

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

A CHAIN-REDS solution for accessing computational services

Rafael Mayo García, CIEMAT

Cancun / 26-28 May 2014

CHAIN-REDS: A legacy from CHAIN



- ▶ CHAIN-REDS is an EC (306819) funded project
 - ▶ ~ 2.1 M€
 - ▶ 1 December 2012 – 30 months
- ▶ Structured in
 - ▶ WP 1 Project Management
 - ▶ WP 2 Dissemination, Training and Outreach
 - ▶ WP 3 Interoperation and coordination of e-Infrastructures
 - ▶ WP 4 Data Infrastructures
 - ▶ WP 5 Support to small groups and emerging communities



- ▶ CHAIN-REDS is an EC (306819) funded project
 - ▶ ~ 2.1 M€
 - ▶ 1 December 2012 – 30 months
- ▶ Structured in
 - ▶ WP 1 Project Management
 - ▶ WP 2 Dissemination, Training and Outreach
 - ▶ WP 3 Interoperation and coordination of e-Infrastructures
 - ▶ WP 4 Data Infrastructures
 - ▶ WP 5 Support to small groups and emerging communities



WP4 'Data infrastructures'

► Partners

- INFN
- CIEMAT
- GRNET
- CESNET
- UBUNTUNET
- CLARA
- IHEP
- ASREN
- SIGMA ORIONIS
- C-DAC



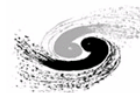
Institute of High Energy Physics
Chinese Academy of Sciences



WP4 'Data infrastructures'

► Partners

- INFN
- CIEMAT
- GRNET
- CESNET
- UBUNTUNET
- CLARA
- IHEP
- ASREN
- SIGMA ORIONIS } Europe
- C-DAC



Institute of High Energy Physics
Chinese Academy of Sciences



WP4 'Data infrastructures'

- ▶ INFN
- ▶ CIEMAT
- ▶ GRNET
- ▶ CESNET
- ▶ UBUNTUNET → Africa
- ▶ CLARA
- ▶ IHEP
- ▶ ASREN
- ▶ SIGMA ORIONIS } Europe
- ▶ C-DAC

WP4 'Data infrastructures'

- ▶ INFN
- ▶ CIEMAT
- ▶ GRNET
- ▶ CESNET
- ▶ UBUNTUNET → Africa
- ▶ CLARA → Latin America
- ▶ IHEP
- ▶ ASREN
- ▶ SIGMA ORIONIS
- ▶ C-DAC

WP4 'Data infrastructures'

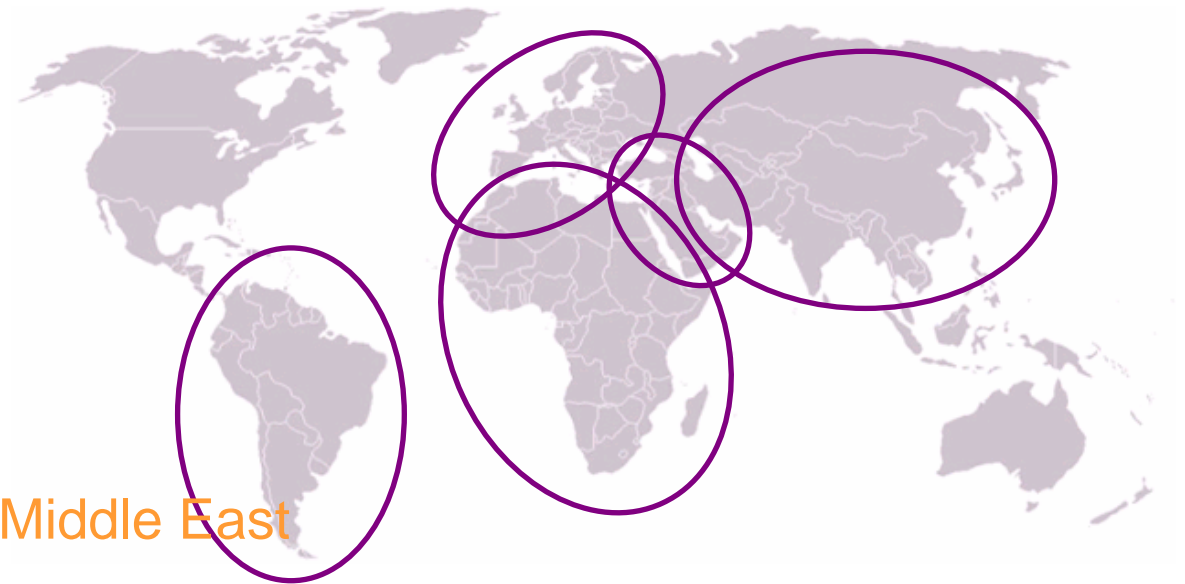
- ▶ INFN
 - ▶ CIEMAT
 - ▶ GRNET
 - ▶ CESNET
 - ▶ UBUNTUNET
 - ▶ CLARA → Latin America
 - ▶ IHEP → Asia
 - ▶ ASREN
 - ▶ SIGMA ORIONIS
 - ▶ C-DAC → Asia
-

WP4 'Data infrastructures'

- ▶ INFN
 - ▶ CIEMAT
 - ▶ GRNET
 - ▶ CESNET
 - ▶ UBUNTUNET
 - ▶ CLARA
 - ▶ **IHEP** → Asia
 - ▶ **ASREN** → Middle East
 - ▶ SIGMA ORIONIS
 - ▶ **C-DAC** → Asia
-

WP4 'Data infrastructures'

- ▶ INFN
- ▶ CIEMAT
- ▶ GRNET
- ▶ CESNET
- ▶ UBUNTUNET
- ▶ CLARA
- ▶ IHEP
- ▶ ASREN → Middle East
- ▶ SIGMA ORIONIS
- ▶ C-DAC



