



# **Science Gateways for Open Science**

Roberto Barbera – University of Catania and INFN  
([roberto.barbera@ct.infn.it](mailto:roberto.barbera@ct.infn.it))

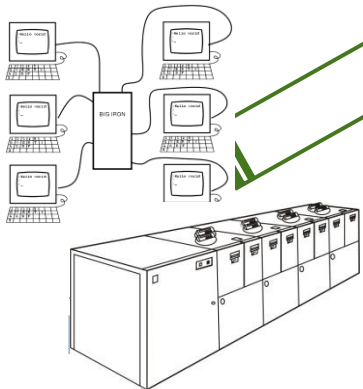
**DPHEP Collaboration Workshop**  
**CERN – 8 June 2015**

- ▶ Introductory concepts and driving considerations
- ▶ INFN approach to Open Access and Open Science
- ▶ Open Science with the Catania Science Gateway Framework
- ▶ Summary and conclusions

# Evolution of Distributed Computing

Cost of hw  
Cost of networks

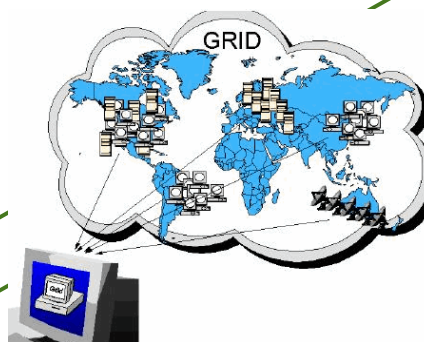
**Mainframe  
Computing**



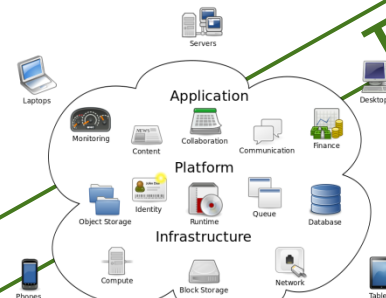
**80's-90's  
Cluster  
Computing**



**90's-00's  
Grid  
Computing**



**00's-10's  
Cloud  
Computing**



Power of COTS  
WAN bandwidth

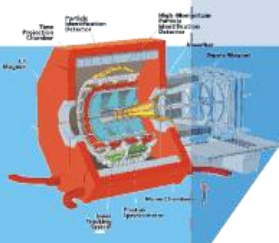
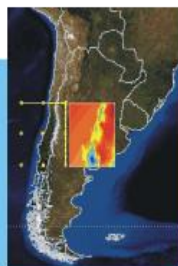
Time



⇒ **Virtual Research Communities**



⇒ **e-Infrastructure**

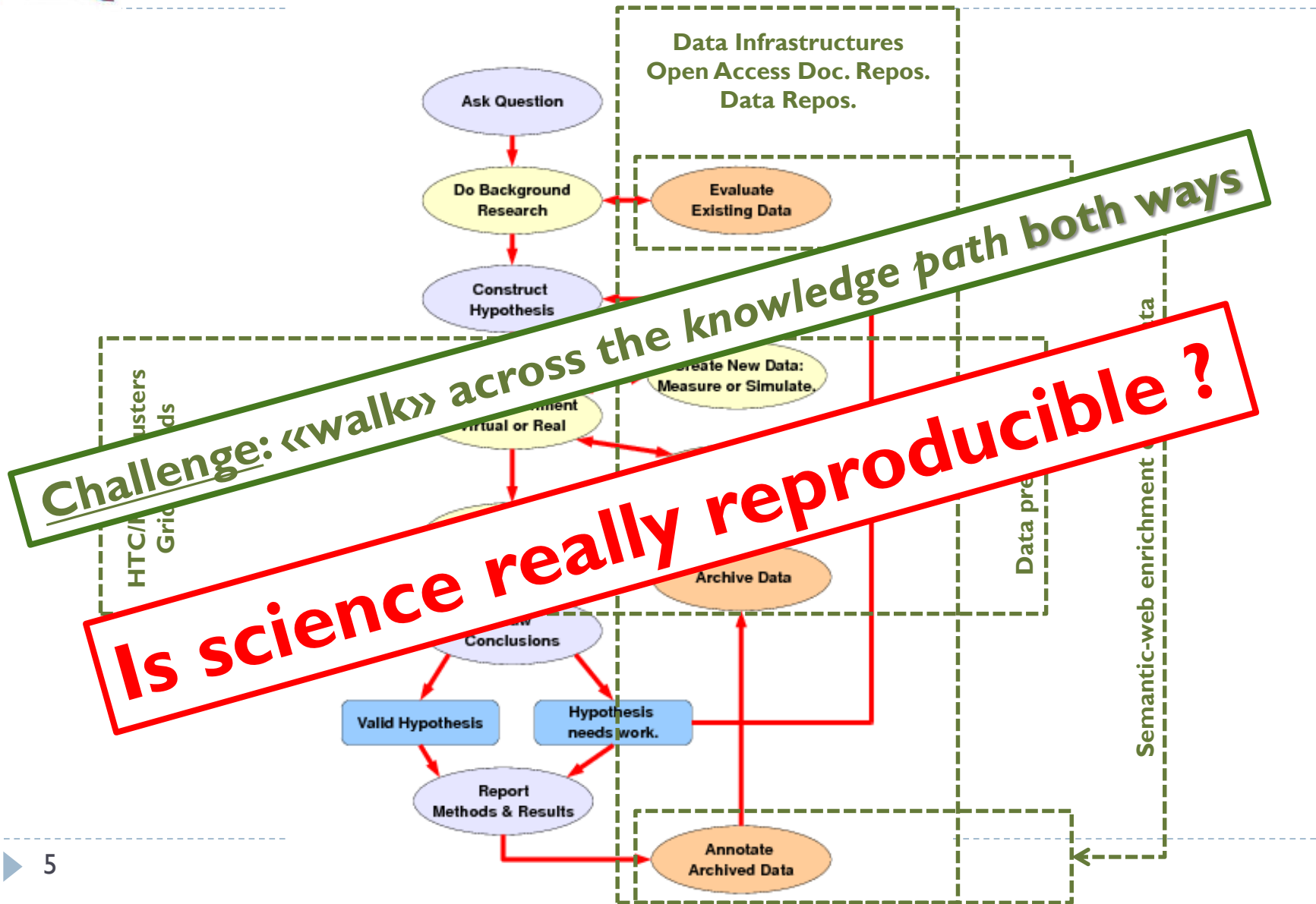


⇒ **Applications**

⇒ **Data**

⇒ **Instruments/sensors**

# How e-Infrastructures support the Scientific Method



# COMMENT

**AVIAN INFLUENZA** Shift expertise to track mutations where they emerge **p.534**

**EARTH SYSTEMS** Past climates give valuable clues to future warming **p.537**

**HISTORY OF SCIENCE** Descartes' lost letter tracked using Google **p.540**

**OBITUARY** Wylie Vale and an elusive stress hormone **p.542**



## Reproducibility crisis

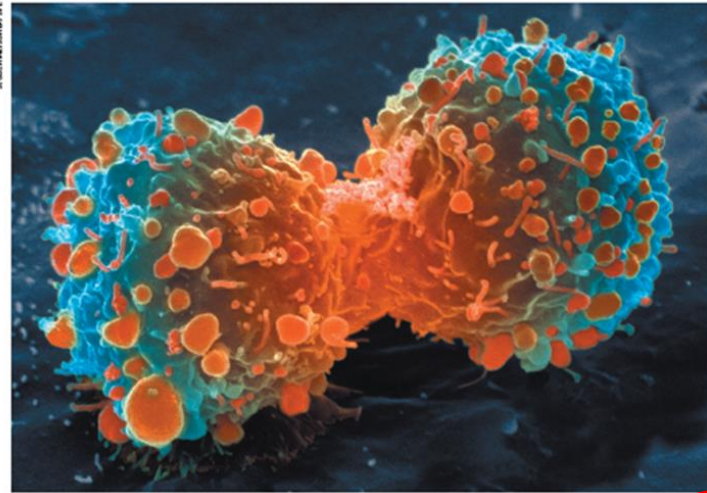
### Must try harder

*Too many sloppy mistakes are creeping into scientific papers. Lab heads must look more rigorously at the data — and at themselves.*

### Error prone

*Biologists must realize the pitfalls of working with massive amounts of data.*

If a



Many landmark findings in preclinical oncology research are not reproducible, in part because of inadequate cell lines and animal models.

