

CDS in the Open Data world

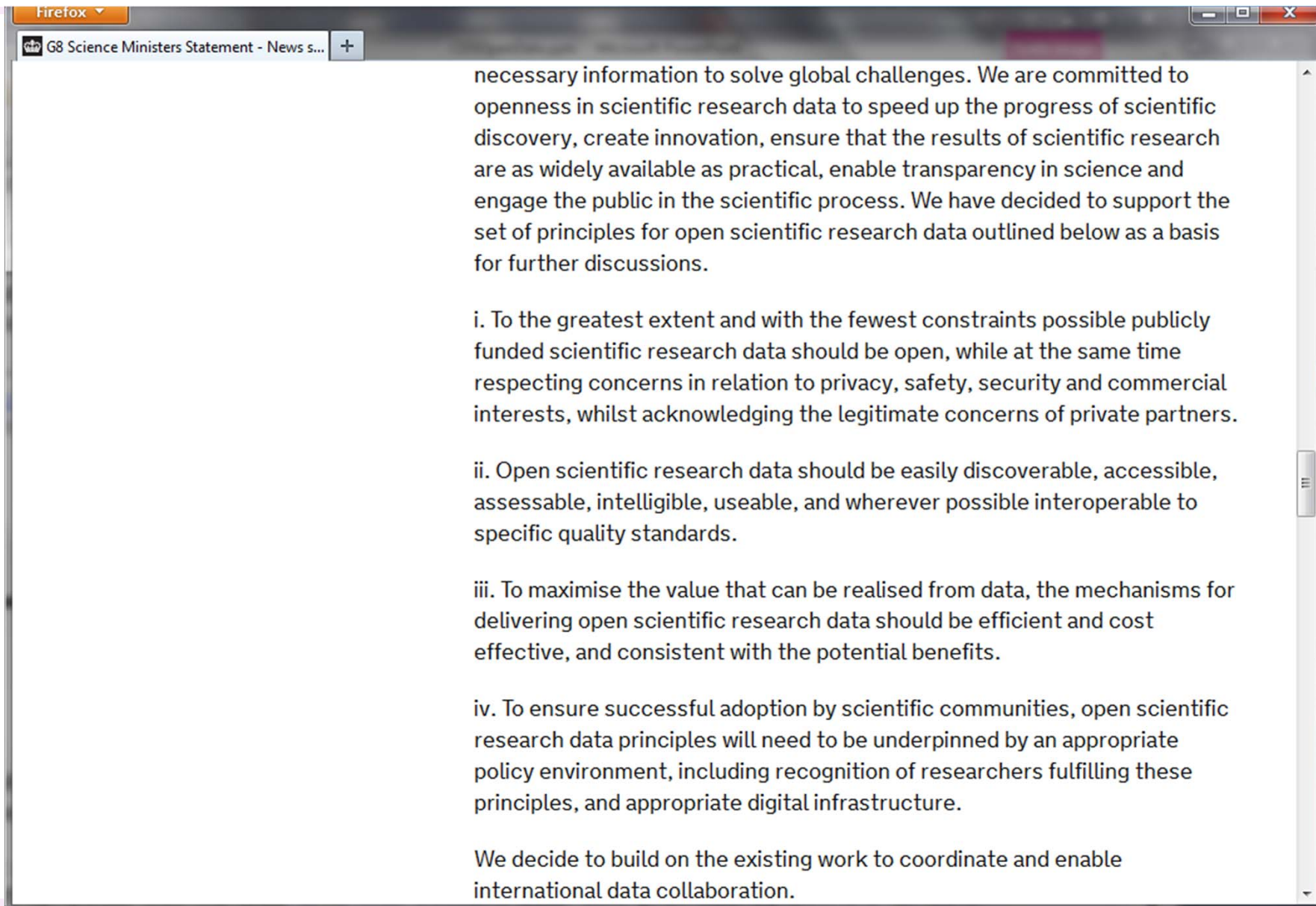
F. Genova and the CDS team

“Open data” is a hot topic

- Strong statement from the G8 Science Ministers (12 June 2013)

The screenshot shows a Mozilla Firefox browser window displaying the GOV.UK website. The address bar shows the URL: <https://www.gov.uk/government/news/g8-science-ministers-statement>. The page title is "G8 Science Ministers Statement - News stories - GOV.UK". The main content area features the heading "News story" followed by "G8 Science Ministers Statement". Below this, it states "Organisation: Foreign & Commonwealth Office" and "Page history: Published 13 June 2013". The main text begins with "On 12 June the Royal Society hosted the first ever G8 joint Science Ministers and national science academies meeting in London". A blue G8 logo is visible in the bottom left corner of the page content. The "Introduction" section starts with "We, the G8 Science Ministers met in London on Wednesday 12 June with Presidents of our respective national science academies, as part of the UK's G8 Science Ministers Meeting..."