WP4 'Data infrastructures'

- ▶ Current work on Data Infrastructure can be found in the public project Deliverables
 - ▶ D4.1 Trans-continental Data Infrastructures and Data repositories
 - ▶ D4.2 Analysis of Data Infrastructures and Data repositories
 - ▶ D4.3 Use cases of the identified Data Infrastructures and Data repositories (coming soon)
 - ▶ It describes the CHAIN-REDS solution for data workflows
- ▶ Available at <http://www.chain-project.eu/deliverables>

WP4 'Data infrastructures'

- ▶ CHAIN-REDS has established official collaborations (MoUs) with other VRC-related communities

- ▶ AgINFRA
- ▶ DCH-RP
- ▶ EarthServer
- ▶ EIFL
- ▶ ENGAGE
- ▶ EUDAT



WP4 'Data infrastructures'

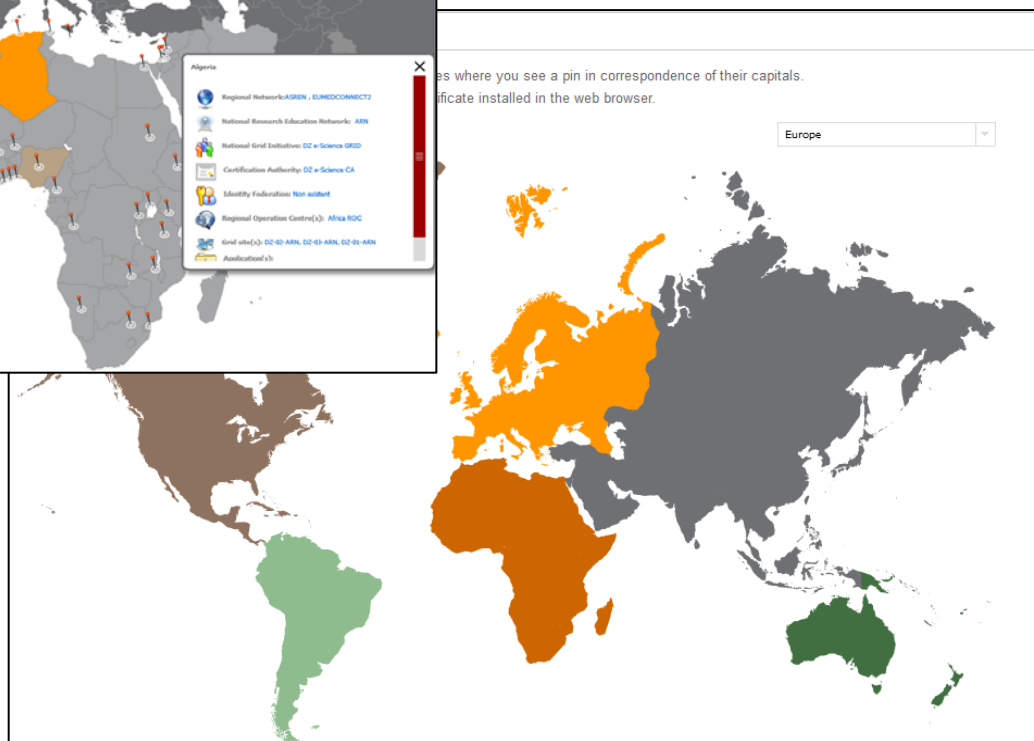
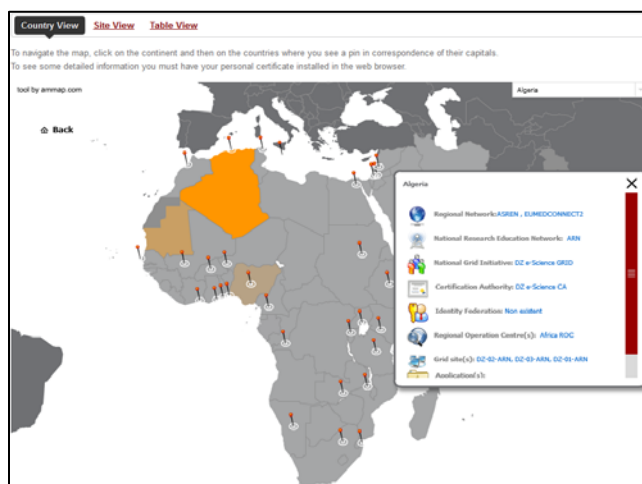
- ▶ Such work has been focused on scientific codes, but the developments that come afterwards can be easily applied to services of interest to the NRENs and Universities communities
 - ▶ Access to services, data, tools (IdF)
 - ▶ Semantic enriched search that allows new knowledge (SSE)
 - ▶ Friendly front-end for the execution of tasks (SG)
 - ▶ Retrieval of results and further storage of them (PID)
- ▶ The scientific side of these services are of interest to SCALAC

WP4 'Data infrastructures'

- ▶ Let us imagine an student looking for his/her record or a network engineer operating and NREN
 - ▶ Access to services, data, tools (IdF)
 - ▶ Login into the academia record or the NREN database
 - ▶ Semantic enriched search that allows new knowledge (KB/SSE)
 - ▶ Find a qualification, but know if others are available; now some statistics of an NREN, but know if others are available about the ROC
 - ▶ Friendly front-end for the execution of tasks (SG)
 - ▶ A tool for analytics process of qualifications (by a Prof.) or data network usage
 - ▶ Retrieval of results and further storage of them (PID)
 - ▶ A new digital object (the previous analysis) is uniquely identified

