



WP4 – DADI Data Access, Discovery and Interoperability

Françoise Genova et l'équipe WP4

ASTERICS

- Astronomy ESFRI Research Infrastructure Cluster
- Horizon2020 project Call September 2014 – Topic :
Implementation of cross-cutting solutions for clusters of ESFRI
research infrastructures
- Focus: ESFRIs - SKA, CTA, KM3Net, close links with
E-ELT, plus EGO (new messengers)
- 22 partners, 15 M€, including 4.5 M€ for WP4, 4 years, began 1
May 2015, coordinated by ASTRON (M. Garrett)
- Major collaboration Astronomy-Astrophysics/Astroparticle
physics
- Multi-wavelength/multi-messenger
- <https://www.asterics2020.eu/home>

WP4 high level objectives

Make the ESFRI and pathfinder project data available for **discovery** and **usage** by the whole astronomical community, **interoperable** in a **homogeneous international framework**, and accessible with a set of **common tools**.

European VO teams **AND** ESFRI/pathfinder teams are involved in **ALL** activities.

Fully aligned with the IVOA work and current priorities

- DADI (WP4) - Françoise Genova (CNRS-OAS):
4.5 M€



ASTERICS WP4: DADI (Data Access, Discovery and Interoperability)

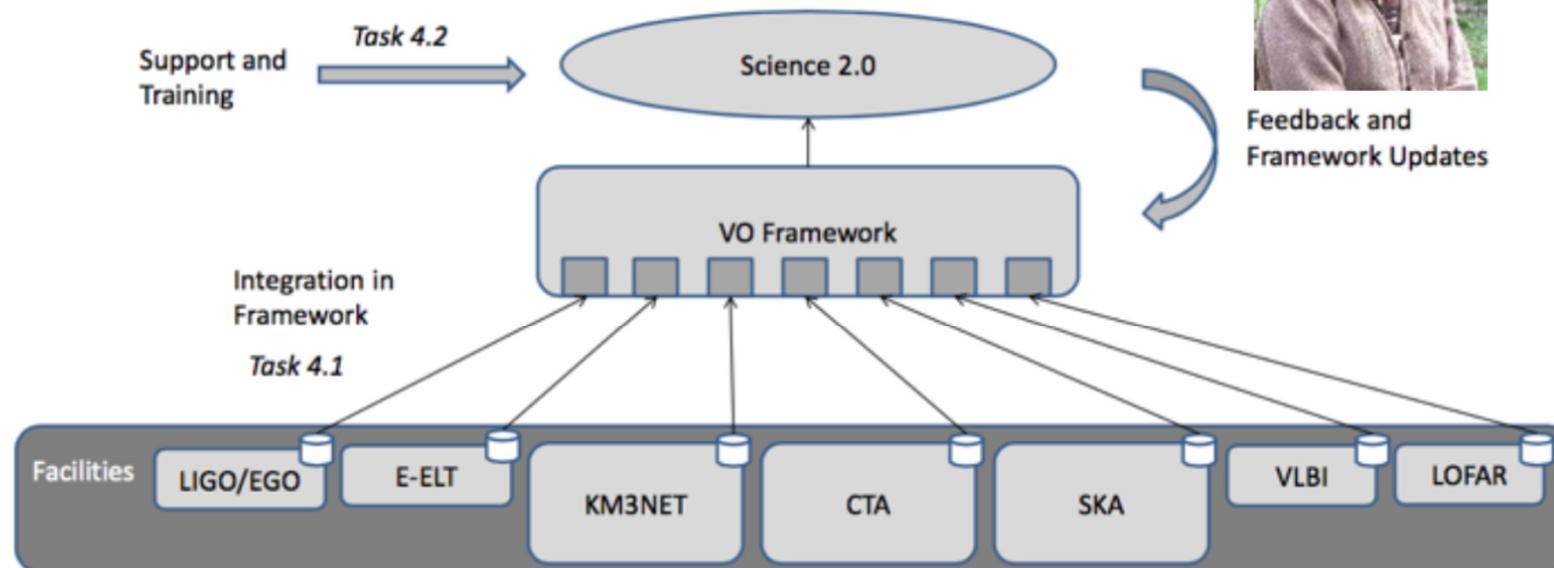


Figure 6: The ESFRI projects integrated in the VO Framework offers users uniform access.