The G8 Science Ministers' statement



necessary information to solve global challenges. We are committed to openness in scientific research data to speed up the progress of scientific discovery, create innovation, ensure that the results of scientific research are as widely available as practical, enable transparency in science and engage the public in the scientific process. We have decided to support the set of principles for open scientific research data outlined below as a basis for further discussions.

- i. To the greatest extent and with the fewest constraints possible publicly funded scientific research data should be open, while at the same time respecting concerns in relation to privacy, safety, security and commercial interests, whilst acknowledging the legitimate concerns of private partners.
- ii. Open scientific research data should be easily discoverable, accessible, assessable, intelligible, useable, and wherever possible interoperable to specific quality standards.
- iii. To maximise the value that can be realised from data, the mechanisms for delivering open scientific research data should be efficient and cost effective, and consistent with the potential benefits.
- iv. To ensure successful adoption by scientific communities, open scientific research data principles will need to be underpinned by an appropriate policy environment, including recognition of researchers fulfilling these principles, and appropriate digital infrastructure.

We decide to build on the existing work to coordinate and enable international data collaboration.

Astronomy at the forefront

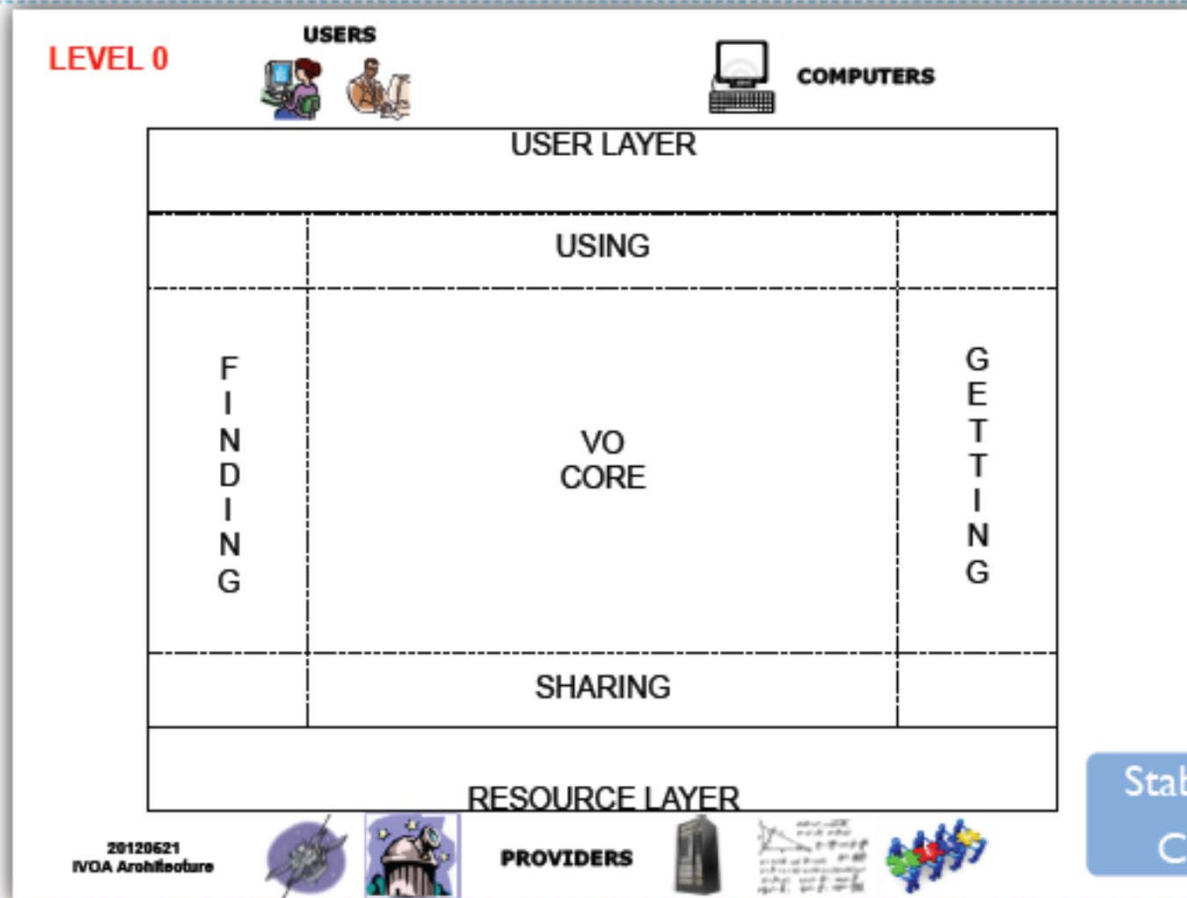


- Astronomy has been one of the pioneers of Open Scientific Data
- We demonstrate that on-line resources are widely used if the data is reusable
- Exchange standards, data formats and shared protocols are critical
 - FITS has been the basis of data exchange, discussion begun in 1976
 - The bibcode (CDS/NED) allowed exchange of bibliographic references in 1988
- The VO allows seamless access to data and VO tools provide integration of data from different origins
 - VO architecture well understood
- At the core of astronomy scientific needs
 - Multi-wavelengths, multi-technique astronomy
 - Comparison of theoretical models with observations
 - Time variability
 - Etc.
- An example at the “political” level and in practice
 - A new paradigm on how science is done: we have a practical experience
 - Knowledge of how to establish interoperability at the international level