Knowledge Base: Infrastructure

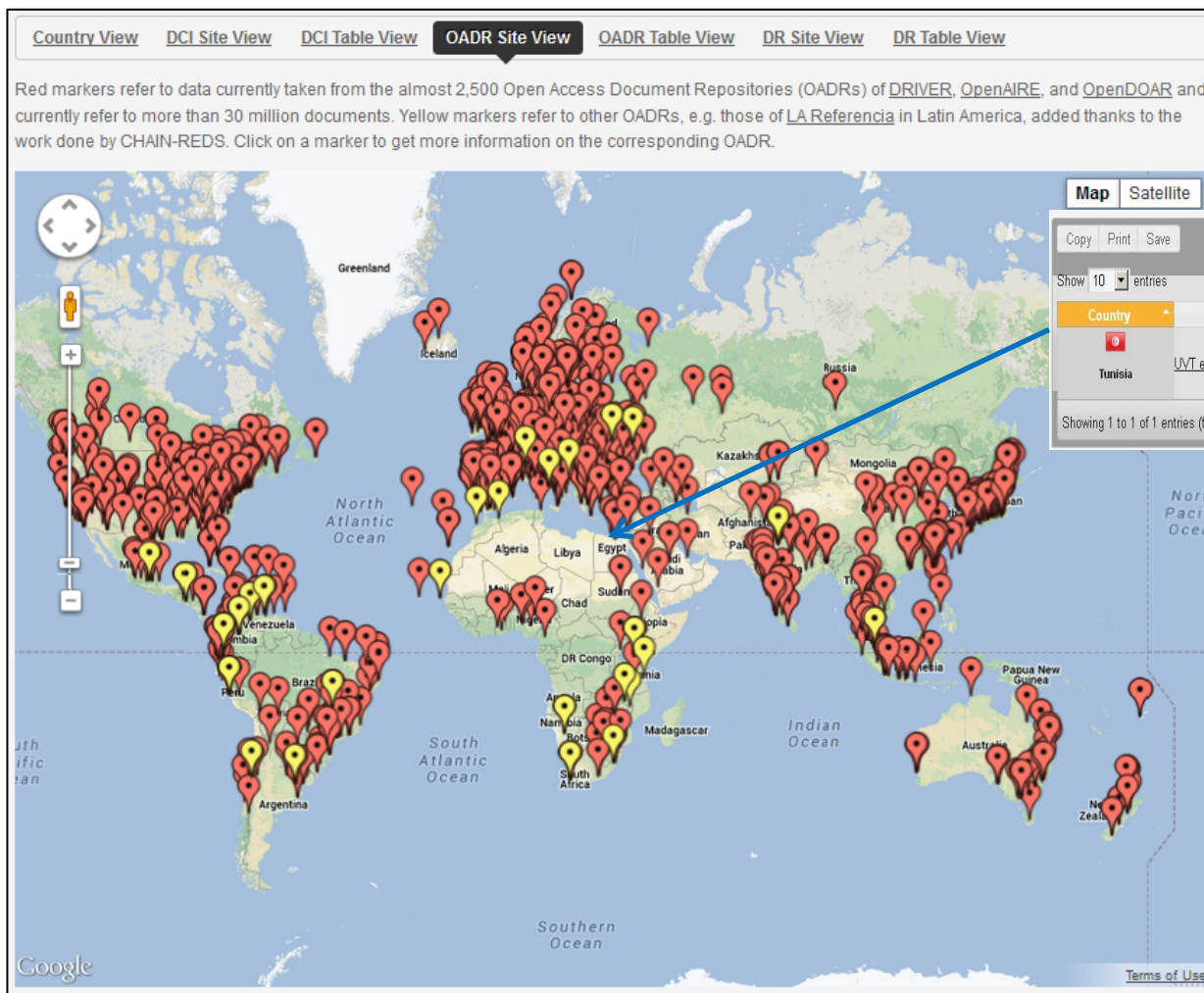
- ▶ Extend the CHAIN-REDS Knowledge Base (BS) with Data capabilities <http://www.chain-project.eu/knowledge-base>



- ▶ RREN(s)
- ▶ NREN
- ▶ NGI
- ▶ CA(s)
- ▶ Ident. Fed(s)
- ▶ ROC(s)
- ▶ Grid site(s)
- ▶ Application (s)

Knowledge Base: Document & Data repositories

- ~3,200 repos
- >33 M docs



Registry of Open Access
Repositories (ROAR)