Raise stan

pre

**47/53 «landmark» publications could not be replicated**  
[Begley & Ellis, Nature 483, 2012]

case for open computer programs

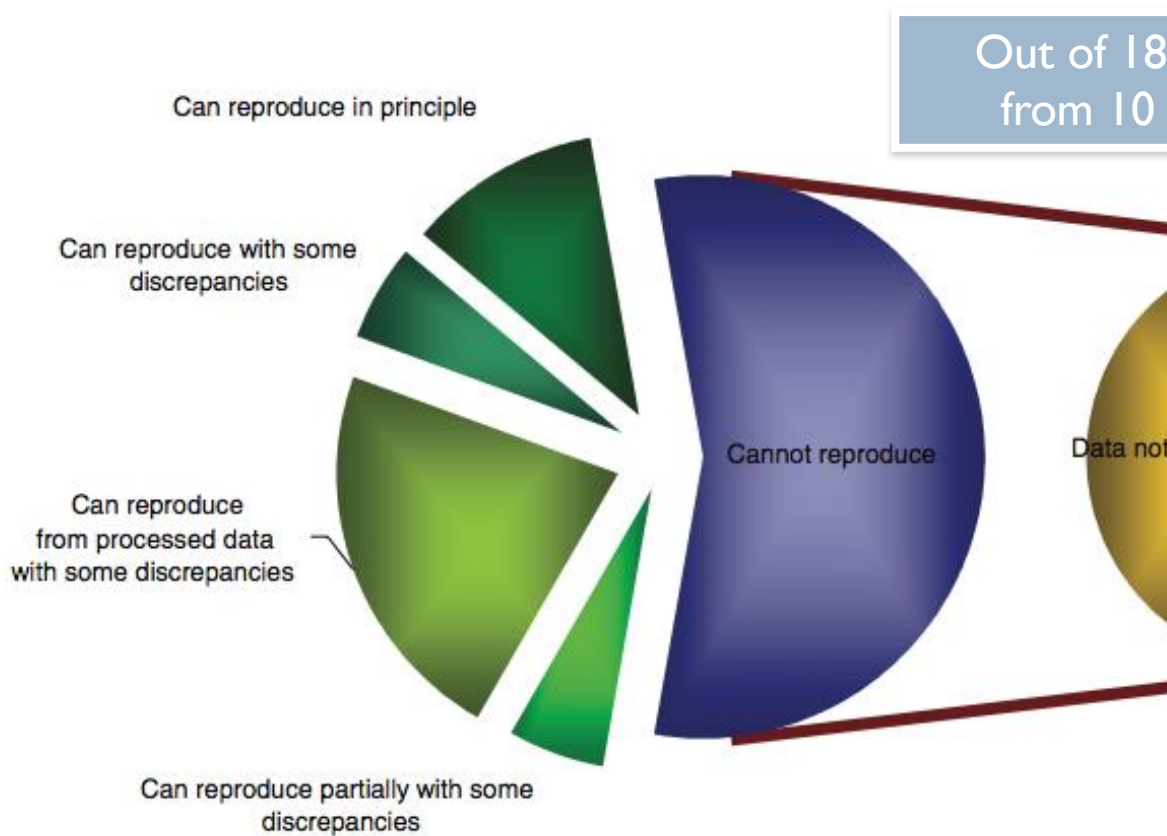
### Six red flags for suspect work

C. Glenn Begley explains how to recognize the preclinical papers in which the data won't stand up.

Know when your numbers are significant

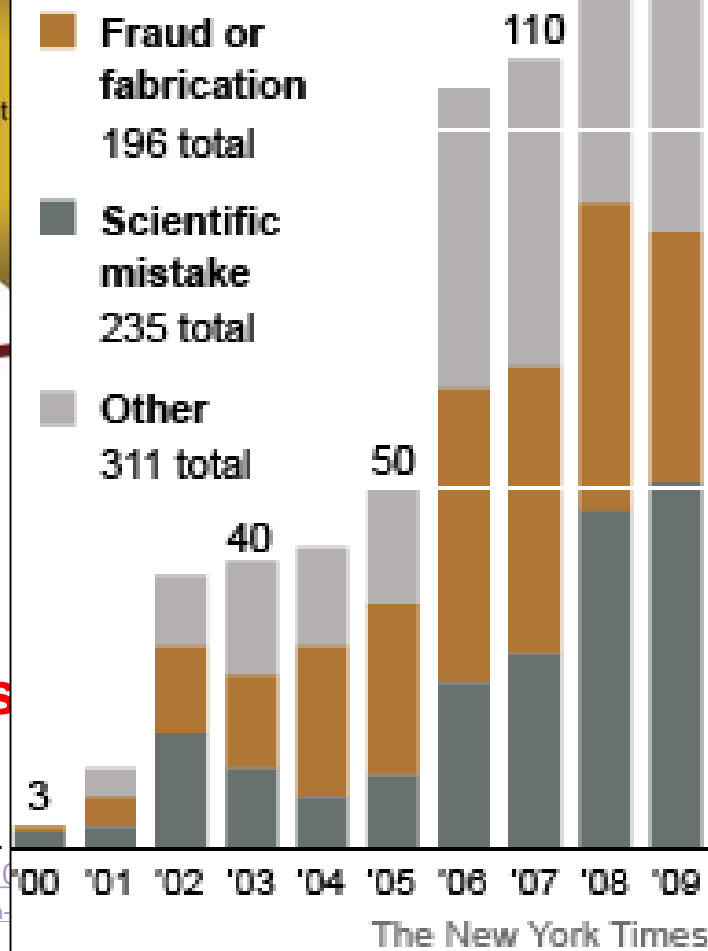


# Reproducibility crisis



## Retractions On the Rise

A study of the PubMed database found that the number of articles retracted from scientific journals increased substantially between 2000 and 2009.



## More retractions:

- >15x increase between 2000-2009
- At current rate, by 2045 as many papers published as retracted

1. Ioannidis et al., 2009. Repeatability of published microarray gene expression analyses.  
 2. Science publishing: The trouble with retractions <http://www.nature.com/news/2011/1111>  
 3. Bjorn Brembs: Open Access and the looming crisis in science <https://theconversation.com/open->

# Repeatability and Reproducibility are not all

repeat

same  
experiment  
same lab

replicate

same  
exper  
lab

experiment  
different set up

reproduce

different  
experiment  
some of same

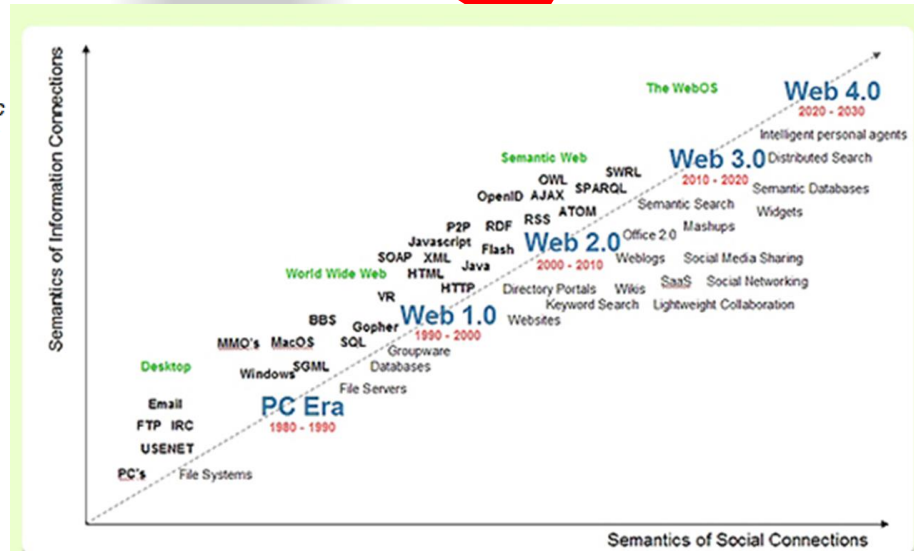
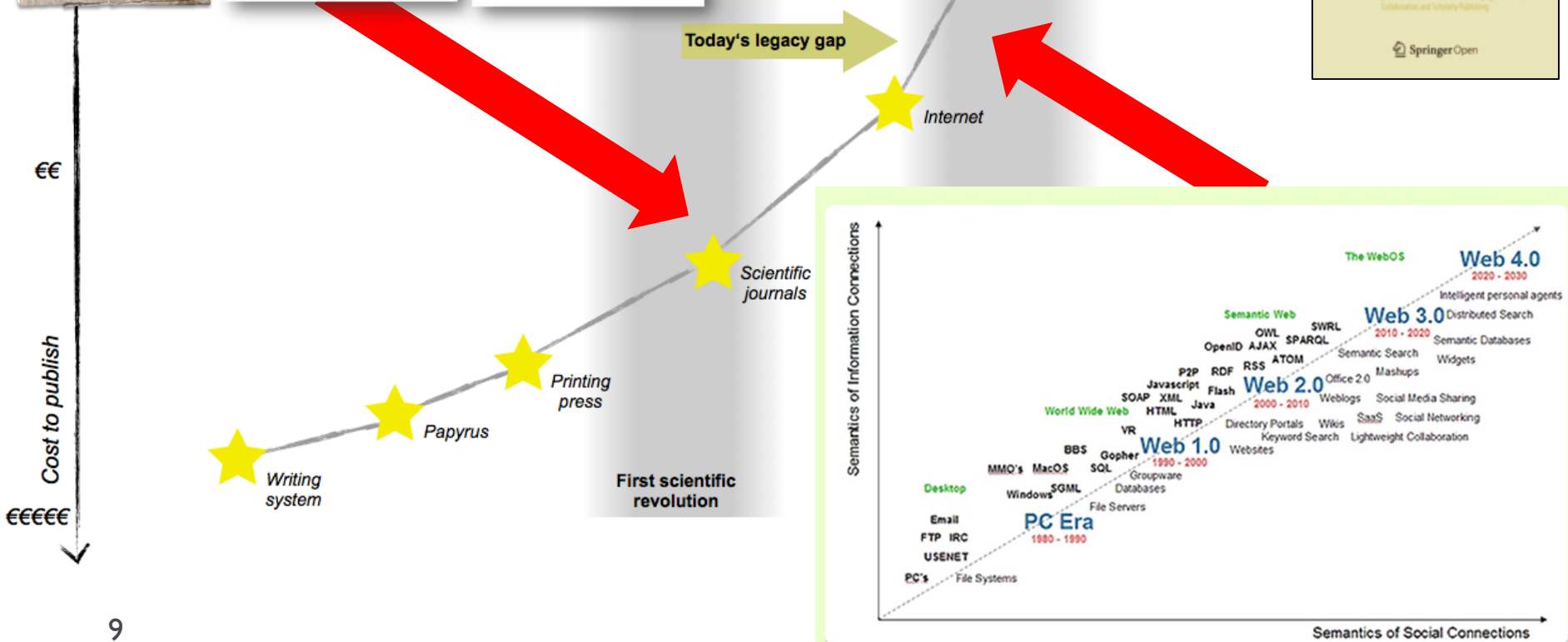
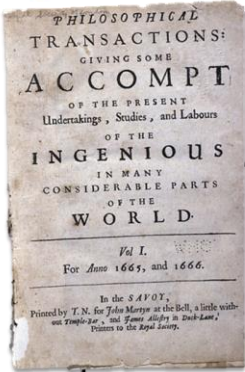
reuse

**How can modern science cope with this ?**



# Evolution of Scientific Research

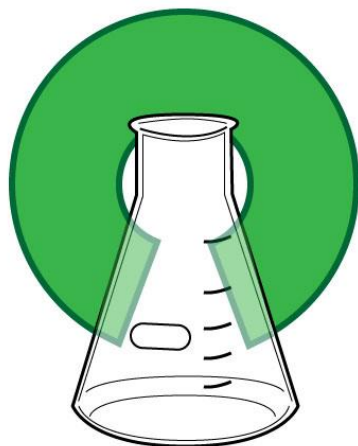
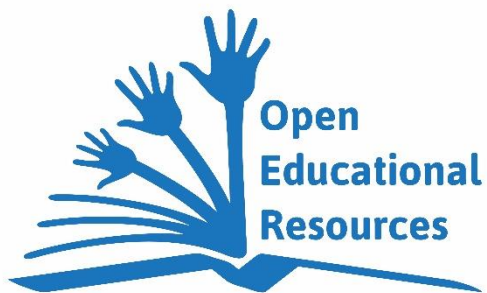
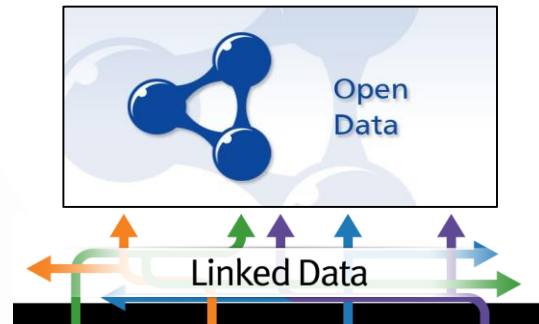
(<http://book.openingscience.org>)



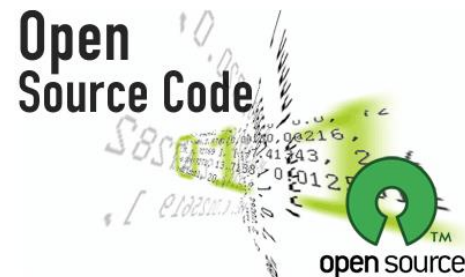
Source: Radar Networks & Nova Spivack, 2007 - [www.radarnetworks.com](http://www.radarnetworks.com)