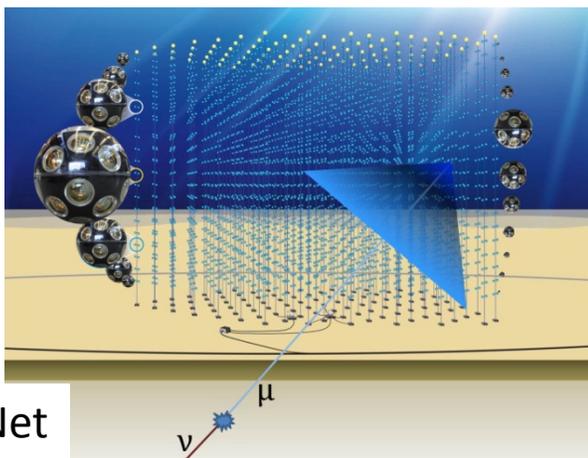
Who is involved



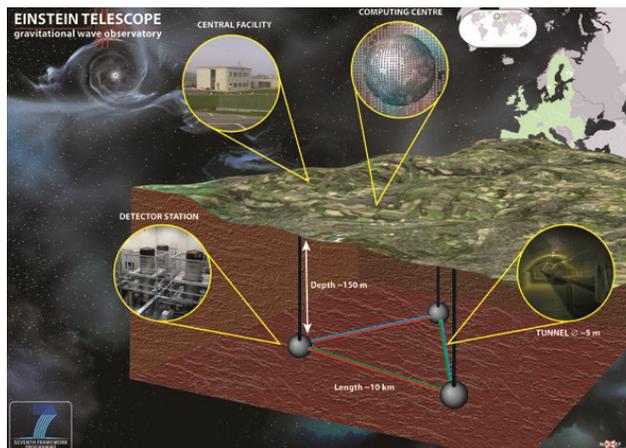
CTA



SKA



KM3Net



Who's involved

- Euro-VO partners, i.e. VO initiatives from France (CNRS/OAS- CDS+UNISTRA), Germany (UHEI), Italy (INAF), Spain (INTA), UK (UEDIN)
- Representatives of ESFRI and pathfinders
 - CTA (CNRS/LUTH + OBSPAR)
 - EGO/VIRGO and ET (CNRS/APC)
 - KM3Net (CNRS/CPPM)
 - SKA (ASTRON)
- ESO is associated to the project
- ESA (ESAC) is working in close collaboration with Euro-VO