IVOA Architecture Level 0



IVOA Architecture Level 0



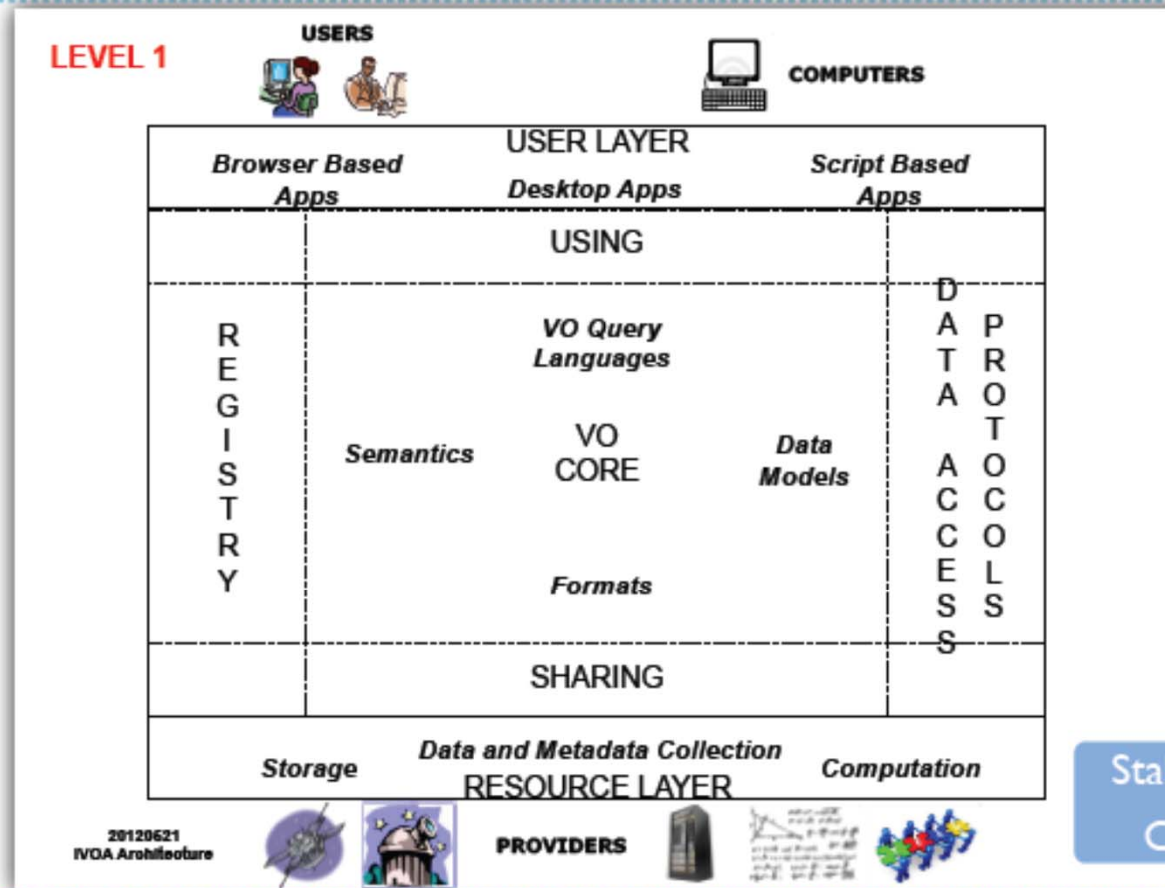
► State of TCG 2013-09-26

Séverin Gaudet – 6

IVOA Architecture Level 1

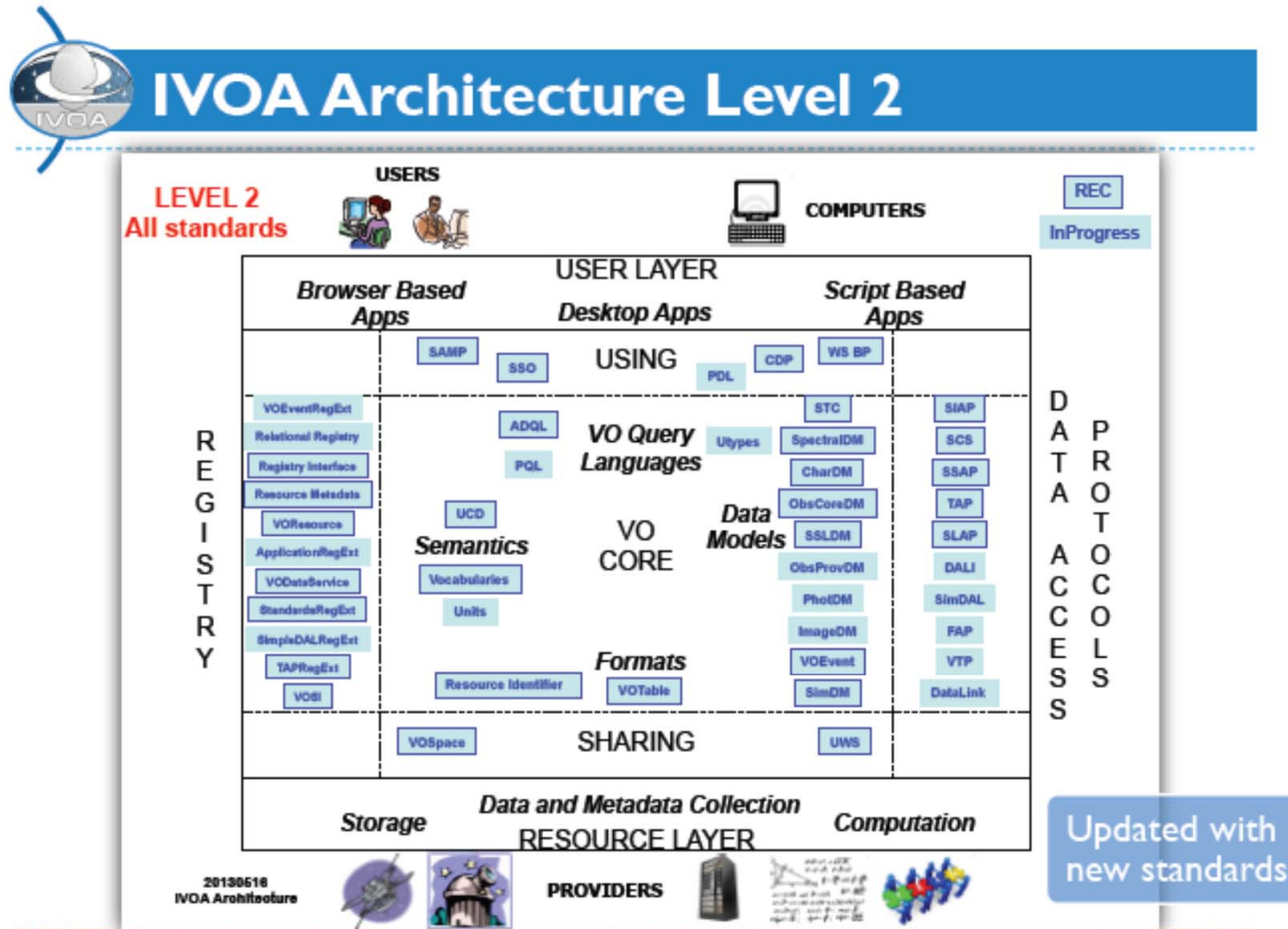


IVOA Architecture Level 1



Stable Since Creation

IVOA Architecture Level 2



▶ State of TCG 2013-09-26

Séverin Gaudet – 8

Sharing lessons learnt with EC Parliament members

The screenshot shows a web browser window displaying the website for the 4th European Innovation Summit | 2012. The page features the Knowledge4Innovation logo, navigation links (About, Events, Membership, Contact), and a main banner for the summit. The banner includes the text: "4th European Innovation Summit", "Building Bridges - Creating Synergies", "9-10 October 2012", and "European Parliament, Brussels". Below the banner, there is a paragraph of text describing the summit, followed by a "Photo Gallery" section with a photo of a man speaking at a microphone.