# o-Science (Open Science)

**OPEN**  **ACCESS**

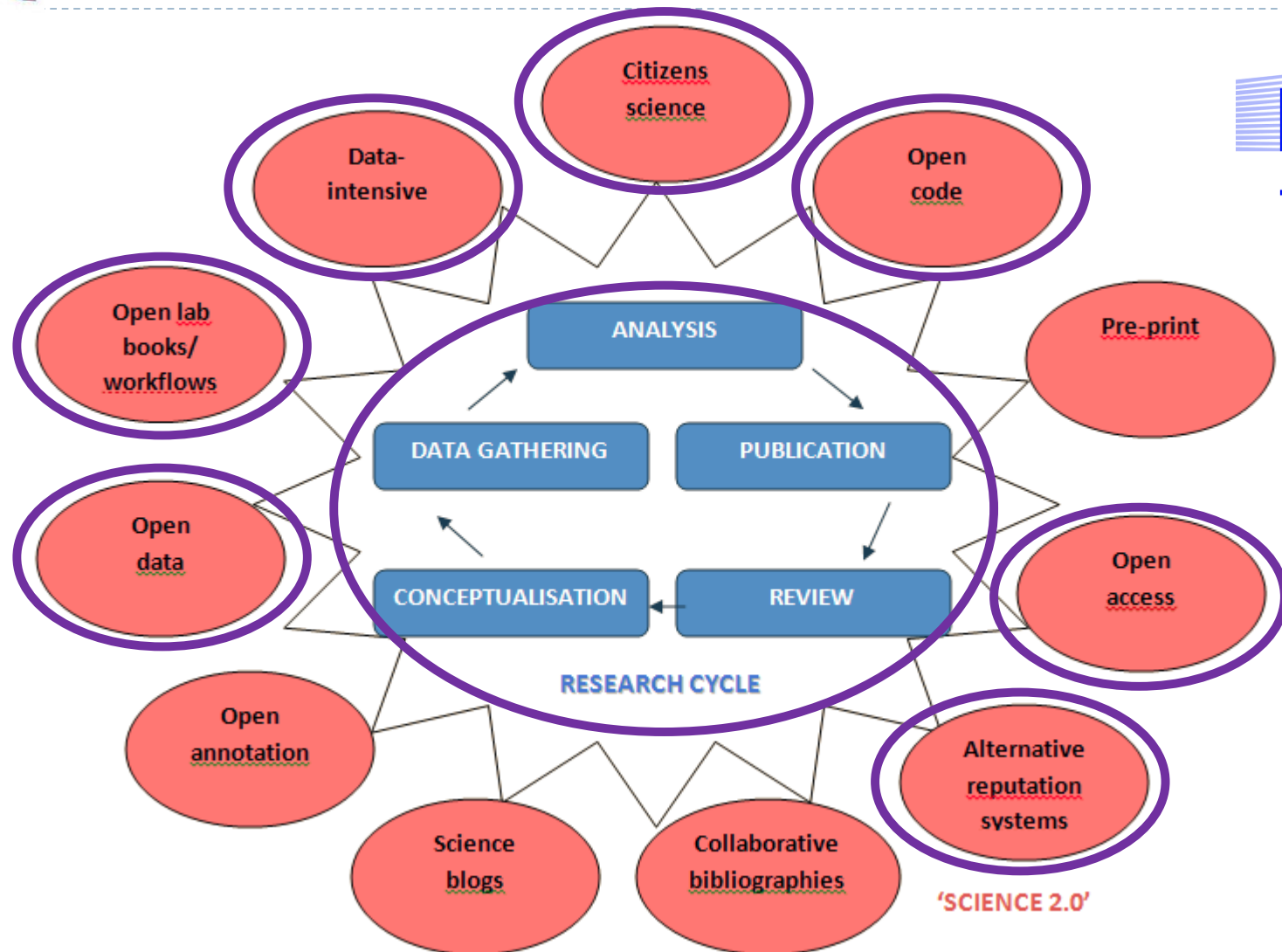


open science

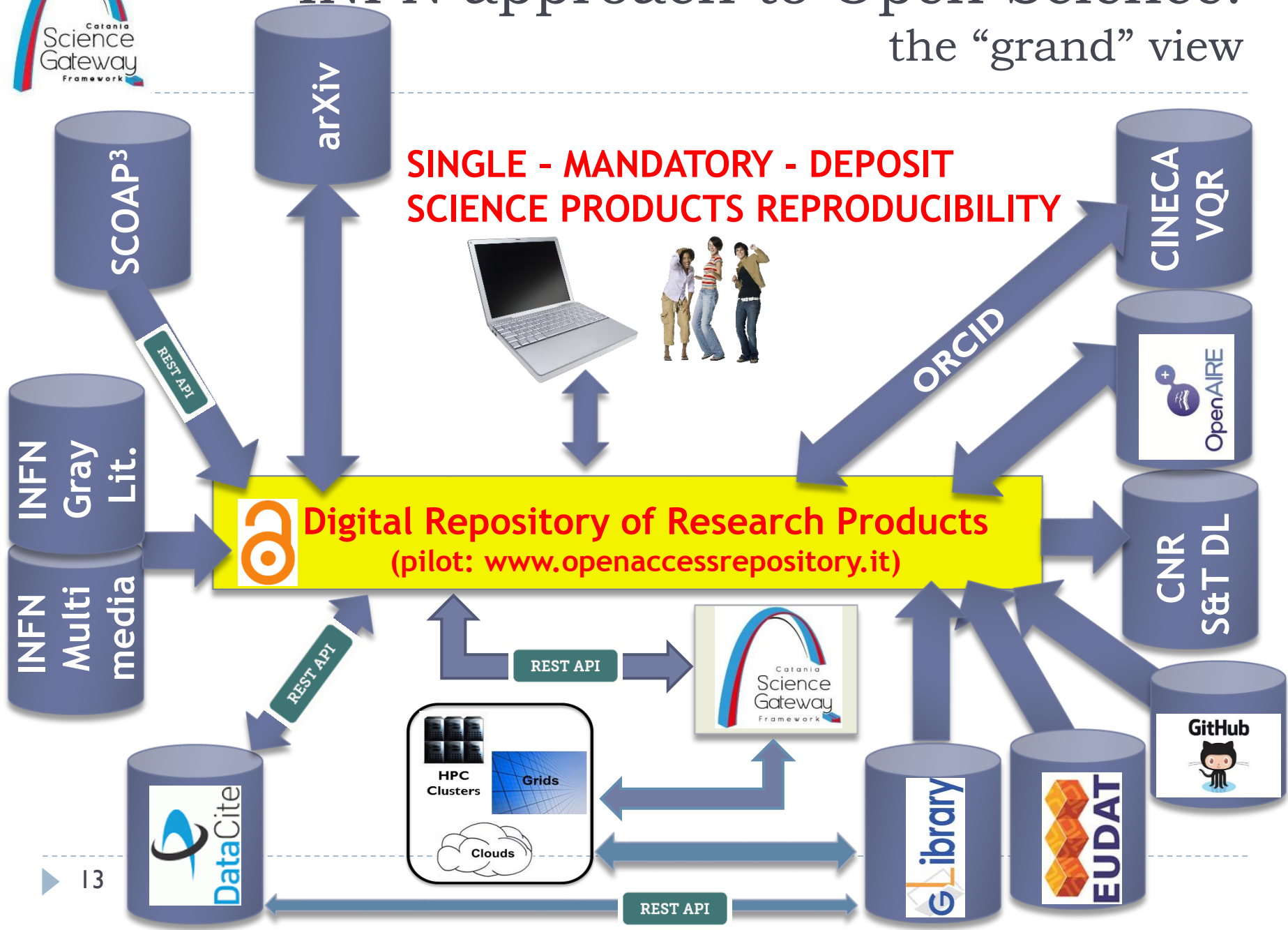


# Science in transition → towards Science 2.0

(<http://ec.europa.eu/research/consultations/science-2.0/background.pdf>)



# INFN approach to Open Science: the “grand” view





# INFN approach to Open Science: requirements, choices and motivations

---

## ▶ Requirements:

- ▶ Open source
- ▶ Standard compliant
- ▶ Well supported
- ▶ Scalable up to  $O(10^6)$ - $O(10^7)$  resources (to begin with)

## ▶ Choice:

- ▶ Invenio ([www.invenio-software.org](http://www.invenio-software.org)) + our «add-ons»

## ▶ Motivations:

- ▶ Fully compliant with OAI-PMH and Marc21 standards
- ▶ Co-developed by an international collaboration comprising institutes such as CERN, DESY, EPFL, FNAL, SLAC and used by about 30 scientific institutions worldwide
- ▶ ZENODO (OpenAIRE flagship repository) and SCOAP<sup>3</sup> repositories are based on Invenio
- ▶ The CERN Document Server (<http://cds.cern.ch/>) contains more than

▶ <sup>14</sup> 1.3 million documents




BETA


bera :: [logout](#)

Search
Submit
Personalize ▾
Help
Administration ▾

**Search 4,347 records for:**

any field ▾
Search
Browse

[Search Tips](#) :: [Advanced Search](#)

- [Audio-Video Recordings](#) (0)  
INFN (0) Others (0)
- [Datasets](#) (185)  
INFN (185) Others (0)
- [Images](#) (0)  
INFN (0) Others (0)
- [Presentations](#) (4)  
INFN (3) PSTS (0) Others (1)
- [Posters](#) (4)  
INFN (4) Others (0)
- [Publications](#) (3,968)  
INFN (1,147) PSTS (1) Others (2,820)
- [Software](#) (186)  
INFN (186) Others (0)

**Automatic ingestion in place from:**



papers





REST API



data



**ABOUT THIS SITE**

Welcome to the beta version of the Open Access Repository managed and operated by [INFN](#). Feel free to browse all the features and contents of this site as well as to upload your own open access documents and data. To sign-up or sign-in, click on [login](#).

**CERTIFICATION AND COMPLIANCE**

This site is both an OAI conforming repository and an official OpenDOAR data provider. It is also one of the official [OpenAIRE archives](#), compliant with version 3.0 of its [guidelines](#).

**SEE ALSO**

[INFN](#), [PSTS](#)



Open Access Repository :: [Search](#) :: [Submit](#) :: [Personalize](#) :: [Help](#)  
 Info :: [Terms of use](#) :: [Privacy Policy](#) :: [Support/Feedback](#)  
 Powered by [Invenio](#) v1.1.3.15-fe13-dirty  
 Maintained by INFN Catania [librarian@openaccessrepository.it](mailto:librarian@openaccessrepository.it)  
 Last updated: 20 Oct 2014, 14:03

This site is also available in the following languages:  
[English](#) [Italiano](#)  
 This is a Service Provider of:





federated authentication



# Resources' automatic harvesting & ingestion: the SCOAP<sup>3</sup> use case

## Step 1

### Call SCOAP<sup>3</sup> API

```
public static HttpMethod callSCOAP3(String date, int jrec, int num_rec) {  
    HttpMethod method = null;  
    method = new GetMethod("http://repo.scoap3.org/search?of=xm&p=datecreated:" + date + "&jrec=" + jrec + "&rg=" + num_rec);  
    return method;  
}
```

*Connect to the SCOAP<sup>3</sup> HTTP endpoint and pull required information*

## Step 2

### Get SCOAP<sup>3</sup> records

```
public static void getRecordsScoop3(String date, int jrec, int num_rec) {  
    String responseXML = null;  
    HttpClient client = new HttpClient();  
    HttpMethod method = callSCOAP3(date, jrec, num_rec);  
    try {  
        client.executeMethod(method);  
        if (method.getStatusCode() == HttpStatus.SC_OK) {  
            method.getResponseBody();  
            responseXML = convertStreamToString(method.getResponseBodyAsStream());  
            FileWriter fw = new FileWriter("MARCXML_SCOAP3_" + date + "/marcXML_scoap3_" + jrec + "_" + num_rec + ".xml");  
            fw.append(responseXML);  
            fw.close();  
        }  
    }  
}
```

*Save retrieved records in local XML file*

## Step 3

### Format SCOAP<sup>3</sup> records

*Create compliance with Open Access Repository scheme (modify MarcXML tags)*

## Step 4

### Ingest final data

*Ingest final, reformatted data in Open Access Repository*

# (Single) resources upload and DOI registration



Open Access Repository

BETA

Search

Submit

Personalize

Help

Administration

Home > Submit > Datasets, Posters, Presentations, Publications, Software > Submit New Record

## Submit New Record

Datasets, Posters, Presentations, Publications, Software

Submit New Record

page: 1

SUMMARY(2)

### Submit a resource:

Your resource will be given a reference number automatically.  
However, if it has other reference numbers, please enter them here:  
(one per line)

Digital Object Identifier

10.5072/oar.it/1412597948.03

Reserve a DOI

Type of publication

Article

Title of my paper

\*Resource Title:

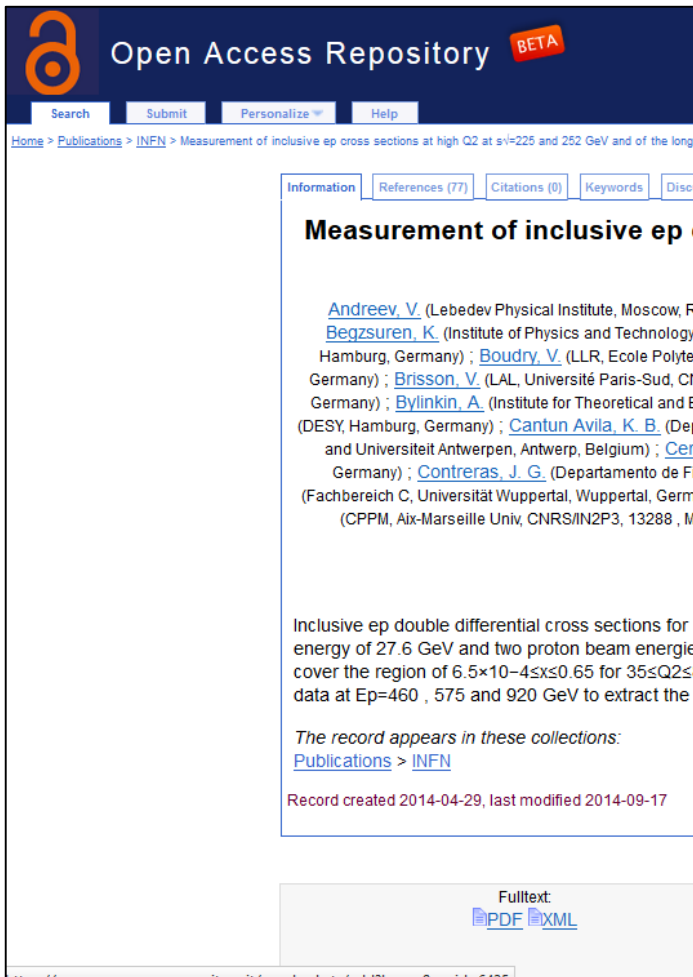
\*Author of the Resource: (one per line)