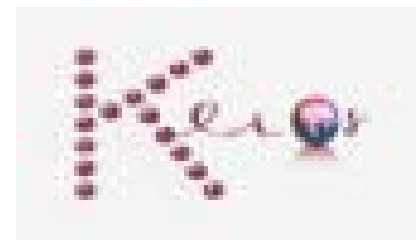
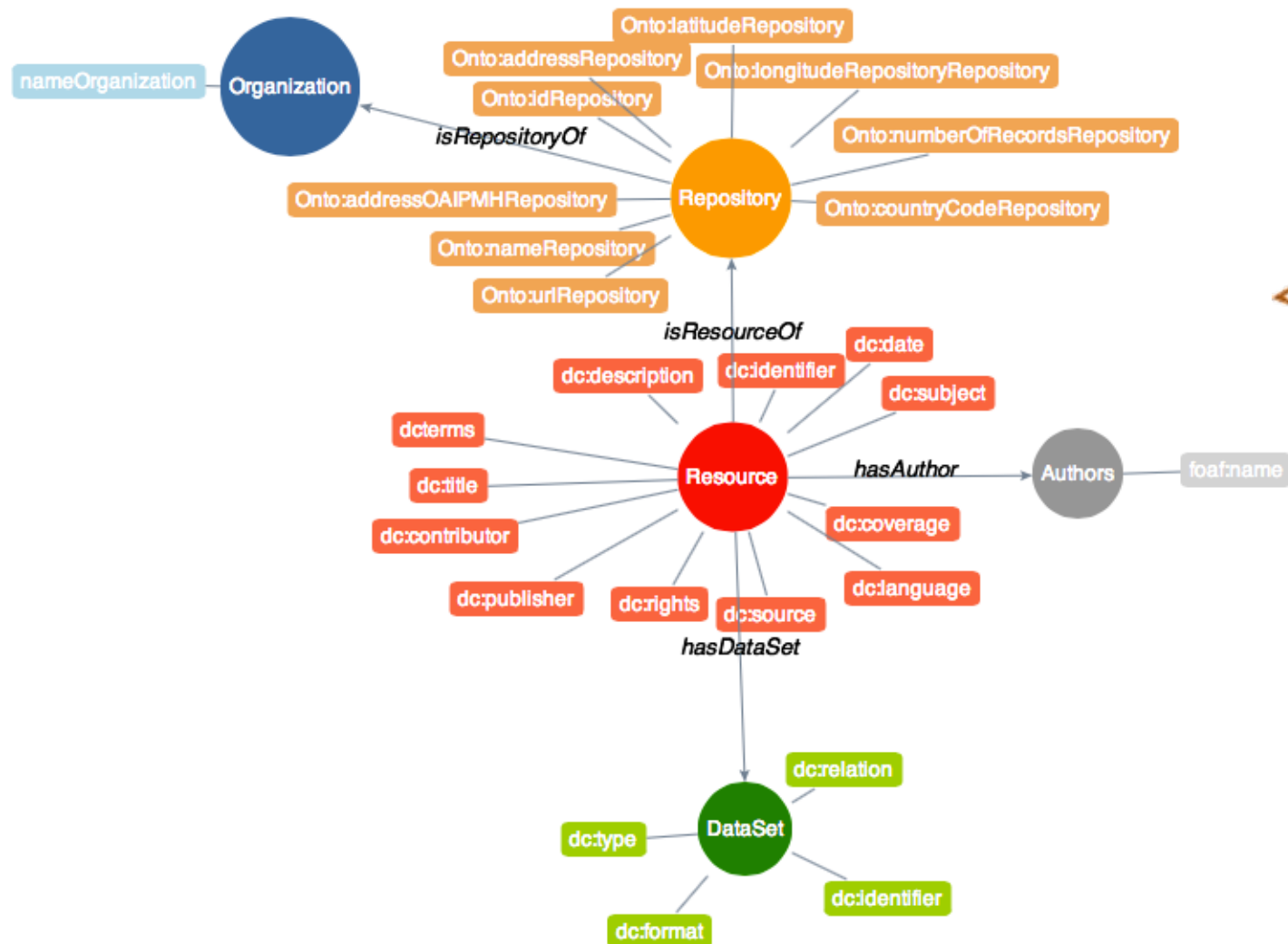
- ▶ About Open Access Data Repositories, standards have been promoted
 - ▶ OAI-PMH for metadata retrieval
 - ▶ Dublin Core as metadata schema
 - ▶ SPARQL for semantic web search
 - ▶ VOTable (XML) as potential standard for the interchange of data represented as a set of tables
 - ▶ Persistent Identifiers (PID)



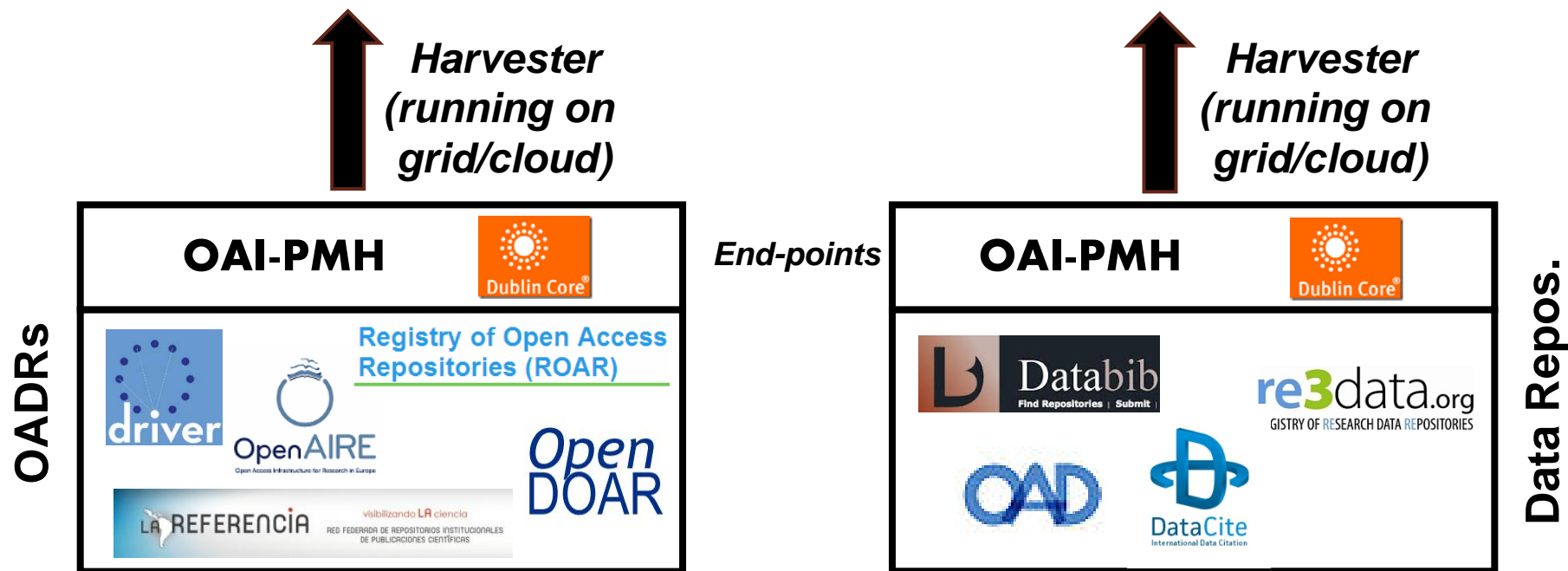
- ▶ The adopted standards have been implemented in the CHAIN-REDS KB
 - ▶ La Referencia repository is included in the KB