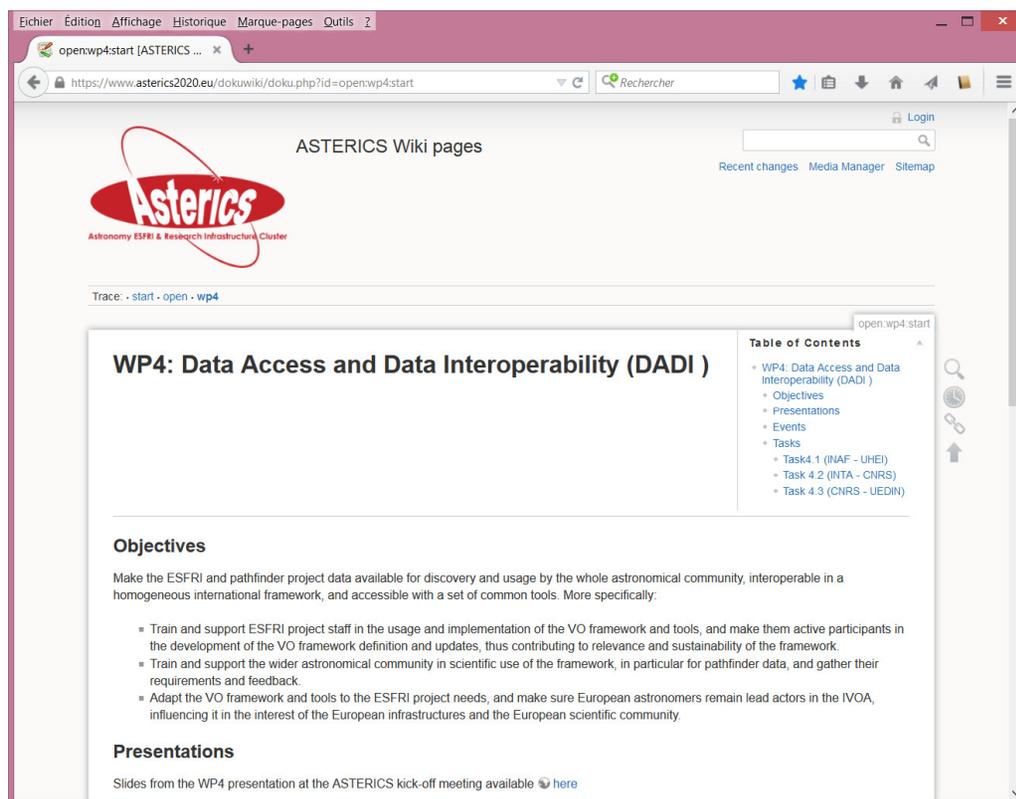
Targets

Three Tasks in support to three complementary targets

- Task 4.1: Support to astronomy ESFRI facilities, their pathfinders and other infrastructures of pan-European interest for implementation of their data in the VO framework (INAF/UHEI)
- Task 4.2: Support to the astronomical community (CNRS-CDS/INTA)
- Task 4.3: Updates of the VO framework from feedback and requirements (CNRS-CDS/UEDIN)

DADI Web site

<https://www.asterics2020.eu/dokuwiki/doku.php?id=open:wp4:start>



The screenshot shows a web browser window displaying the DADI web site. The browser's address bar shows the URL <https://www.asterics2020.eu/dokuwiki/doku.php?id=open:wp4:start>. The page title is "ASTERICS Wiki pages". The main content area features the "Astetrics" logo and the heading "WP4: Data Access and Data Interoperability (DADI)". Below the heading, there is a section for "Objectives" and a section for "Presentations". The "Objectives" section states: "Make the ESFRI and pathfinder project data available for discovery and usage by the whole astronomical community, interoperable in a homogeneous international framework, and accessible with a set of common tools. More specifically:" followed by a list of objectives. The "Presentations" section states: "Slides from the WP4 presentation at the ASTERICS kick-off meeting available [here](#)". A "Table of Contents" sidebar is visible on the right side of the page, listing the main page and sub-sections: "WP4: Data Access and Data Interoperability (DADI)", "Objectives", "Presentations", "Events", "Tasks", "Task 4.1 (INAF - UHEI)", "Task 4.2 (INTA - CNRS)", and "Task 4.3 (CNRS - UEDIN)".

Activities during the period

- DADI is structured around the organisation of Workshops (Deliverables) and participation in external events (IVOA meetings, which are Milestones, plus ADASS, RDA)
- Intense activity during the first 9 months
 - to create DADI community
 - to introduce newcomers to the VO
 - to share information on the partners' expertise, activities and needs
 - to begin to identify topics of common interest and collaboration paths
 - to perform technological work on the initial priorities, in particular multi-dimensional data (IVOA standards) – defining standards and agreeing on them takes time!