Barbera, R

openaccessrepository.it  
is a registered domain of:



# Examples of document and data resources



**Open Access Repository** BETA

Search Submit Personalize Help

Home > Publications > INFN > Measurement of inclusive ep cross sections at high Q2 at s<sup>1/2</sup>=225 and 252 GeV and of the long

Information References (77) Citations (0) Keywords Disc

## Measurement of inclusive ep cross sections at high Q<sup>2</sup> at s<sup>1/2</sup>=225 and 252 GeV and of the long

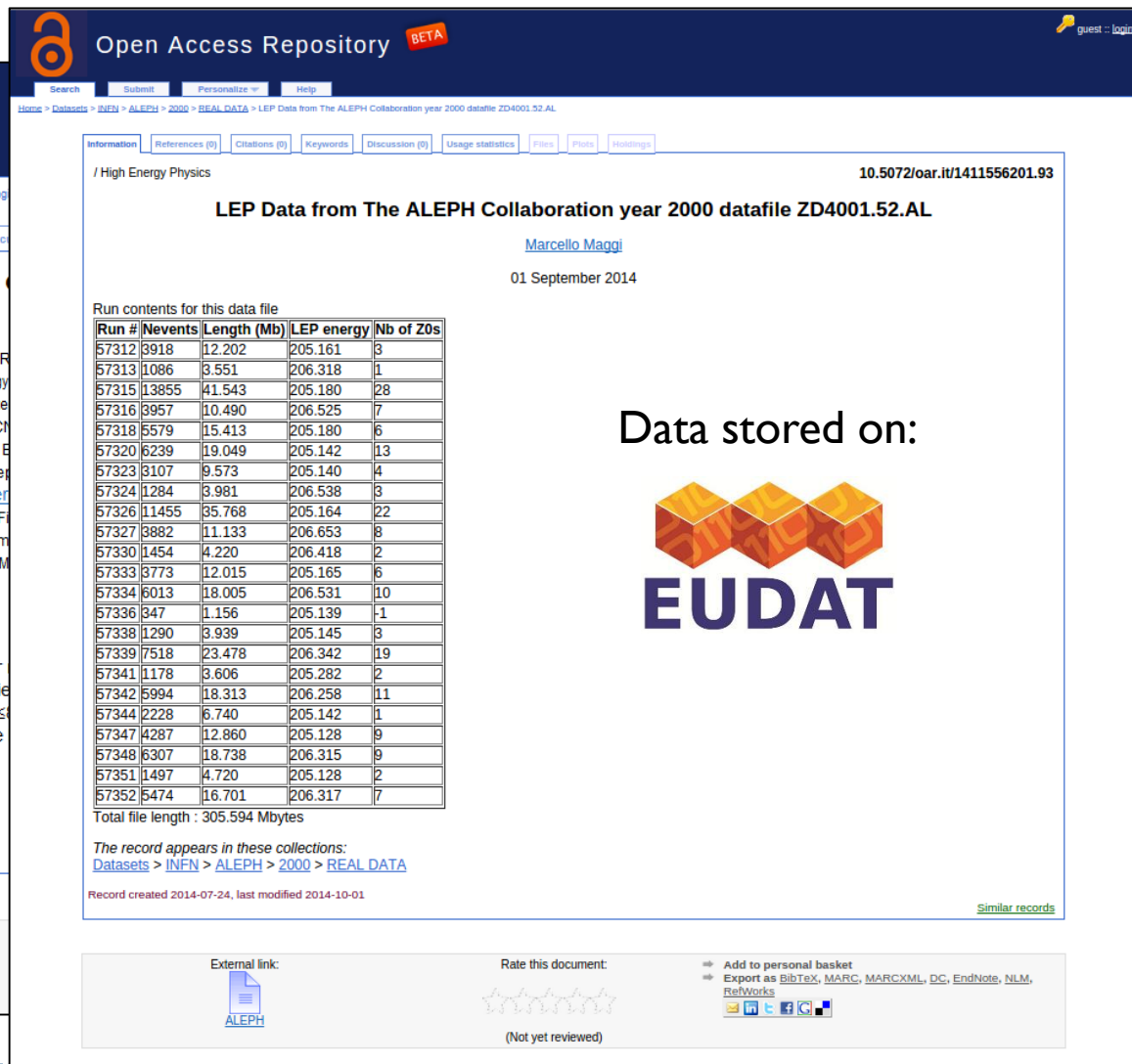
[Andreev, V.](#) (Lebedev Physical Institute, Moscow, R); [Begzsuren, K.](#) (Institute of Physics and Technology Hamburg, Germany); [Boudry, V.](#) (LLR, Ecole Polytechnique, France); [Brisson, V.](#) (LAL, Université Paris-Sud, CNRS, France); [Bylinkin, A.](#) (Institute for Theoretical and Experimental Physics, Moscow, Russia); [Cantun Avila, K. B.](#) (Department of Physics and Universiteit Antwerpen, Antwerp, Belgium); [Cerrito, M.](#) (DESY, Hamburg, Germany); [Contreras, J. G.](#) (Departamento de Física, Universidad de Sevilla, Sevilla, Spain); [Couture, J.](#) (Fachbereich C, Universität Wuppertal, Wuppertal, Germany); [Couture, J.](#) (CPPM, Aix-Marseille Univ, CNRS/IN2P3, 13288, Marseille, France)

Inclusive ep double differential cross sections for the production of a muon and a neutrino at energy of 27.6 GeV and two proton beam energies of 225 and 252 GeV. The data cover the region of 6.5×10<sup>-4</sup> ≤ x ≤ 0.65 for 35 ≤ Q<sup>2</sup> ≤ 300 GeV<sup>2</sup> and 0.05 ≤ y ≤ 0.95. Data at Ep=460, 575 and 920 GeV to extract the

The record appears in these collections:  
[Publications](#) > [INFN](#)

Record created 2014-04-29, last modified 2014-09-17

Fulltext:  
[PDF](#) [XML](#)



**Open Access Repository** BETA

Search Submit Personalize Help

Home > Datasets > INFN > ALEPH > 2000 > REAL DATA > LEP Data from The ALEPH Collaboration year 2000 datafile ZD4001.52.AL

Information References (0) Citations (0) Keywords Discussion (0) Usage statistics Files Plots Holdings

/ High Energy Physics 10.5072/oar.it/1411556201.93

## LEP Data from The ALEPH Collaboration year 2000 datafile ZD4001.52.AL

[Marcello Maggi](#)  
01 September 2014

Run contents for this data file

Run #	Nevents	Length (Mb)	LEP energy	Nb of Z0s
57312	3918	12.202	205.161	3
57313	1086	3.551	206.318	1
57315	13855	41.543	205.180	28
57316	3957	10.490	206.525	7
57318	5579	15.413	205.180	6
57320	6239	19.049	205.142	13
57323	3107	9.573	205.140	4
57324	1284	3.981	206.538	3
57326	11455	35.768	205.164	22
57327	3882	11.133	206.653	8
57330	1454	4.220	206.418	2
57333	3773	12.015	205.165	6
57334	6013	18.005	206.531	10
57336	347	1.156	205.139	-1
57338	1290	3.939	205.145	3
57339	7518	23.478	206.342	19
57341	1178	3.606	205.282	2
57342	5994	18.313	206.258	11
57344	2228	6.740	205.142	1
57347	4287	12.860	205.128	9
57348	6307	18.738	206.315	9
57351	1497	4.720	205.128	2
57352	5474	16.701	206.317	7

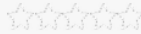
Total file length : 305.594 Mbytes


The record appears in these collections:  
[Datasets](#) > [INFN](#) > [ALEPH](#) > [2000](#) > [REAL DATA](#)

Record created 2014-07-24, last modified 2014-10-01

[Similar records](#)

External link: [ALEPH](#)

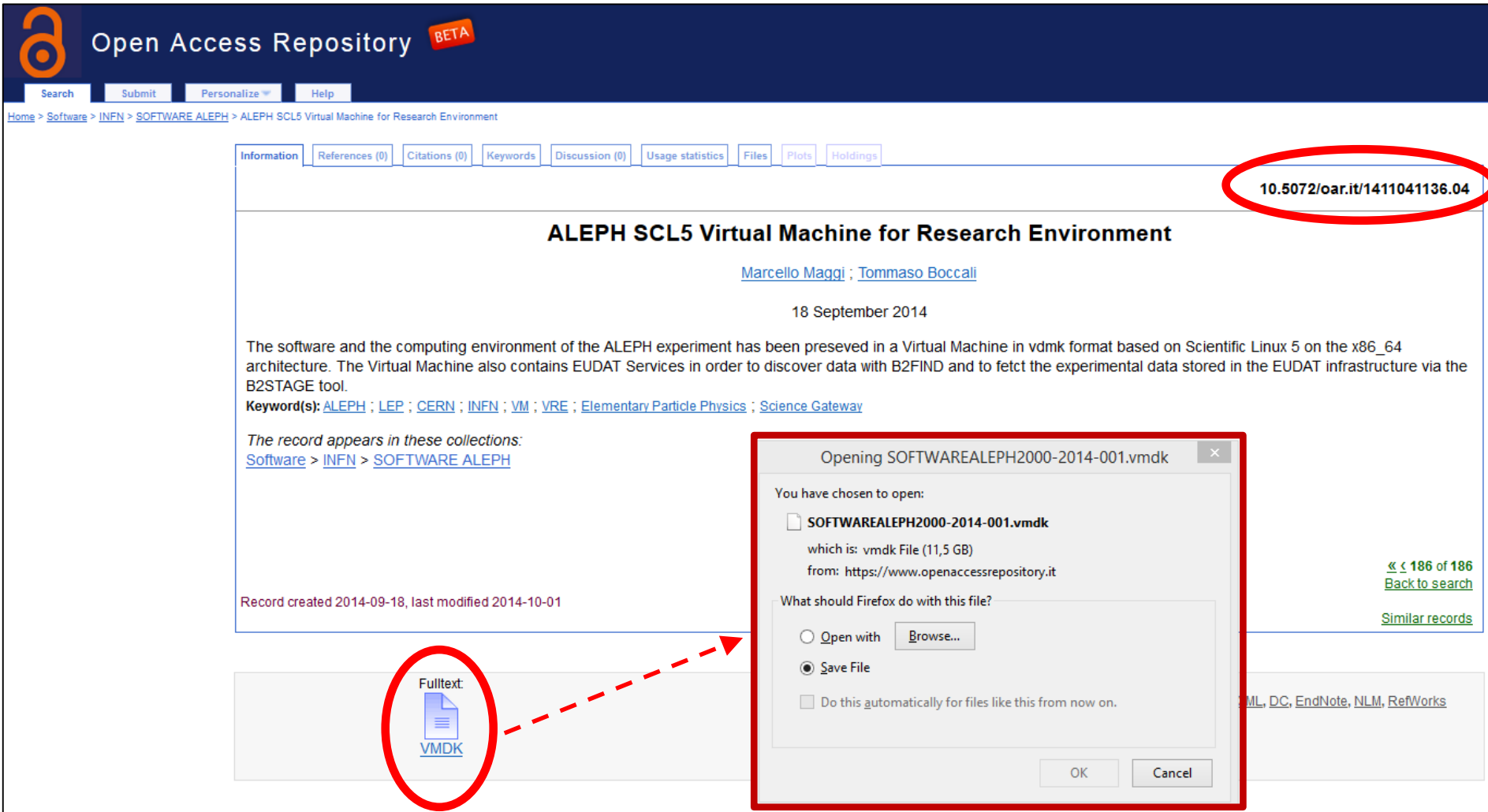
Rate this document:  (Not yet reviewed)

Add to personal basket  
 Export as BibTeX, MARC, MARCXML, DC, EndNote, NLM, RefWorks  


Data stored on:



# Example of software resources: the ALEPH Virtual Research Environment



**Open Access Repository** BETA

Search Submit Personalize Help

Home > Software > INFN > SOFTWARE ALEPH > ALEPH SCL5 Virtual Machine for Research Environment

Information References (0) Citations (0) Keywords Discussion (0) Usage statistics Files Plots Holdings

**10.5072/oar.it/1411041136.04**

## ALEPH SCL5 Virtual Machine for Research Environment

[Marcello Maggi](#) ; [Tommaso Boccali](#)


18 September 2014

The software and the computing environment of the ALEPH experiment has been preseved in a Virtual Machine in vmdk format based on Scientific Linux 5 on the x86\_64 architecture. The Virtual Machine also contains EUDAT Services in order to discover data with B2FIND and to fetct the experimental data stored in the EUDAT infrastructure via the B2STAGE tool.