- ▶ Developments on (Open Access) Document and Data Repositories have been carried out
 - ▶ A semantic web enrichment
 - ▶ A semantic search engine



Semantic Search Engine architecture




- ▶ The semantic search engine on CHAIN-REDS linked data is available
 - ▶ Allows searching among the semantically-enriched metadata coming from the OADRs and DRs included in the KB

Semantic Search on Linked Data

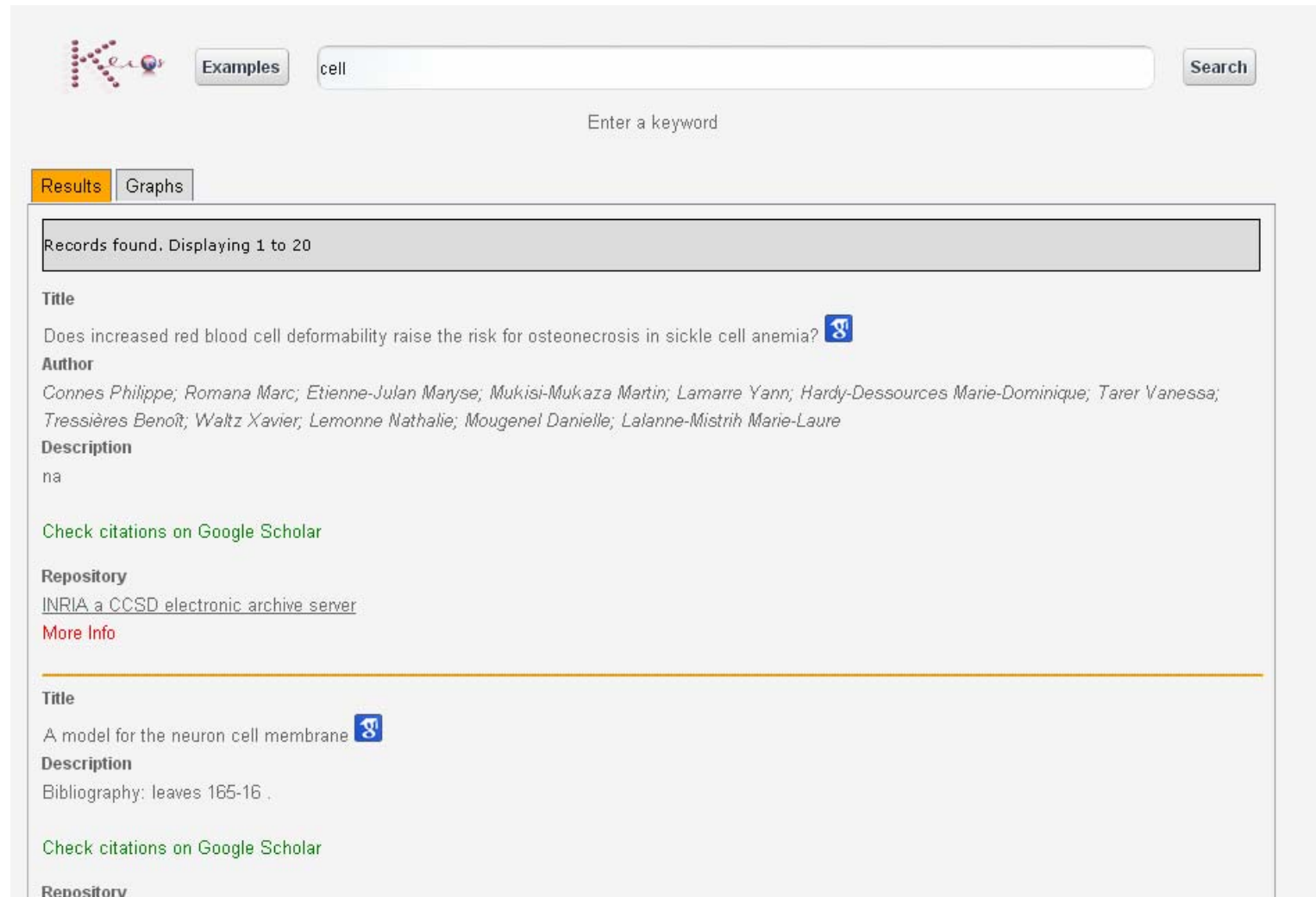
Here you can search across the thousands of semantically enriched [Open Access Document Repositories](#) and [Data Repositories](#) included in the [CHAIN Knowledge Base](#). Some exemplar keywords are: cardiology, cornea, geology, linked data, natural gas.

Select a language: English ▼



Submit Query

Enter a keyword or select a language and then choose a subject



The screenshot shows the Keri search interface. At the top left is the Keri logo. Next to it is an 'Examples' button. To the right is a search input field containing the text 'cell'. Further right is a 'Search' button. Below the search bar is a placeholder text 'Enter a keyword'. Underneath the search bar are two tabs: 'Results' (which is active and highlighted in orange) and 'Graphs'. Below the tabs is a grey box containing the text 'Records found. Displaying 1 to 20'. The first search result is displayed below this box. It has a 'Title' section with the text 'Does increased red blood cell deformability raise the risk for osteonecrosis in sickle cell anemia?' followed by a small blue square icon with a white 'g'. Below the title is an 'Author' section with a list of names: 'Connes Philippe; Romana Marc; Etienne-Julan Maryse; Mukisi-Mukaza Martin; Lamarre Yann; Hardy-Dessources Marie-Dominique; Tarer Vanessa; Tressières Benoît; Waltz Xavier; Lemonne Nathalie; Mougeneil Danielle; Lalanne-Mistrih Marie-Laure'. Below the author list is a 'Description' section with the text 'na'. Below the description is a green link that says 'Check citations on Google Scholar'. Below this is a 'Repository' section with the text 'INRIA a CCSD electronic archive server'. Below the repository is a red link that says 'More Info'. A horizontal orange line separates this result from the next one. The second search result is displayed below the line. It has a 'Title' section with the text 'A model for the neuron cell membrane' followed by a small blue square icon with a white 'g'. Below the title is a 'Description' section with the text 'Bibliography: leaves 165-16'. Below the description is a green link that says 'Check citations on Google Scholar'. Below this is a 'Repository' section which is currently empty.