4th European Innovation Summit | 2012 | Knowledge4Innovation - Mozilla Firefox

Fichier Édition Affichage Historique Marque-pages Outils ?

4th European Innovation Summit | 2012 | ...

www.knowledge4innovation.eu/4th-european-innovation-summit-2012

Google

knowledge4innovation

About Events Membership Contact

4th European Innovation Summit | 2012



4th European Innovation Summit
Building Bridges - Creating Synergies
9-10 October 2012
European Parliament, Brussels

knowledge4innovation
K4I Forum of the European Parliament

4th European Innovation Summit | 2012

- > 4th EIS | Partners
- > 4th EIS | Programme and Speakers

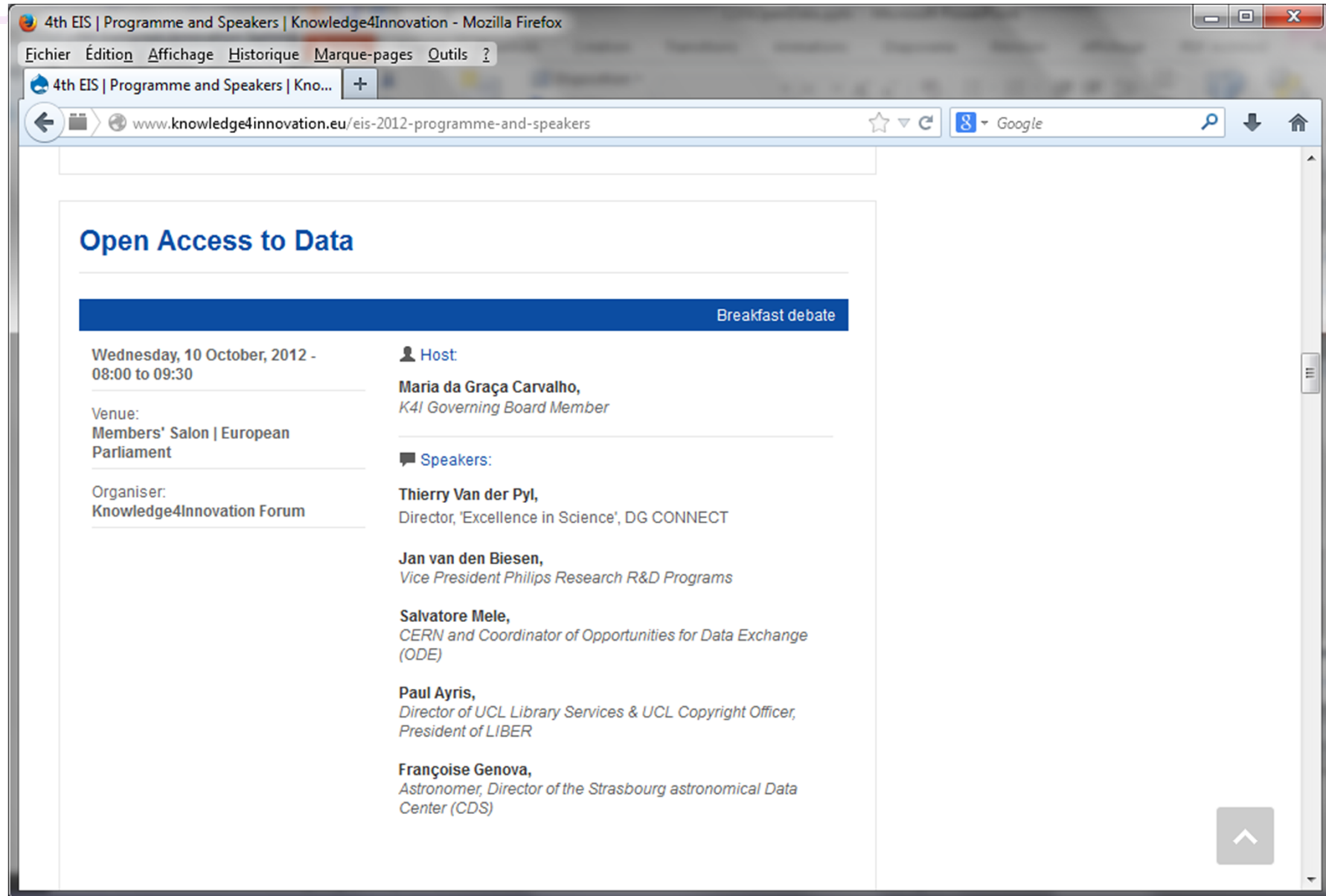
Photo Gallery:



knowledge4innovation
K4I Forum of the European Parliament

Participants largely agreed that, given their potential as a major driver of jobs and growth, the relevant EU policies and programmes in support of innovation ought to be revised. In order to better achieve this objective, it is crucial that all efforts, from the European to the local levels, are aligned. It is no surprise that, to a large extent, the discussions related to Horizon2020, the new Cohesion Policy Framework, the Programme for the Competitiveness of Enterprises and SMEs (COSME) and the Multiannual Financial Framework (MFF), which are currently under debate in the European Parliament.

Members of the European Parliament, who are dealing with innovation-related Commission proposals, contributed to the debate, including **Lambert van Nistelrooij**, Chair of the K4I Forum of the European Parliament and key rapporteur for the Cohesion Policy 2014-2020 as well as co-negotiator for Horizon 2020, **Ioannis Tsoukalas**, shadow rapporteur for the regulation establishing the European Institute for Innovation and Technology (EIT), **Maria Da Graça Carvalho**, rapporteur for the Specific Programme Implementing Horizon2020, **Christian Ehler**, rapporteur for Horizon2020 Rules for Participation, **Kent Johansson**, shadow rapporteur for Horizon2020, **Ivailo Kalfin**, co-rapporteur on the Multi-Annual



4th EIS | Programme and Speakers | Knowledge4Innovation - Mozilla Firefox

Fichier Édition Affichage Historique Marque-pages Outils ?

4th EIS | Programme and Speakers | Kno... +

www.knowledge4innovation.eu/eis-2012-programme-and-speakers

Google

Open Access to Data

Breakfast debate

Wednesday, 10 October, 2012 - 08:00 to 09:30

Venue:
Members' Salon | European Parliament

Organiser:
Knowledge4Innovation Forum

Host:
Maria da Graça Carvalho,
K4I Governing Board Member

Speakers:

Thierry Van der Pyl,
Director, 'Excellence in Science', DG CONNECT

Jan van den Biesen,
Vice President Philips Research R&D Programs

Salvatore Mele,
CERN and Coordinator of Opportunities for Data Exchange (ODE)

Paul Ayris,
Director of UCL Library Services & UCL Copyright Officer, President of LIBER

Françoise Genova,
Astronomer, Director of the Strasbourg astronomical Data Center (CDS)