**Keyword(s):** [ALEPH](#) ; [LEP](#) ; [CERN](#) ; [INFN](#) ; [VM](#) ; [VRE](#) ; [Elementary Particle Physics](#) ; [Science Gateway](#)

*The record appears in these collections:*  
[Software](#) > [INFN](#) > [SOFTWARE ALEPH](#)

Record created 2014-09-18, last modified 2014-10-01

Fulltext:  
  
VMDK

Opening SOFTWAREALEPH2000-2014-001.vmdk

You have chosen to open:

- SOFTWAREALEPH2000-2014-001.vmdk  
which is: vmdk File (11,5 GB)  
from: <https://www.openaccessrepository.it>

What should Firefox do with this file?

- Open with [Browse...](#)
- Save File
- Do this automatically for files like this from now on.

OK Cancel

« 186 of 186  
[Back to search](#)  
[Similar records](#)

ML, DC, EndNote, NLM, RefWorks



## Registration Record

element name	element value
Base URL	<a href="http://www.openaccessrepository.it/oai2d">http://www.openaccessrepository.it/oai2d</a>
Repository Name	Open Access Repository
Protocol Version	2.0
Email	librarian@openaccessrepository.it
Registration Date	2014-05-06T10:47:29Z
Date Last Validated	Tue May 6 10:47:29 2014
OAI Repository ID	www.openaccessrepository.it

If you are the maintainer of this repository, you may to update the information recorded to match new information exposed via the Identify response by running the validation/registration process again. Go to the [validation page](#) and select "Register this site".

*Fri May 23 09:57:55 2014*

"Open Access Repository"

Any Subject Area  Any Content Type  Any Repository Type

Any Country  Any Language  Any Software

Full records  1  per page. Sort by:

To search the *contents* of the repositories listed in OpenDOAR, please see our [Content Search](#) page.

Result 1 of 1.

Page: << Previous 1 Next >>

## [Open Access Repository](#)

**URL:** <http://www.openaccessrepository.it/>

**Organisation:** INFN

**Address:** Via Santa Sofia, 62, Catania

**Country:** Italy

**Location:** *Latitude:* 37.526700 & *Longitude:* 15.073100, [Google Map](#)

**Description:** Open Access repository of INFN publications and data, to be eventually extended to other organisations. The interface is in English.

**Type:** Institutional - Operational

**Size:** 2026 items (2014-05-19)

**OAI-PMH:** <http://www.openaccessrepository.it/oai2d>

**Software:** invenio

**Subjects:** Physics and Astronomy

**Content:** Articles; Conferences; Datasets; Multimedia; Software; Special

**Languages:** English


**Contacts:** 1. Roberto Barbera ([roberto.barbera@ct.infn.it](mailto:roberto.barbera@ct.infn.it)), Administrator  
2. Roberto Barbera ([librarian@openaccessrepository.it](mailto:librarian@openaccessrepository.it)), Administrator

**OpenDOAR ID:** 3061, *Last reviewed:* 2014-05-16, [Suggest an update for this record](#)  
Link to this record: <http://opendoar.org/id/3061/>



# Visibility and compliance (compliant with OpenAIRE guidelines)

## Browse Compatibility Test Results

 Refresh (F5) to see the progress of the latest activities.



Guidelines 3.0

Type	Status	Score	User	Started	Duration	Repository	Rule Set	Reports
OAI Content Validation	finished	100	roberto.barbera@ct.infn.it	2014-05-22 15:53:18	5 mins 8 s ecs	http://www.openaccessrepository.it/oai2d	Custom	<a href="#">Show Results</a>
OAI Usage Validation	finished	100	roberto.barbera@ct.infn.it	2014-05-22 15:53:18	1 secs	http://www.openaccessrepository.it/oai2d	Custom	<a href="#">Show Results</a>

 **Re: Registration of a data provider**

**Repository Technical Support**

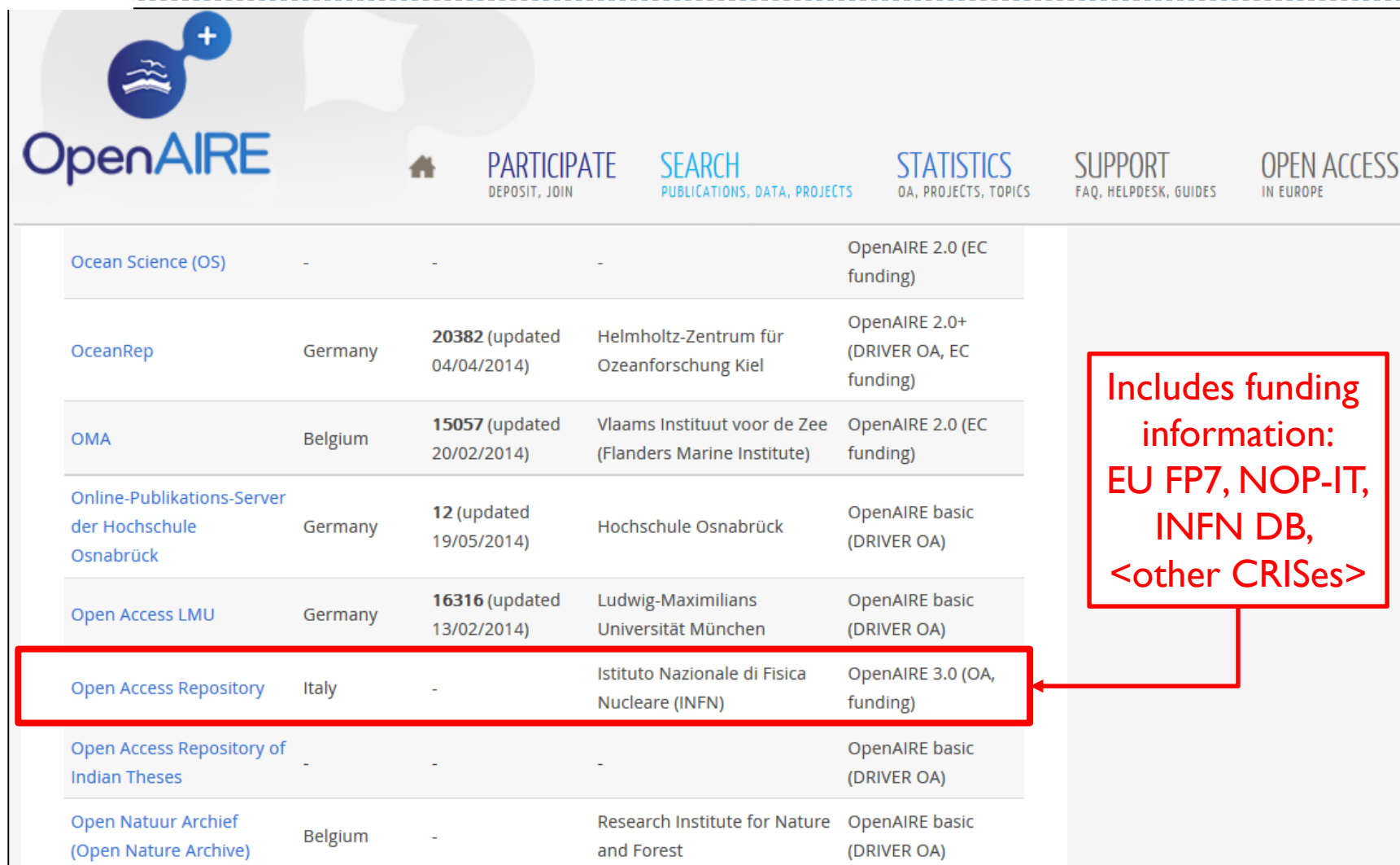
Dear Roberto,

23 May 2014, 08:51

your repository seems to have successfully passed the registration test.

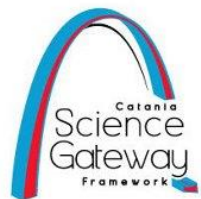
" A new repository/journal is ready to be added in the list of OpenAIRE compliant repositories/journals:  
Open Access Repository

# Visibility and compliance (registered as an OpenAIRE data provider)



OpenAIRE		PARTICIPATE		SEARCH		STATISTICS		SUPPORT		OPEN ACCESS	
		DEPOSIT, JOIN		PUBLICATIONS, DATA, PROJECTS		OA, PROJECTS, TOPICS		FAQ, HELPDESK, GUIDES		IN EUROPE	
Ocean Science (OS)	-	-	-	-	-	-	OpenAIRE 2.0 (EC funding)				
OceanRep	Germany	20382 (updated 04/04/2014)	Helmholtz-Zentrum für Ozeanforschung Kiel				OpenAIRE 2.0+ (DRIVER OA, EC funding)				
OMA	Belgium	15057 (updated 20/02/2014)	Vlaams Instituut voor de Zee (Flanders Marine Institute)				OpenAIRE 2.0 (EC funding)				
Online-Publikations-Server der Hochschule Osnabrück	Germany	12 (updated 19/05/2014)	Hochschule Osnabrück				OpenAIRE basic (DRIVER OA)				
Open Access LMU	Germany	16316 (updated 13/02/2014)	Ludwig-Maximilians Universität München				OpenAIRE basic (DRIVER OA)				
Open Access Repository	Italy	-	Istituto Nazionale di Fisica Nucleare (INFN)				OpenAIRE 3.0 (OA, funding)				
Open Access Repository of Indian Theses	-	-	-				OpenAIRE basic (DRIVER OA)				
Open Natuur Archief (Open Nature Archive)	Belgium	-	Research Institute for Nature and Forest				OpenAIRE basic (DRIVER OA)				