Keri

Examples

cell


Search

Enter a keyword

Results Graphs

Records found. Displaying 1 to 20

Title

Does increased red blood cell deformability raise the risk for osteonecrosis in sickle cell anemia? 

Author

Connes Philippe; Romana Marc; Etienne-Julan Maryse; Mukisi-Mukaza Martin; Lamarre Yann; Hardy-Dessources Marie-Dominique; Tarer Vanessa; Tressières Benoît; Waltz Xavier; Lemonne Nathalie; Mougeneil Danielle; Lalanne-Mistrih Marie-Laure

Description

na


[Check citations on Google Scholar](#)

Repository

[INRIA a CCSD electronic archive server](#)

[More Info](#)

Title

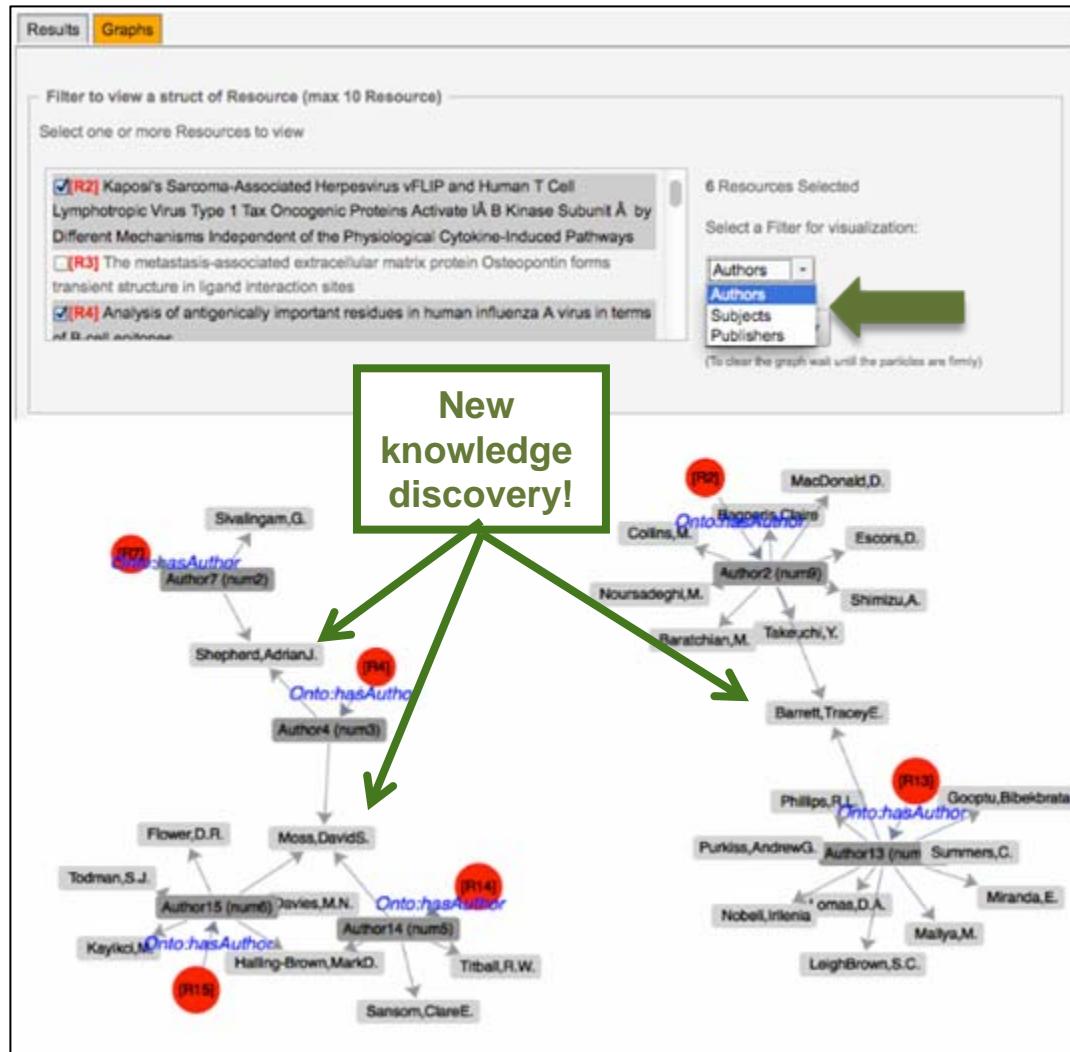
A model for the neuron cell membrane 

Description

Bibliography: leaves 165-16

[Check citations on Google Scholar](#)

Repository



Semantic Search Engine

- ▶ In order to improve the efficiency of the SSE, the results are ranked according to the Jan 2014 edition of the Ranking Web of Repositories
 - ▶ As a result, the CHAIN-REDS KB and SSE are now featured on Semanticweb.com
- ▶ The search can be also refined by keywords
 - ▶ author, subject, type, format, and publisher
- ▶ It is possible to query Google Scholar for each resource found in a given search
 - ▶ The number and the list of citations as well as the year of publication are retrieved, if available