The pioneering role of CDS



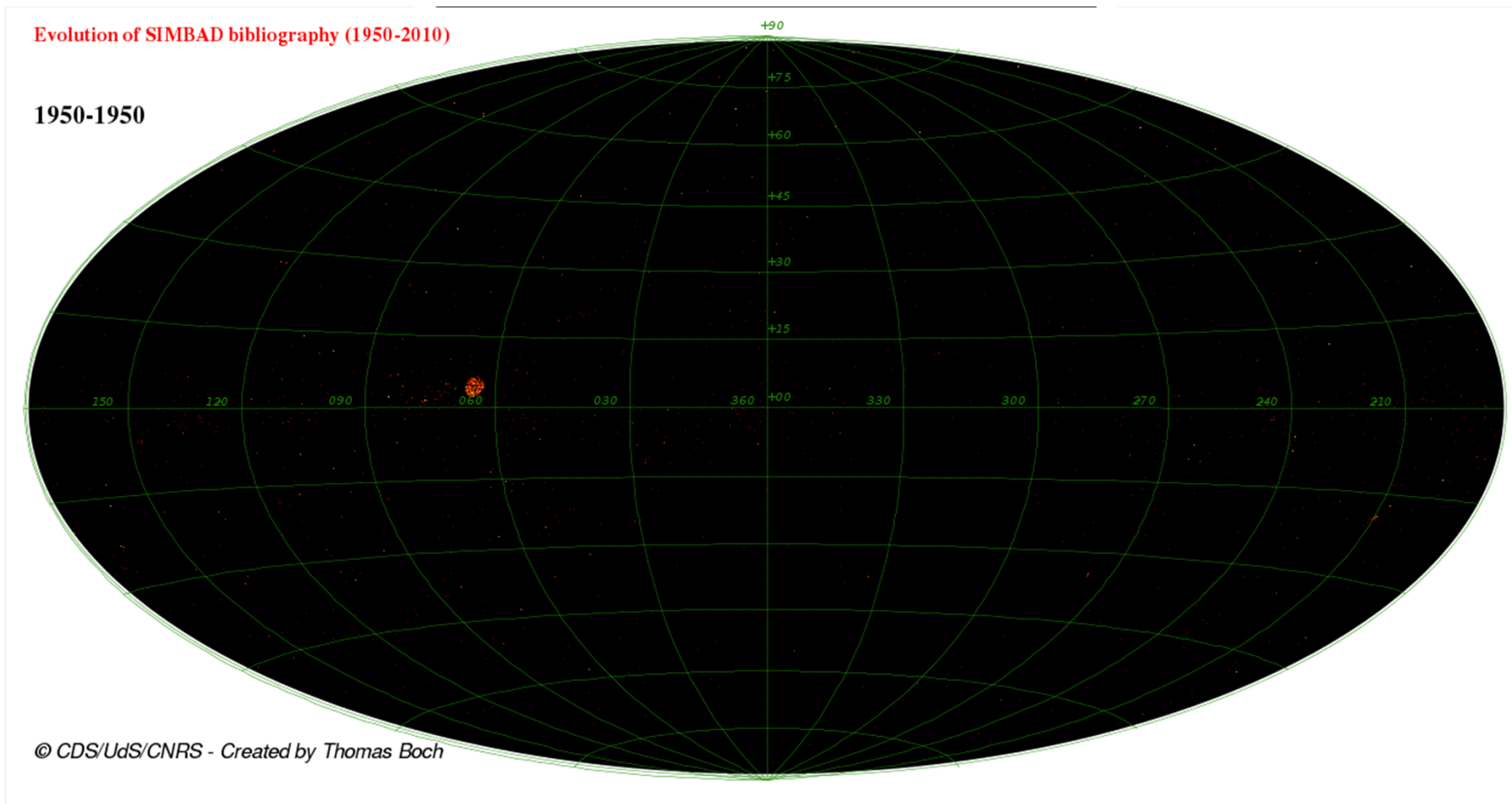
- Provision of **value-added** databases since 1972
- Driven by science needs and data usage, not data preservation
- Early networking of resources
 - SIMBAD name resolver 1993
 - Collaboration with A&A 1993
 - Close collaboration with the journals, the ADS, observatory archives (
 - Aladin as image data integrator, before the VO
- Major role in the VO development
 - Support to IVOA (Exec, SCSP, WG/IG leadership)
 - Very active participation in the development of VO standards and tools
 - Our services are major VO building blocks