Includes funding information:  
EU FP7, NOP-IT,  
INFN DB,  
<other CRISes>



# The Catania Science Gateway Framework (CSGFF) in a nutshell

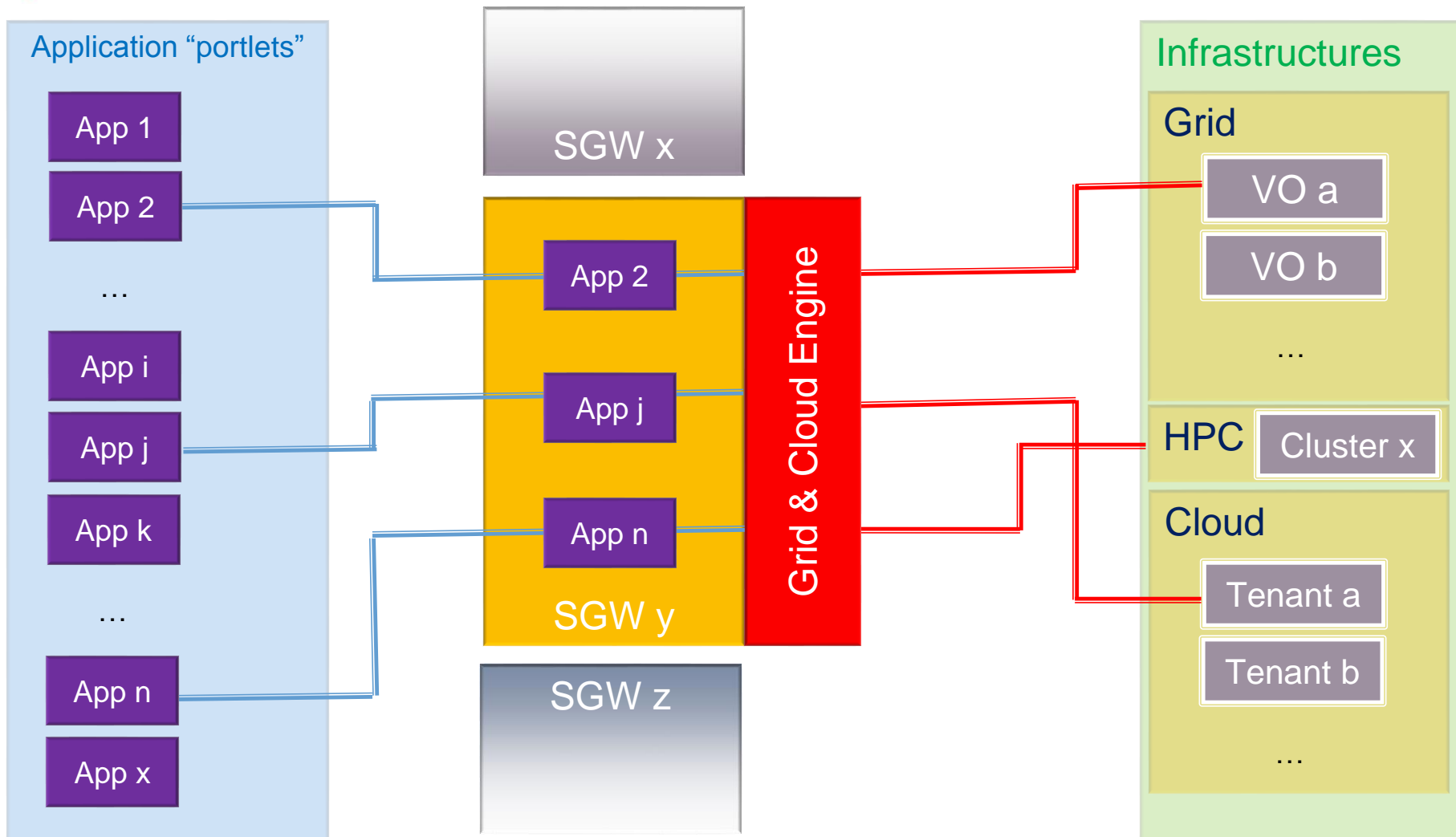


<http://www.catania-science-gateways.it>

To better access e-Infrastructures Worldwide

- ▶ Born in 2010 to hide Grid – and now Cloud - complexity (especially security-wise)
- ▶ Designed to be:
  - ▶ Sustainable (Fully based on standards)
  - ▶ Scalable (e.g., through Glassfish)
  - ▶ Secure (integrated AAAAI)
  - ▶ Interoperable (one system → many infrastructures)
  - ▶ Accessible anytime from anywhere (including mobile devices)

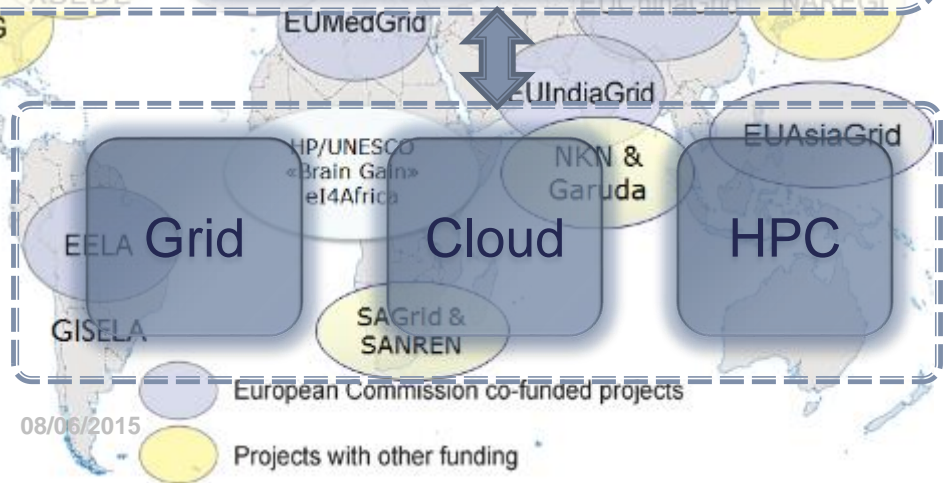
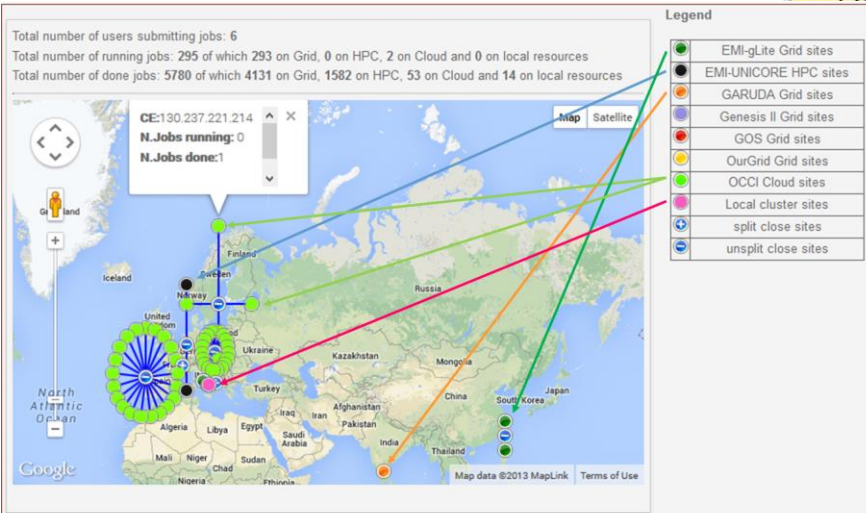
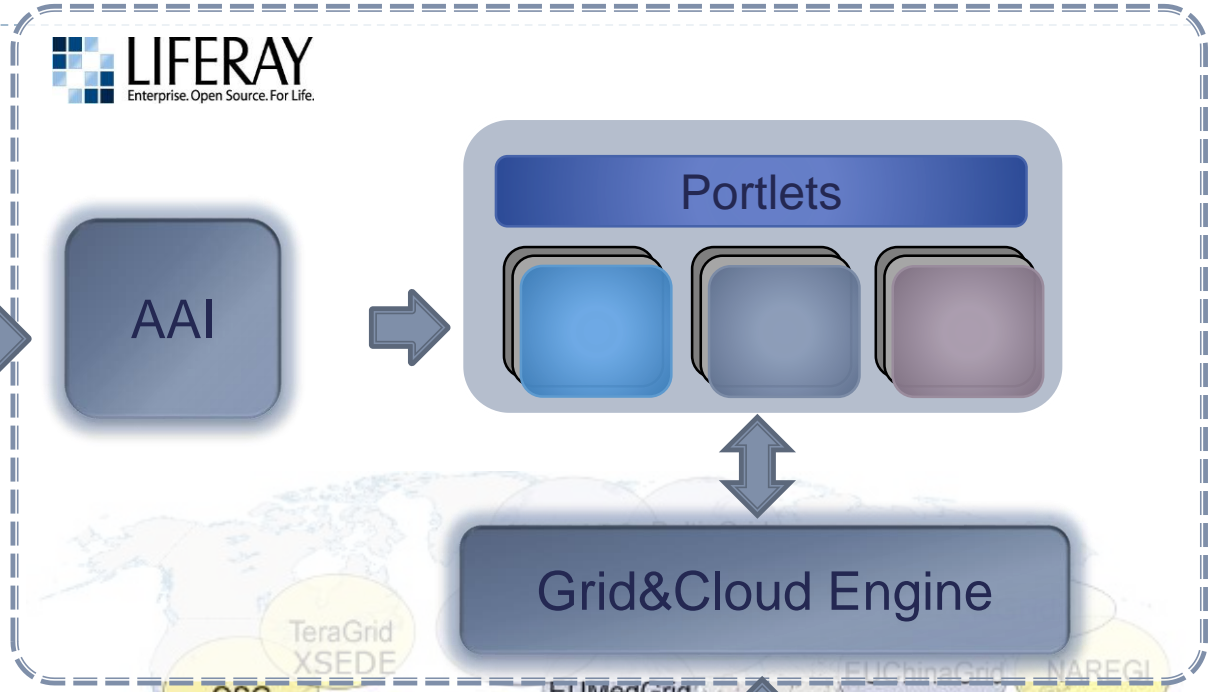
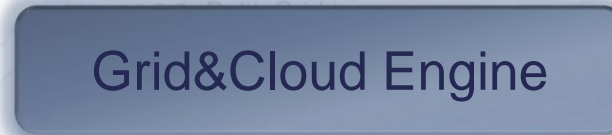
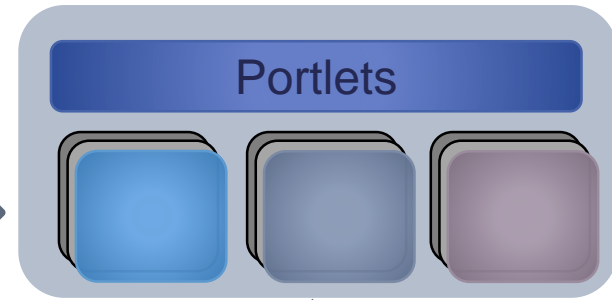
# CSGF Architecture

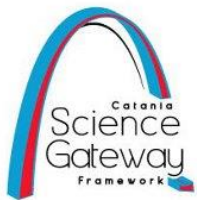


# CSGF Components

Users having different roles and privileges

- Administrators
- Power users
- Basic users
- VRC members
- etc,





access.egi.eu

EGI  
UMP

OpenID



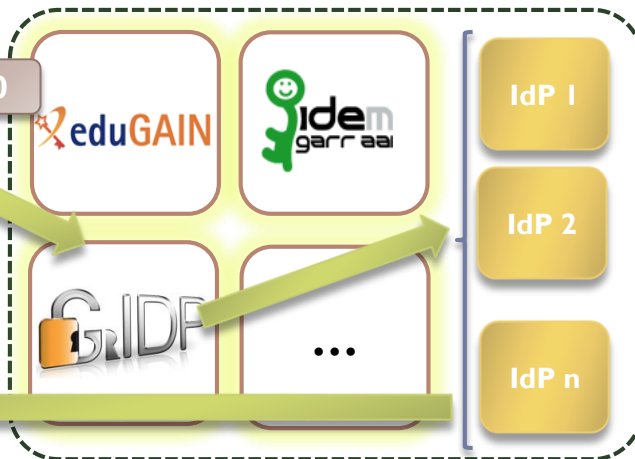
# AAI in the CSGF

AuthN/AuthZ  
handled separately



SAML 2.0

Identity Federations



1. Sign In

brunor

.....

Change password

SIGN IN

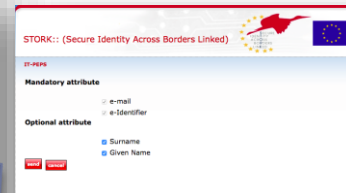
REGISTER



2. Select eID or Your institute IdFs



3. Select your IdF and IdP



eid-stork.eu

4. Use IdP to get authorized by IdP




# Support for the most diverse and widest possible communities of users (1/3)

Sign In

Select the origin country


Your institute

Contacts: [account](#) | [privacy](#)  
[Privacy Policy](#) | [Terms](#)



Return to Full Page

Powered By [Liferay](#)  
with [Catania SG Framework](#)  
is a Service Provider of:



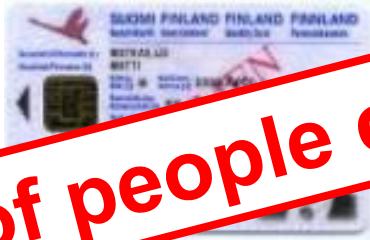
Sign In

28

Sign In Your institute



# Support for the most diverse and widest possible communities of users (2/3)



Väestorekisterikeskus  
Population Register Centre



POŠTA<sup>R</sup>CA  
Certifikatska agencija Pošte Slovenije



Agència Catalana  
de Certificació

Millions of people can become potential users

A new way of thinking about citizen science

# The OAR Knowledge Workflow



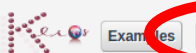
# The OAR Knowledge Workflow: data search & discovery

Co-ordination & H  
for Research and E

Home Project User Pages

Semantic Search on Linked Data

Here you can search in more than  
Document Repositories and Data  
Repositories. Click on the Exampl



Results Graphs

Records found. Displaying 1 to 20

Title

Analysis with the CHAIN-REDS S

Author

Roberto Barbera; Maggi Marcello

Description

This is the software of a JSR 286  
the title. Click on the External link

Check this record on Google Scholar

Repository

[Open Access Repository](#)

[More Info](#)

## Analysis with the CHAIN-REDS Science Gateway of LEP Data from the ALEPH Collaboration year 2000 datafile ZD4001.17.AL

### General Information

Authors: RobertoBarbera ; MaggiMarcello() ; BrunoRiccardo() ; RicceriRita() ; InseraGiuseppina() ; CarrubbaCarla()

Description : This is the software of a JSR 286 compliant portlet installed on the CHAIN-REDS Science Gateway that allows the analysis of the dataset mentioned in the title. Click on the External link below to be automatically re-directed to the execution page of the portlet on the Science Gateway.