Deliverables

- One deliverable/task during the period
- Deliverables are Workshops, the « text deliverable» is provided several weeks afterwards

#	Title	Lead partner	Due date	Actual date and location
D4.1	First DADI Technology Forum	CNRS/CDS (Task 4.3)	September 2015	Held 17-18 Sept. Strasbourg Del. 3 Nov.
D4.2	FIRST ASTERICS European School	INTA (Task 4.2)	November 2015	Held 15-17 Dec. Madrid Del. being finalized
D4.3	First ESFRI Forum & Training Event	INAF (Task 4.1)	November 2015	Held 3-4 Dec. Trieste Del. Being finalized

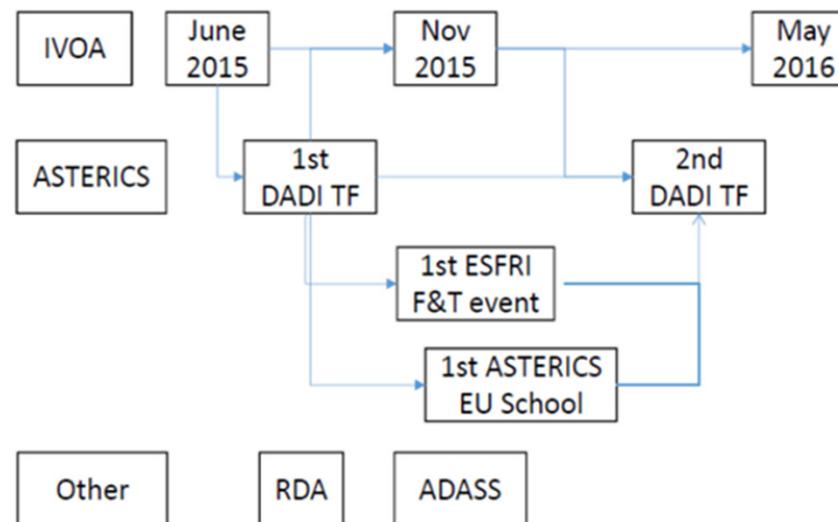
Milestones and other relevant meetings

- The IVOA Interoperability meetings are milestones

M2	IVOA Sesto (Italy)	14-19 June 2015
M5	IVOA Sydney (Australia)	30 October-1 November 2015

- Other relevant meetings
 - ADASS XXVth Meeting, Sydney, 25-29 November 2015
 - RDA Sixth Plenary, Paris (France), 23-25 October 2015

DADI events in a nutshell



D4.1 First Technology Forum

- 17-18 September 2015, Strasbourg
- 34 participants (except KM3Net-VLVvT), including ESO and ESA
- DADI kick-off
 - Information about ASTERICS and DADI
 - Presentation of the VO Framework by the IVOA Chair (C. Arviset, ESA)
 - Presentation of partners' expertise and on-going work
 - « Hack-a-thon » i.e. discussions by small groups on topics identified on the spot
- Outcomes
 - List of topics of interest for the ESFRI Forum D4.3
 - Preparation of Sydney IVOA meeting
 - Status of multi-D standards
 - HiPS strategy



D4.3 First ESFRI Forum & Training Event

- 3-4 December 2015, Trieste
- 25 participants, all DADI partners and ESO
- Centered on the ESFRI and pathfinder requirements
- Organised around the topics identified during the Tech Forum
 - Multi-dimensional data access
 - Time domain data access (<> WP5)
 - VO registry of resources
 - Authentication & Autorisation (<>WP3)
- VO standards and tools already in use
- Discussion on specific support activities initiated with the ESFRIs and pathfinders