Semantic Search Engine

- ▶ Single and Parallel semantic search are available
 - ▶ Single: the usual semantic search service described before
 - ▶ Parallel: the new parallel semantic search service that allow users to search in parallel across the millions of resources contained in the CHAIN-REDS Knowledge Base and in the ENGAGE Platform

- ▶ Parallel semantic search engines have been made available also in others Science Gateways
 - ▶ agINFRA (CHAIN-REDS Knowledge Base & OpenAgris repository)
 - ▶ DCH-RP (CHAIN-REDS Knowledge Base & Europeana, Cultura Italia and Isidore repositories)

Semantic Search Engine


You will search in CHAIN-REDS-KB , Engage



Enter a keyword

CHAIN-REDS-KB

Records found. Displaying 1 to 20


Title
Does increased red blood cell deformability raise the risk for osteonecrosis in sickle cell anemia? 

Author
Connes Philippe; Romana Marc; Etienne-Julan Maryse; Mukisi-Mukaza Martin; Lamarre Yann; Hardy-Dessources Marie-Dominique; Tarer Vanessa; Tressières Benoît; Waltz Xavier; Lemonne Nathalie; Mougene! Danielle; Lalanne-Mistrih Marie-Laure;

Description
na

[Check citations on Google Scholar](#)

Repository
[INRIA a CCSD electronic archive server](#)
[More Info](#)

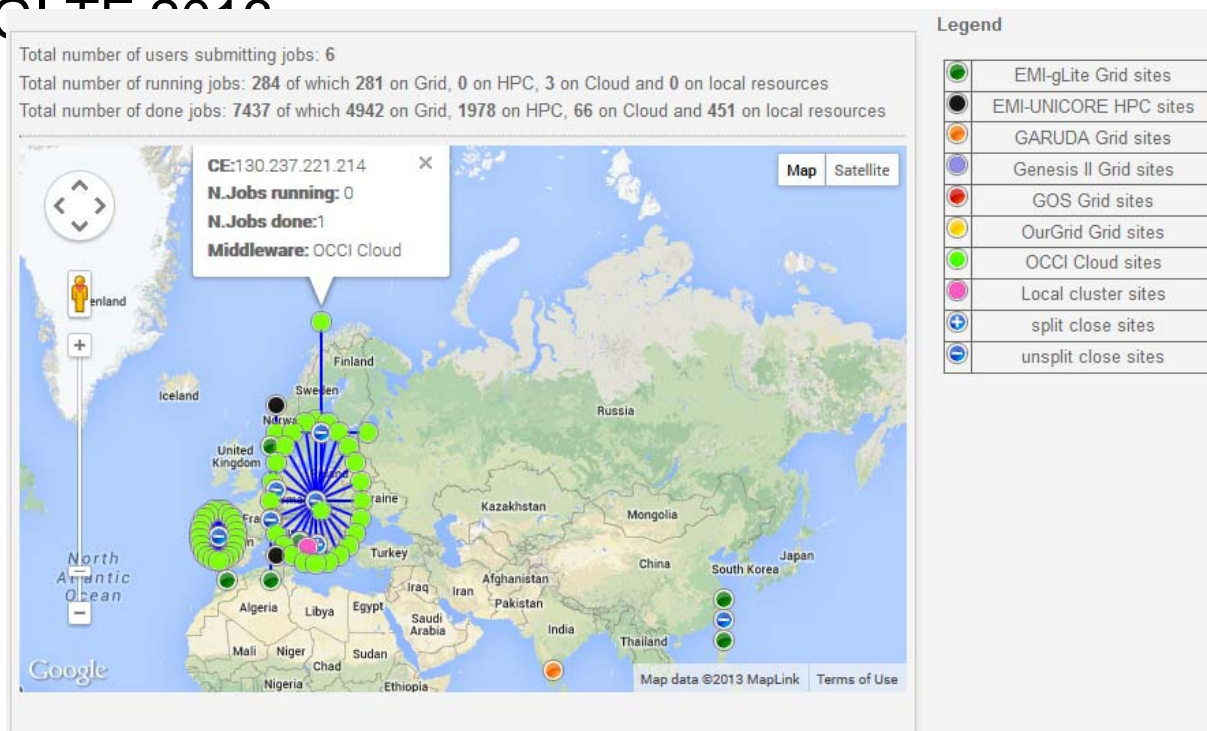
Title
A model for the neuron cell membrane 

Description
Bibliography: leaves 165-16 .