Identifier : irods://data.repo.cineca.it:1247/CINECA01/home/EUDAT\_STAFF/Aleph\_Test/ZD4001.17.AL

Identifier : <http://www.openaccessrepository.it/record/7343>

Subject : High Energy Physics

Language : eng



Date Information ↗

Dataset Information ↗

Repository Information ↗

Information from Google Scholar ↗



# The OAR Knowledge Workflow: data inspection

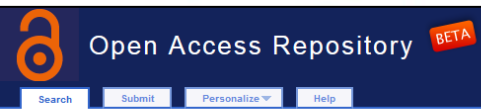


Home > Datasets > INFN > ALEPH > 2000 > REAL DATA > LEP Data from The ALEPH Collaboration year 2000 datafile ZD4001.52.AL > References

Information | **References (3)** | Citations (0) | Keywords | Discussion (0) | Usage statistics | Files | Plots | Holdings

**LEP Data from The ALEPH Collaboration year 2000 datafile ZD4001.52.AL - Marcello Maggi - REALDATA-2014-001**

- [1] [\[Search for neutral Higgs bosons\]](#)
- [2] [\[ALEPH SCL5 Virtual Machine\]](#)
- [3] [\[Analysis with the CHAIN-REDS Science Gateway\]](#)



Home > Software > INFN > SOFTWARE ALEPH > Analysis with the CHAIN-REDS Science Gateway of LEP Data from the ALEPH Collaboration year 2000 datafile ZD4001.17.AL

Information | References (0) | Citations (0) | Keywords | Discussion (0) | Usage statistics | Files | Plots | Holdings

/ High Energy Physics 10.5072/oar.it/1412595236.96

## Analysis with the CHAIN-REDS Science Gateway of LEP Data from the ALEPH Collaboration year 2000 datafile ZD4001.17.AL

[Roberto Barbera](#) ; [Bruno Riccardo](#) ; [Carrubba Carla](#) ; [Inserra Giuseppina](#) ; [Maggi Marcello](#) ; [Ricceri Rita](#)


This is the software of a JSR 286 compliant "portlet" installed on the [CHAIN-REDS Science Gateway](#) that allows the analysis of the dataset mentioned in the title. Click on the External link below to be automatically re-directed to the execution page of the portlet on the Science Gateway.

Keyword(s): [ALEPH](#) ; [Science Gateway](#) ; [Elementary Particle Physics](#) ; [INFN](#) ; [CERN](#)

The record appears in these collections:  
[Software](#) > [INFN](#) > [SOFTWARE ALEPH](#)

Record created 2014-10-01, last modified 2014-10-06 [Similar records](#)

External link:  **RUN PAGE**

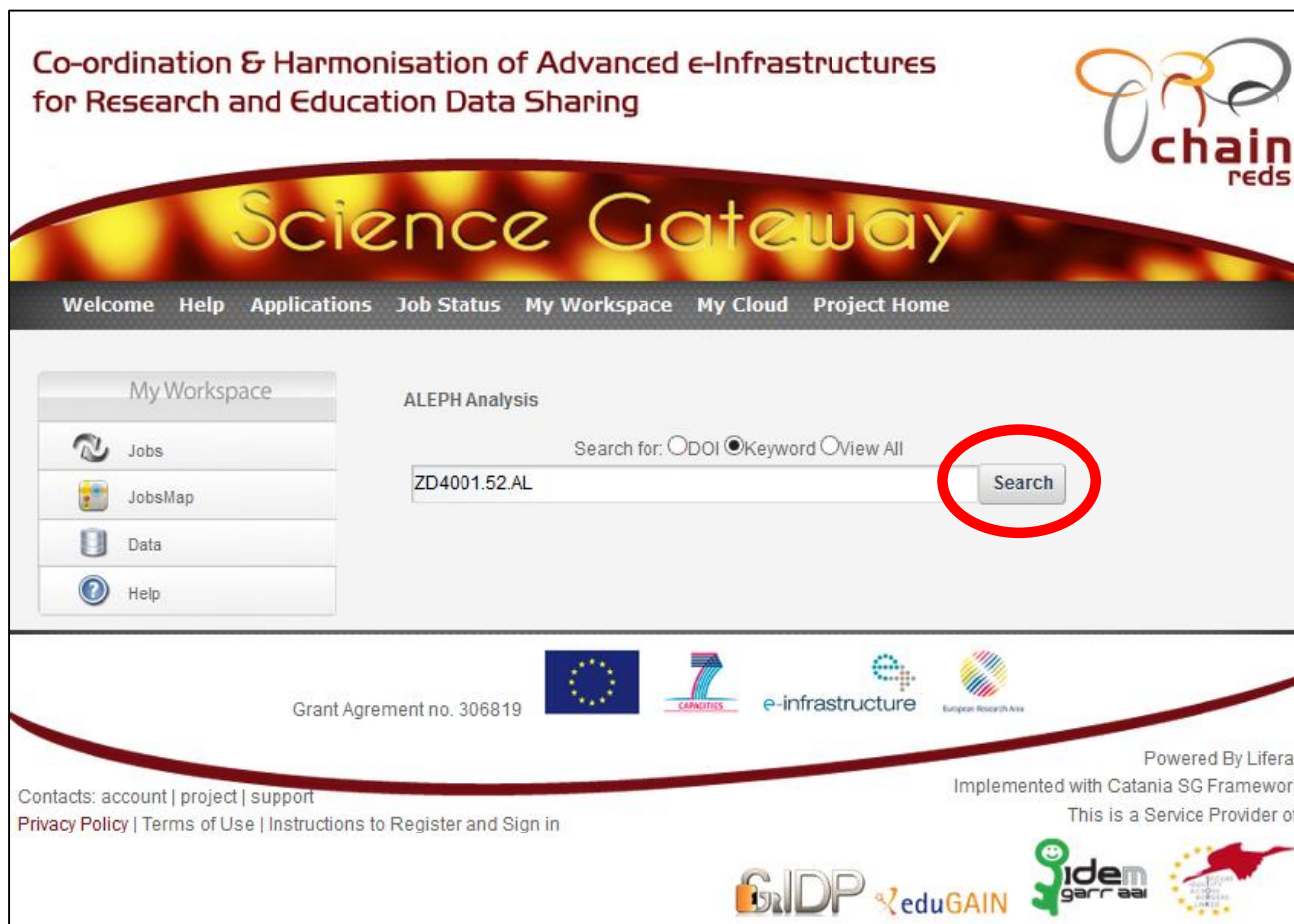
Rate this document:  (Not yet reviewed)

[Add to personal basket](#)  
[Export as BibTeX, MARC, MARCXML, DC, EndNote, NLM, RefWorks](#)  


1. From OAR it is possible to select an "analysis" as simply as any other resources in the archive
2. Clicking on **RUN PAGE**, the researcher can either reproduce or extend that particular analysis using a Science Gateway

# The OAR Knowledge Workflow: data analysis (1/2)

The Science Gateway collects from OAR, and allows user browse, metadata associated to the dataset(s) associated to that particular analysis



Co-ordination & Harmonisation of Advanced e-Infrastructures  
for Research and Education Data Sharing

chain  
reds

## Science Gateway

Welcome Help Applications Job Status My Workspace My Cloud Project Home

My Workspace

- Jobs
- JobsMap
- Data
- Help

ALEPH Analysis

Search for:  DOI  Keyword  View All

ZD4001.52.AL

Grant Agreement no. 306819

Powered By Liferay  
Implemented with Catania SG Framework  
This is a Service Provider of.

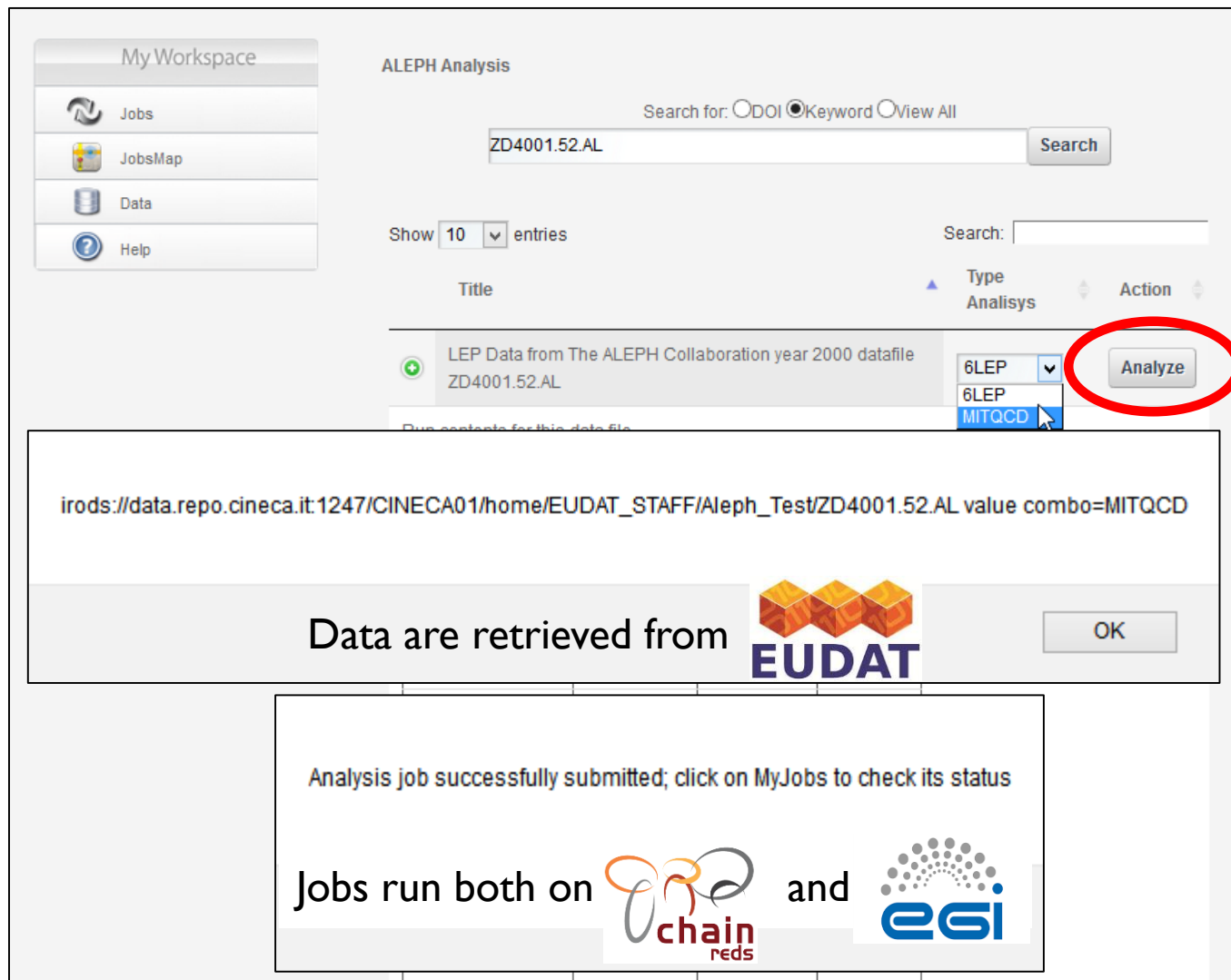
Contacts: account | project | support  
Privacy Policy | Terms of Use | Instructions to Register and Sign in

SRDP eduGAIN idem garr aai

# The OAR Knowledge Workflow: data analysis (2/2)

Using the JSAGA adaptor for all OCCI-compliant cloud-middlewares, the Science Gateway starts a dedicated VM already configured with all the experiment software

Both the CHAIN-REDS Cloud Testbed and the EGI Federated Cloud can be used as e-Infrastructures



The screenshot shows the 'ALEPH Analysis' interface. On the left is a 'My Workspace' sidebar with 'Jobs', 'JobsMap', 'Data', and 'Help' options. The main area has a search bar with 'ZD4001.52.AL' and a search button. Below the search bar, there's a 'Show 10 entries' dropdown and a table of results. The table has columns for 'Title', 'Type', and 'Action'. One entry is visible: 'LEP Data from The ALEPH Collaboration year 2000 datafile ZD4001.52.AL'. The 'Type' column has a dropdown menu with '6LEP' selected and 'MITQCD' highlighted. The 'Action' column has an 'Analyze' button circled in red. Below the table, there's a message box: 'irods://data.repo.cineca.it:1247/CINECA01/home/EUDAT\_STAFF/Aleph\_Test/ZD4001.52.AL value combo=MITQCD'. Below that, a message says 'Data are retrieved from EUDAT' with an 'OK' button. At the bottom, a message says 'Analysis job successfully submitted; click on MyJobs to check its status' and 'Jobs run both on chain reds and egi' with logos for 'chain reds' and 'egi'.