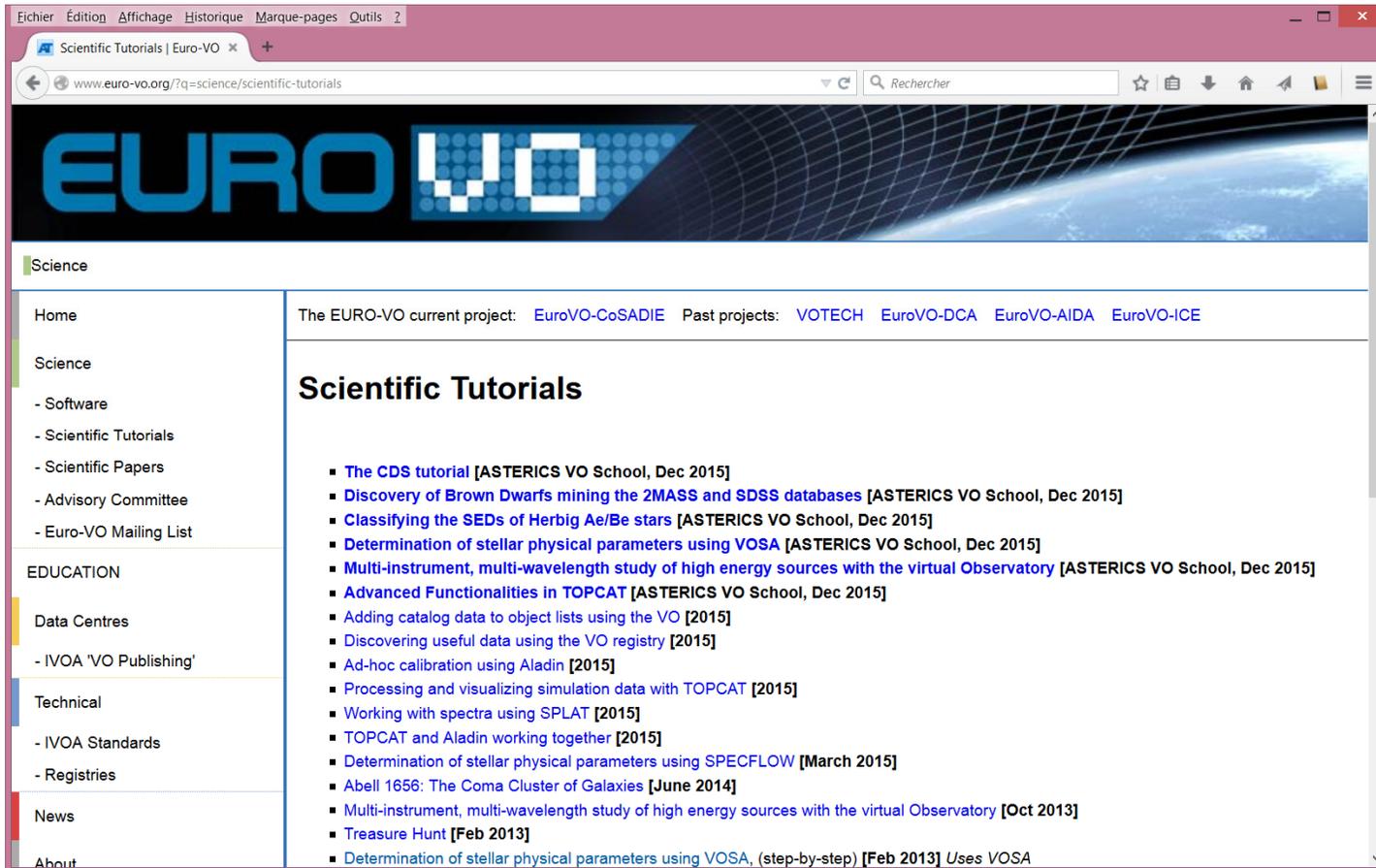
D4.2 First DADI School

- 15-17 December 2015, Madrid
- 43 participants from France, Germany, Italy, Spain, UK, and from Belgium, Greece, Lithuania, Poland, Portugal and Slovakia, also CTA, LOFAR, ESO
- 11 tutors
- Learn enough about the VO to be able to use it in one's own research
 - Short presentation of the IVOA and ASTERICs
 - “Hands-on” tutorials & “Treasure Hunt”
 - Participants' own projects
 - Presentation of some of the projects and feedback session
- Significant preparatory work
 - Update of the tutorials
 - Discussion with the participants about their scientific interests
- Follow-up
 - Two seminars in Athens in 2016
 - The participants will be polled each year to follow the School impact and monitor their usage of the VO