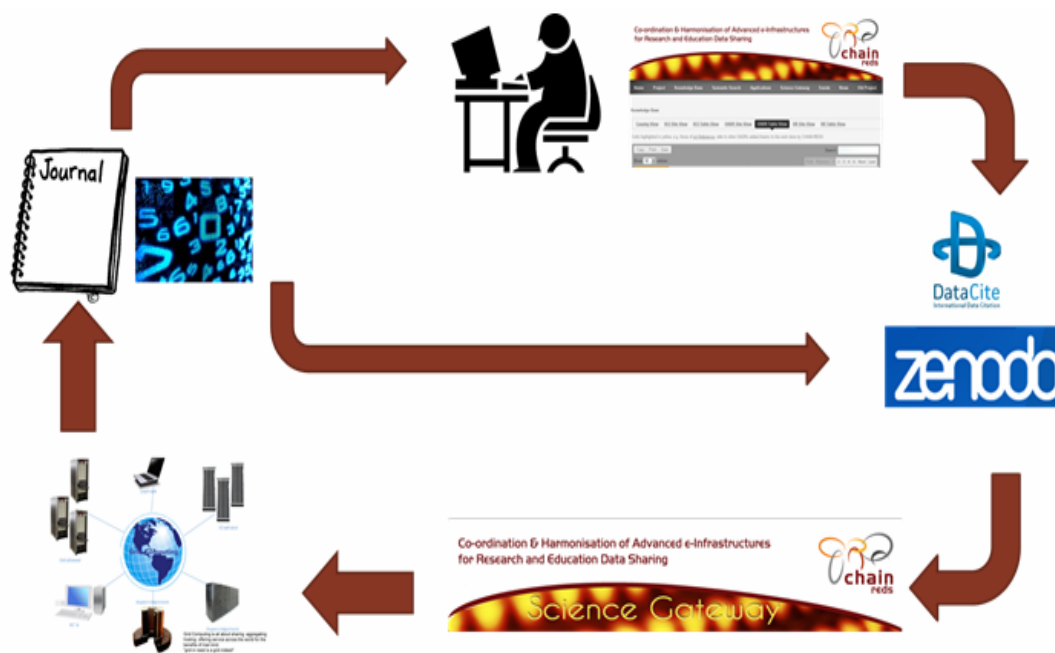
Semantic Search Engine

- ▶ A programmable use of the CHAIN-REDS Semantic Search Engine is also possible by means of a RESTful API
 - ▶ <http://www.chain-project.eu/semantic-search-api>
 - ▶ CHAIN-REDS webpage → Semantic Search → Web
- ▶ Example
 - ▶ http://www.chain-project.eu/virtuoso/api/resources?keyword=<KEYWORD>&limit=<NUMBER_OF_RESOURCES>

- ▶ Data Accessibility, Reproducibility and Trustworthiness (DART)
 - ▶ Based on the interoperability demo performed by CHAIN-REDS at ECTP 2012



- ▶ DART aims at seamlessly perform the cycle
 - ▶ Access to a document → Extraction of associated raw data → Execution of a code taking those data as input → Generation of new results → Upload of the new results and article



- ▶ Search a term/author/document can be directly performed using the CHAIN-REDS KB and/or SSE
- ▶ Datasets found can be used as an input in the CHAIN-REDS Science Gateway
 - ▶ Accessed by Identity Federation services
- ▶ Jobs are executed on Grid and Cloud platforms
- ▶ Output is retrieved and, if desired, can be identified by a PID service
 - ▶ GRNET service

- ▶ DART has been successfully tested using the Molon portlet
 - ▶ The MPI-Mainz UV/VIS Spectral Atlas of Gaseous Molecules of Atmospheric Interest
 - ▶ Cross sections → Molecular Absorption coefficients



**The MPI-Mainz UV/VIS Spectral Atlas
of Gaseous Molecules of Atmospheric Interest**

www.uv-vis-spectral-atlas-mainz.org

Hannelore Keller-Rudek¹, Geert K. Moortgat², Rolf Sander², Rüdiger Sörensen¹

¹Satellite Group
²Atmospheric Chemistry Division
Max-Planck Institute for Chemistry
Mainz, Germany

[Home](#)

[Cross Sections](#)

[Quantum Yields](#)

[Contact, Impressum, Acknowledgements, Citation](#)

Scientific background

The photolysis rates of gaseous trace species in the atmosphere are important parameters in the atmospheric sciences. This is especially true for modeling atmospheric chemistry, as most chemical reactions are directly or indirectly driven by the sun's radiation. Photolysis rates depend on the intensity of the actinic flux, and also on the properties of the absorbing molecules. Photodissociation rate coefficients are governed by the [absorption cross section](#) σ and the [quantum yield](#) q of the photolysis.

The absorption cross section σ is defined by the [Beer-Lambert law](#) describing the attenuation of light by a homogeneous absorbing system:

$$I = I_0 \exp(-(\sigma \cdot d \cdot n))$$

where I_0 and I are the incident and transmitted light intensities, d is the absorption path length (in cm), n is the concentration of the absorber (in molecule/cm³), and σ is the absorption cross section (in cm² molecule⁻¹).

The quantum yield q is the probability that a particular photochemical process will occur following the absorption of a photon by the molecule.

Both, σ and q , depend on wavelength, temperature, and pressure. Thus, knowing their values under atmospheric conditions is essential. In addition, the experimental determination of kinetic and photochemical parameters of many elementary reactions requires the precise knowledge of the absorption cross sections of many species.

Search Cross Sections

Species Search:

▶

Identifier Search:

▶

Reference Search:

▶

Full Text Search:

▶

- ▶ CHAIN-REDS is targeting scientific communities with a worldwide presence
- ▶ CHAIN-REDS has been working on developing tools with exploit current data capabilities
 - ▶ Knowledge Base
 - ▶ Semantic Web Enrichment
 - ▶ Semantic Search Engine
 - ▶ Science Gateway

- ▶ CHAIN-REDS has developed the Data Accessibility, Reproducibility and Trustworthiness (DART) challenge
- ▶ DART intends to easily reproduce the whole computing process devoted to either perform or reproduce a (previous) complete research
 - ▶ Look for available data and references
 - ▶ Used them as input data on a multiplatform application
 - ▶ Execute that application seamlessly
 - ▶ Retrieve the results and stored them with a Persistent Identifier

- ▶ DART is running with a proof-of-principle application from the chemical physics domain
 - ▶ Our acknowledgement to the MPI-Mainz UV/VIS Spectral Atlas repository owners
 - ▶ DART will be proposed to the regional researchers and collaborative initiatives of CHAIN-REDS in order to be adapted to their needs
 - ▶ It can be applied to scientific codes, but administrative services too
 - ▶ DART is part of the EUDAT Workflow Working Group
-

- ▶ DART concept can be easily adapted to other academic purposes
 - ▶ NREN
 - ▶ University Secretariat
 - ▶ Administrative tools

- ▶ A DART video can be found at the CHAIN-REDS webpage

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

Thank you !

www.chain-project.eu
proj-office@chain-project.eu – rafael.mayo@ciemat.es