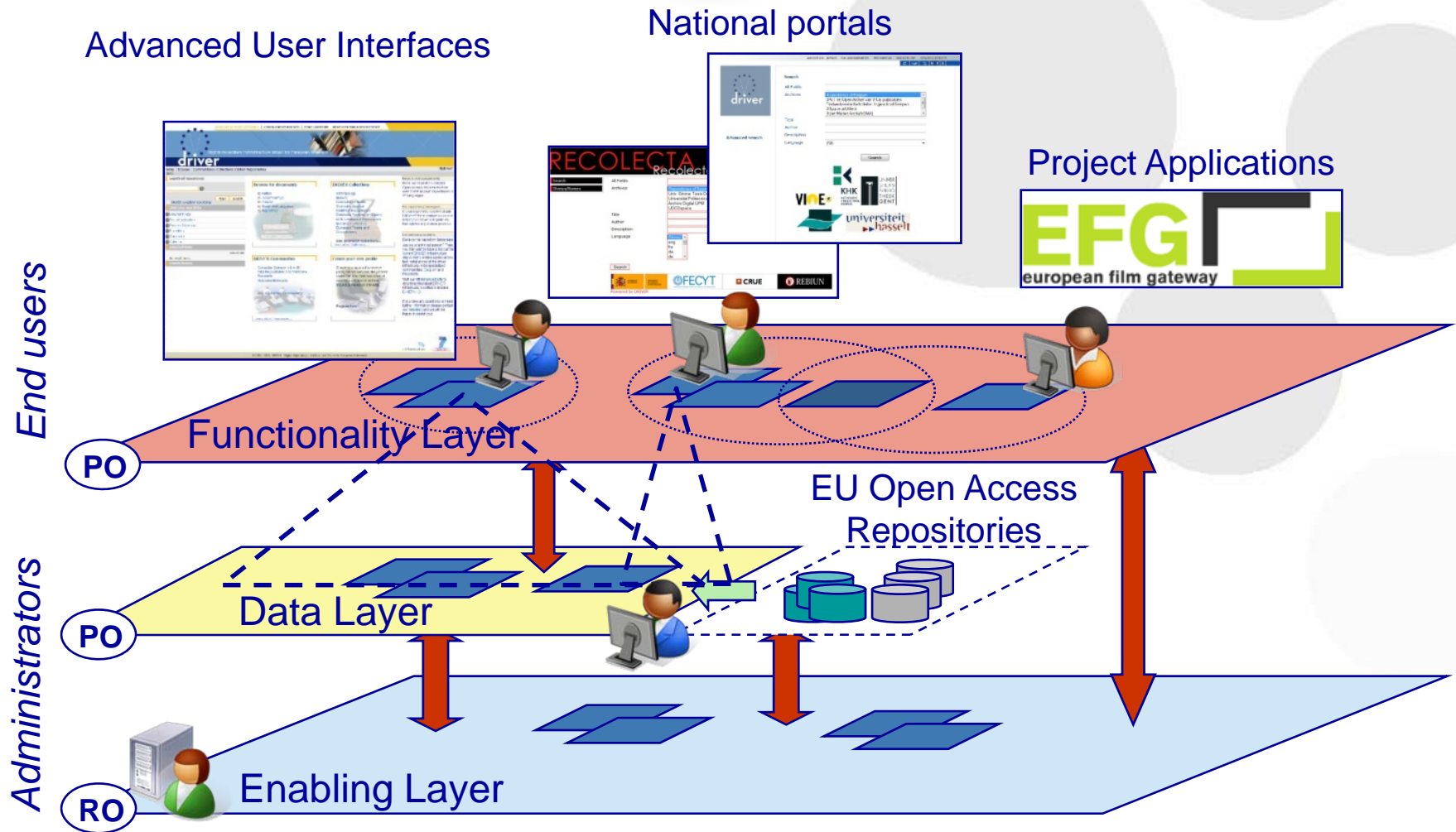


Metadata Formats

# DRIVER Infrastructure benefits

- Architecture enabling *sustainable* business models for *very-large organizations*
  - ⊙ Openness: enabling extension/update/integration of content and functionality resources
  - ⊙ Scalability: in content, functionality and users
  - ⊙ Sharing: communities can share their resources and reuse others to fulfil their needs
  - ⊙ Self-administration (autonomic infrastructure concept): low application installation, maintenance and administration cost

# DRIVER production infrastructure



# DRIVER Aggregation Framework

- Metadata curation
  - Transformator Services, user friendly interfaces
  - Vocabularies, functions, XSLT
- Harvesting from both Repositories and existing aggregators (subject-based or area-based)
  - Provenance information
- Delegation to national correspondents
- DRIVER Guidelines for Repository Managers (v2.0)

# Status of production infrastructure

## ○ Content

- ⊙ More than 184 harvested repositories (more than twice to come) over 21 countries
- ⊙ More than 835,393 open access documents

## ○ Services

- ⊙ Currently 25 different kinds of Services
- ⊙ Production release: 36 service running instances over 9 nodes located at CNR, ICM and NKUA