Tutorials

<http://www.euro-vo.org/?q=science/scientific-tutorials>



The screenshot shows a web browser window displaying the Euro-VO website. The browser's address bar shows the URL www.euro-vo.org/?q=science/scientific-tutorials. The page features a large header with the "EURO VO" logo and a background image of a planet. Below the header, there is a navigation menu on the left with categories like Home, Science, EDUCATION, Technical, News, and About. The main content area is titled "Scientific Tutorials" and lists various tutorial topics, each with a date in brackets. The list includes:

- [The CDS tutorial](#) [ASTERICS VO School, Dec 2015]
- [Discovery of Brown Dwarfs mining the 2MASS and SDSS databases](#) [ASTERICS VO School, Dec 2015]
- [Classifying the SEDs of Herbig Ae/Be stars](#) [ASTERICS VO School, Dec 2015]
- [Determination of stellar physical parameters using VOSA](#) [ASTERICS VO School, Dec 2015]
- [Multi-instrument, multi-wavelength study of high energy sources with the virtual Observatory](#) [ASTERICS VO School, Dec 2015]
- [Advanced Functionalities in TOPCAT](#) [ASTERICS VO School, Dec 2015]
- [Adding catalog data to object lists using the VO](#) [2015]
- [Discovering useful data using the VO registry](#) [2015]
- [Ad-hoc calibration using Aladin](#) [2015]
- [Processing and visualizing simulation data with TOPCAT](#) [2015]
- [Working with spectra using SPLAT](#) [2015]
- [TOPCAT and Aladin working together](#) [2015]
- [Determination of stellar physical parameters using SPECFLOW](#) [March 2015]
- [Abell 1656: The Coma Cluster of Galaxies](#) [June 2014]
- [Multi-instrument, multi-wavelength study of high energy sources with the virtual Observatory](#) [Oct 2013]
- [Treasure Hunt](#) [Feb 2013]
- [Determination of stellar physical parameters using VOSA, \(step-by-step\)](#) [Feb 2013] Uses VOSA

D4.4 Second Technology Forum

- 7-8 mars 2016, Edimbourg
- 26 participants
- Technical discussions
 - Talks & ‘open’ Hack-a-Thon breakout sessions on different subjects, including several discussed in this AOV meeting
 - <https://www.asterics2020.eu/dokuwiki/doku.php?id=open:wp4:wp4techforum2> being completed



IVOA

- DADI technological activities (Task 4.3) are performed in the IVOA context
- Sesto (June): presentation of DADI and work on multi-d standards
- Sydney (October)
 - Many meeting highlights relevant to DADI
 - Well prepared by the Tech Forum
 - Significant impact of participants linked to DADI
 - At least one chair or co-chair of all the Groups present at the meeting except one
 - Significant participation in the Groups

ADASS XXVth

- The place to be to discuss with large projects about their data
- ASTERICS talk (F. Pasian)
- IVOA talk (C. Arviset)
- VO ubiquitous in the talks and posters, very positively mentioned in B. Schmidt keynote address
- Two ASTERICS posters

IVOA Provenance Data Model: Hints from the CTA provenance prototype

Michèle Ségouin¹, Mathieu Scaife², Maricela Lopez³,
 François Dubillard⁴, Catherine Besson⁵, Johan Bergman⁶
¹ICFO/PIRA, ²SOAR/PIRA, ³ADFF/PIRA

Abstract

We present the last developments on the IVOA Provenance data model, mainly based on the W3C PROV concept. In the context of the Cherenkov astronomy, the data processing stages imply both assumptions and comparisons. As a consequence, Provenance information is crucial to the end user in order to interpret the high level data products. The Cherenkov Telescope Array (CTA), currently in preparation, is thus a perfect test case for the development of an IVOA standard on Provenance information. Contrary to previous Cherenkov experiments, CTA will serve as an open observatory to a wide astronomical community and high level data products, sky maps, sky maps will be made available through the Virtual Observatory. We describe general use-cases for the computational Provenance in the CTA production pipeline and explore the proposed W3C notations the PROV-R format, as well as Provenance access solutions.