Access to research results at CDS

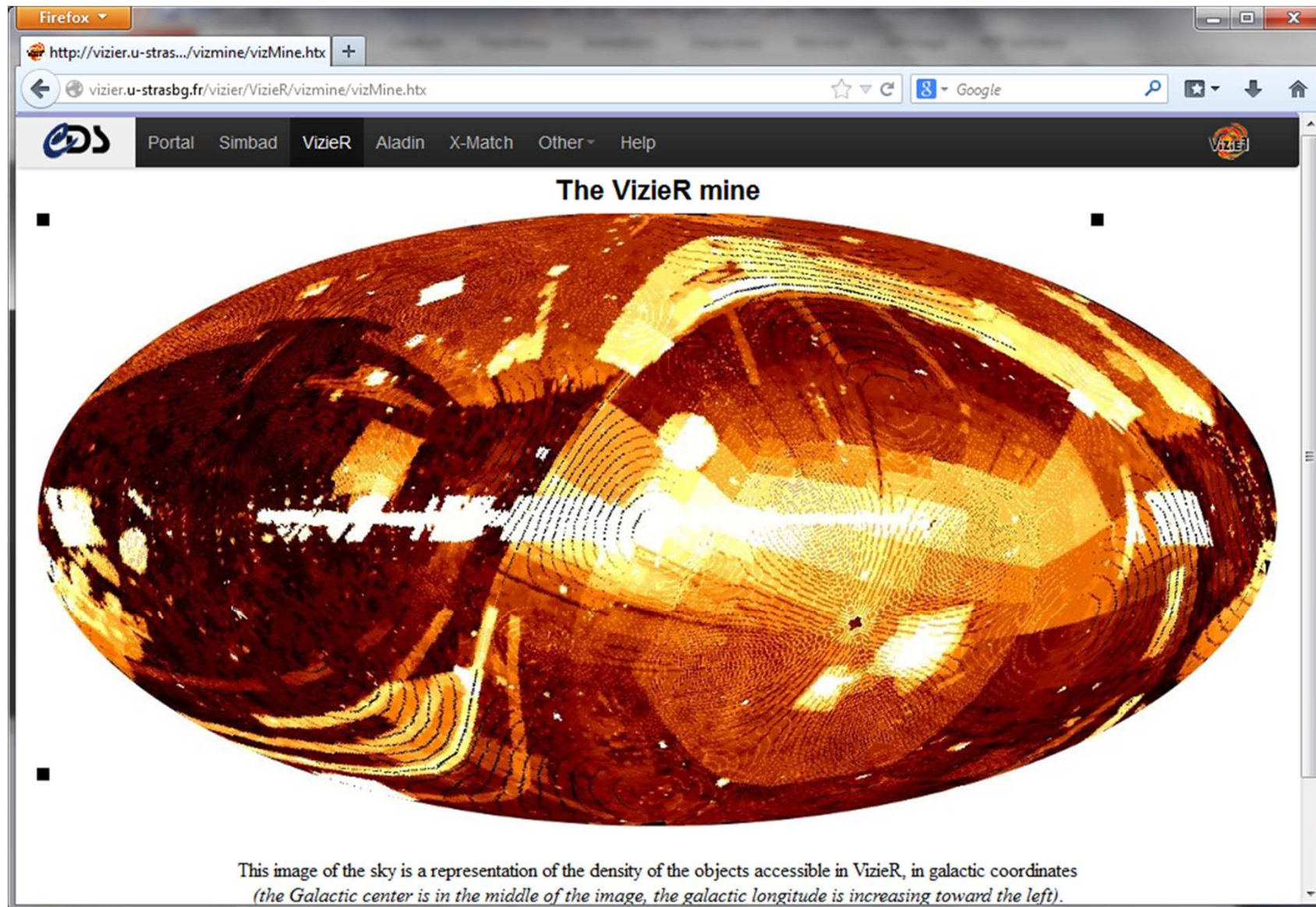


- One core of our activity
 - Extract validated and reusable information from refereed publications
 - Agreement with A&A in 1993
 - On-line publication of tables
 - Table description adopted by the other journals
- Many funding agencies make it mandatory for funded projects to give access to their results
 - Often understood as open access to publications
 - But more and more also access to result data
- Also widespread concern about the “long tail” of data (produced by research, beyond large datasets from instruments and surveys)
- Different ways of providing access to result data
 - “Drop box” – minimal information provided by data providers
 - Metadata allowing data reuse
 - Metadata allowing data integration with other data

Knowledge accumulated in SIMBAD over the years, linked to publications



VizieR data mine: homogeneous view of heterogeneous data



CDS role in data curation

- RESULTS VALIDATED by a refereed publication
- 14% of VizieR catalogues contain data which are not in tabular form (spectra, time series, images, cubes, etc.)
- Fast increase

	July 2012	August 2013	
Time series	812	1026	26%
spectra	268	349	30%
images	87	108	24%

- Agreement with A&A
- Other journals share the same concerns
- Discussion with AAS journals after January AAS : build a common framework with multiple poles

Planned consolidation activities



- Improve data visualisation in addition to SED and images
- Improve discoverability by direct publication in the VO
 - Fine grain indexing
 - Possible implementations to be assessed
- Improve instruction to authors to take data diversity better into account
- Continue the discussion with A&A scientific editors and with the other journals

Sharing expertise and lessons learnt



- Continue to respond positively to offers to participate in meetings, expert and audit groups, etc.
- In astronomy: participation in IVOA, Euro-VO leadership, support to French teams, participation in projects
- Beyond astronomy: RDA, “chargée de mission” INSU, etc.
- E.g. the team proposed new profiles for the job directory of “documentalistes”
- Collaboration with EUDAT on the interface with their “generic” elements of the scientific data infrastructure

Proposal



Include the development of the CDS role in the Open Data endeavour as one of our high level strategy drivers