## ○ Applications

- ⊙ DRIVER Main, Belgium, Spain-Recolecta

# Conclusion and future issues

- The benefits of the infrastructural approach are proved by the growing requests for
  - ⊙ Low-cost deployment of new applications
  - ⊙ Experimentation of deployment of new infrastructure instances
  - ⊙ EFG project use-case
- Towards D-NET 2.0 – December 2009
  - ⊙ Advanced Aggregation framework
  - ⊙ Advanced management for “Enhanced Publications” Compound Objects and ORE

# D-NET 1.0 release

## ○ Technical partners

- Istituto di Scienza e Tecnologia dell'Informazione, National Research Council, Italy
- ICM Research Centre Warsaw, Poland
- National and Kapodistrian University of Athens, Greece
- University of Bielefeld Library, Germany

## ○ Software packages

- Open Source Apache License
- Downloadable and installable with (limited to availability) technical support



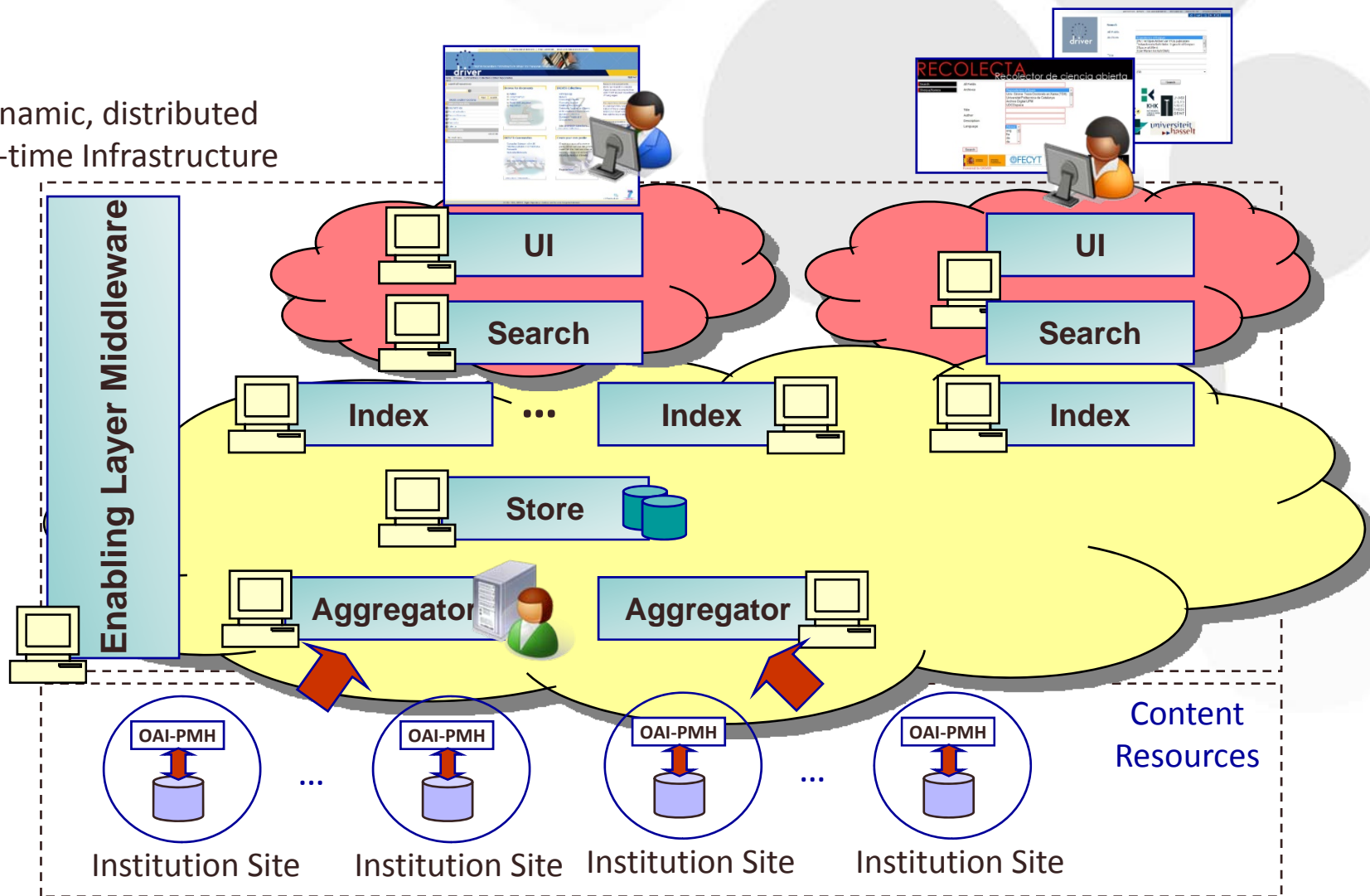
# The DRIVER Infrastructure Demo



driver

# Demo

Dynamic, distributed  
Run-time Infrastructure



# Credits

- Paolo Manghi
- ISTI - CNR
- Speaker's Contact Details
  - ⦿ [paolo.manghi@isti.cnr.it](mailto:paolo.manghi@isti.cnr.it)
  
- DRIVER II Project
- <http://www.driver-community.eu/>
  
- Supported by European Commission