# Reproducibility of ALEPH data with the CHAIN-REDS Science Gateway (1/3)

Co-ordination & Harmonisation of Advanced e-Infrastructures  
for Research and Education Data Sharing



## Science Gateway

Welcome Help Applications Job Status My Workspace My Cloud Project Home

"Hello World" ▶

Computational chemistry ▶

Computer Science and Mathematics ▶

Earth Sciences ▶

High Energy Physics ▶

Life Sciences ▶

Cloud Applications ▶

Data Applications ▶

Other ▶

Full List ▶

By definition, a Science Gateway is a set of tools, applications, and services that are customized to meet the needs of a specific scientific community.

The present Science Gateway is based on the CHAIN and CHAIN-REDS Gateway paradigm and is available worldwide, based on different e-Infrastructures (Grid, HPC, Cloud or simply local resources) at each user application level.

Thanks to the collaboration of the CHAIN-REDS partners (EMED-NEA-CRESCO, EUMEDGRID-Support, FutureGrid, GARUDA, GISELA, IGI, JSAGA, SAGrid and WeNMR projects), a set of Demo Applications has been deployed on various Grid (based on EMI - gLite and UNICORE - Genesis II, Globus, GOS, OurGrid middleware), Cloud (based on OCCI compliant - Okeanos, OpenNebula and OpenStack - stacks) and local (based on the Platform Computing resource manager) resources and you can execute them through this portal in a simple and easy way. Some of the Cloud sites belong to the EGI Federated Cloud.

Contributors



Follow us on Social Networks


This Includes the possibility to access the Science Gateway from within the Social Network page.

# Reproducibility of ALEPH data with the CHAIN-REDS Science Gateway (2/3)

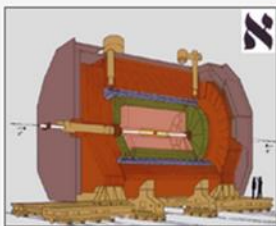
Co-ordination & Harmonisation of Advanced e-Infrastructures  
for Research and Education Data Sharing



## Science Gateway

Welcome Help Applications Job Status Project Home  Sign In

### ALEPH Analysis



ALEPH was a particle physics experiment installed on the Large Electron-Positron collider (LEP) at the CERN laboratory in Geneva/Switzerland. It was designed to explore the physics predicted by the Standard Model and to search for physics beyond it. ALEPH first measured events in LEP in July 1989. LEP operated at around 91 GeV – the predicted optimum energy for the formation of the Z particle. From 1995 to 2000 the accelerator operated at energies up to 200 GeV, above the threshold for producing pairs of W particles. The data taken, consisted of millions of events recorded by the ALEPH detector, allowed precision tests of the electro-weak Standard Model (SM) to be undertaken. The group here concentrated our analysis efforts mainly in Heavy Flavour (beauty and charm) physics, in searches for the the Higgs boson, the particles postulated to generate particle mass, and for physics beyond the SM, e.g. Supersymmetry, and in W physics.

This application perform the search for the production and non-standard decay of a scalar Higgs boson into four tau leptons through the intermediation of the neutral pseudo-scalars Higgs particle.

The analysis was conducted by the ALEPH collaboration with the data collected at centre-of-mass energies from 187 to 209 GeV.




Sign-in to RUN

Results are published in JHEP 1005 (2010) 049 DOI: 10.1007/JHEP05(2010)049

# Reproducibility of ALEPH data with the CHAIN-REDS Science Gateway (3/3)

Co-ordination & Harmonisation of Advanced e-Infrastructures  
for Research and Education Data Sharing



## Science Gateway

Welcome Help Applications Job Status My Workspace My Cloud Project Home

My Workspace

- Jobs
- JobsMap
- Data
- Help

aleph

Search for:  DOI  Keyword  View All

Show  entries

Title	Type	Action
<input type="checkbox"/> LEP Data from The ALEPH Collaboration year 2000 datafile ZD4000.1.AL	6LEP 6LEP MITQCD	<input type="button" value="Analyze"/>
<input type="checkbox"/> LEP Data from The ALEPH Collaboration year 2000 datafile ZD4000.11.AL	6LEP	<input type="button" value="Analyze"/>
<input type="checkbox"/> LEP Data from The ALEPH Collaboration year 2000 datafile ZD4000.12.AL	6LEP	<input type="button" value="Analyze"/>
<input type="checkbox"/> LEP Data from The ALEPH Collaboration year 2000 datafile ZD4000.13.AL	6LEP	<input type="button" value="Analyze"/>

Remember: repeatability and  
reproducibility are not all

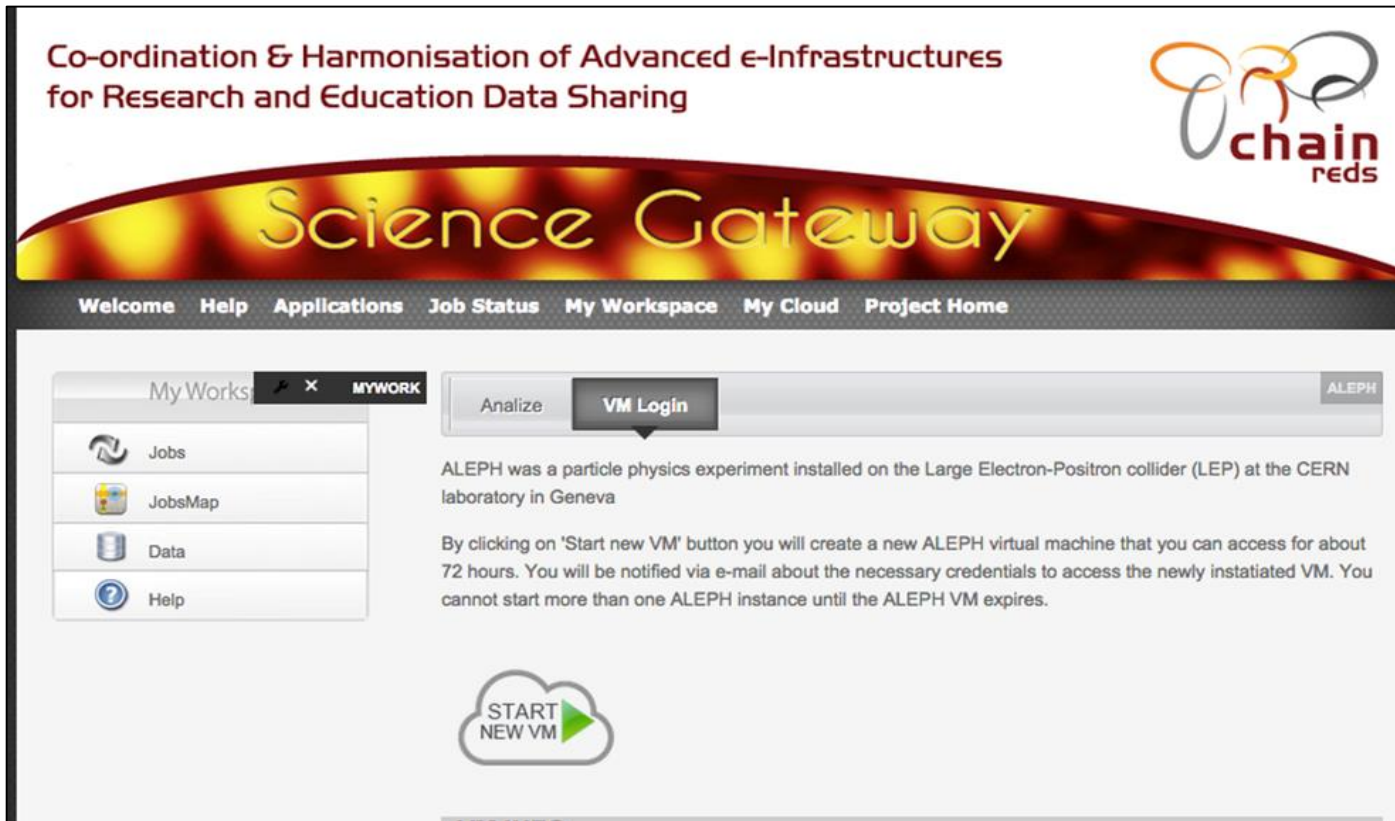
---

**Reusability and «extensibility» matter!**




# Reusability of ALEPH data with the CHAIN-REDS Science Gateway (1/2)

1. From within the CHAIN-REDS Science Gateway an entitled researcher can start a VM already configured to re-use/extend a given scientific analysis
2. The VM was previously distributed both on the CHAIN-REDS Cloud Testbed and on the EGI Federated Cloud using the EGI AppDB



Co-ordination & Harmonisation of Advanced e-Infrastructures  
for Research and Education Data Sharing



## Science Gateway

Welcome Help Applications Job Status My Workspace My Cloud Project Home

My Works ✕ MYWORK

Analyze **VM Login** ALEPH

ALEPH was a particle physics experiment installed on the Large Electron-Positron collider (LEP) at the CERN laboratory in Geneva

By clicking on 'Start new VM' button you will create a new ALEPH virtual machine that you can access for about 72 hours. You will be notified via e-mail about the necessary credentials to access the newly instatiated VM. You cannot start more than one ALEPH instance until the ALEPH VM expires.

START NEW VM

# Reusability of ALEPH data with the CHAIN-REDS Science Gateway (2/2)

1. The VM available for a customizable amount of time during which the user has full access to the dataset(s) and analysis algorithm(s) of the experiment
2. The user receives via email the credentials to access the VM using the protocols (e.g., SSH, VNC, etc.). Clicking on the SSH or VNC links will directly access the cloud-located VM from within the CHAIN-REDS Science Gateway

**The same work is gonna start with LHC's ALICE Experiment  
Orders of magnitude more data and... more complexity**

**New stable analyses (and their results), generated running the VM, may be registered in the OAR (with a DOI) to further extend the analysis catalogue shared across the Virtual Research Community**



# Summary and conclusions

- ▶ Open Science vision can be implemented only if the “openness” paradigm becomes pervasive in research
- ▶ Science outputs’ reproducibility, but also re-usability and extensibility, are key to walk through the “knowledge path” in both directions
- ▶ The INFN Open Access Repository is a pilot data preservation repository of science products meant to serve both researchers and citizen scientists; what makes OAR different from other repositories is its capability to connect to Science Gateways and exploit cloud resources worldwide to easily reproduce/extend scientific analyses
- ▶ The feasibility of the OAR Knowledge Workflow has successfully been tested with ALEPH datasets and will now be benchmarked with ALICE ones
- ▶ For the new tests we plan to explore CERNVM as well as Containers + Docker

Thank you !

- ▶ S. Bianco (INFN LNF)
- ▶ T. Boccali (INFN Pisa)
- ▶ R. Bruno (INFN Catania)
- ▶ C. Carrubba (INFN Catania)
- ▶ G. Inserra (INFN Catania)
- ▶ M. Maggi (INFN Bari)
- ▶ D. Menasce (INFN Milano Bicocca)
- ▶ R. Ricceri (INFN Catania)