Cherenkov Astronomy Context: Complex data

- Observation is indirect
- Reconstruction is based on detailed simulations
- Complex workflow and data level structure
- User need data to understand and reproduce the data

Users need:

- to know what we are talking about: **Data Model**
- to know how data sets were produced: **Provenance description**
- to select data on provenance criteria: **Query**

Provenance Data Models

W3C Provenance Data Model
Simple and general concept to trace the activities performed on Entities by Agents.

VO Provenance Data Model
adapted to astronomy data and data processing activities.

Provenance description

INTEROPERABILITY: STANDARDIZED DESCRIPTION LANGUAGE PROV-R

The W3C model does not include the concepts of Activity Collections or Instances. We get round the problem by describing on the one hand a workflow and on the second hand independently the activities that comprise it. The interoperability relation is used to indicate the beginning and end of the workflow.

Provenance access

The objective is to enable user queries on Provenance information, in a standardized way.

Selection criteria could be:

- Name of attributes of the W3C or IVOA Provenance data model
- Name of the attributes specific to the CTA context (run number, arbitrary conditions, ...)

We need to identify the specific provenance terms for each data level.

The VO user need to query which specific attributes could be a criteria.

IVOA Data access layer project

Bonnafant Francois, Dowler Patrick, Nockle Keith, Tody Douglas [COS, CADC, University Edinburgh, NRAO]
 Contact: francois.bonnafant@cea.fr

IVOA DAL Working Group Project and Objectives

A group formed at the very beginning of international VO efforts. Data is no longer produced in silos.

- data discovery
- data retrieval
- data access processing
- providing other data access, in dynamically composed subsets of service astronomical processes

Interoperability with:

- Data Model: query based on data model concepts, responses are detailed instances of data models
- Resource: use of descriptive metadata
- Registry: standardized description and discovery of the services
- VO: application made support

Historical protocols

- Discovery**
 - simple search: simple HTTP requests for listing of records, VOTABLE support
 - single image browser: Content and other additional parameters concerns
 - VOTABLE query responses for data discovery giving standardized image description, instance and Resource generation
 - single data browser
- Generation**
 - single image browser: the IVOA has expressed in terms of a comprehensive general data model
 - HTTP interoperable resource made query services

Simple Spectral Access

- Standard query parameters in description of data
- Additional spectral related parameters (WAVELENGTH, SPEED, CHANNEL, etc.)
- Query responses in VOTABLE
- Organized in sub-URLs required from Spectrum Data Model packaging: Format, Location, Format, Characteristics, Coordinates, etc.
- Wavelength and speed support with queries
- Virtual data generation
- services speed color or image generate a "fast search" to query parameters "locally"

TAP and ADQL

A standard protocol to retrieve and query relational tables.

ADQL for astrophysics applications made in compliance with IVOA Cherenkov Monitor Services standard.

- Location and Model exposed with the full ACID and ODBC-like descriptions: relational metadata
- Query via ADQL, language parameterized selection of SQL
- The SQL standard data model effectively exposed as a TAP service to support uniform cross observatory data discovery

Recent or current developments of new protocols

- collection of service capabilities (parameters) to provide multi-dimensional data discovery and access
- parameter phrase
- Calculate: Transformation methods to link resources to input datasets
- SQL: parameterized query (SQL) of the OData 1.1 data model for simple interoperable data discovery
- Accession: driving formal data processing for accepting information from the database

Lessons learned from DAL development

- service use cases often involve using a collection of services in a sequence of steps
 - in general, one service never always satisfy requirements
- having clearer interoperability
 - single purpose service capabilities that can be combined to support diverse use cases
- common parameters and features
- additional data Access Layer Interface (DAL) for easy to use and maintaining
- Test a framework of protocols sets
 - standardizable protocols (more complex, more powerful) which align model protocols (simple, implementation flexibility, stability)
 - open generation of Resource Query Language (RQL)
 - Application interoperability: **VOA** handling

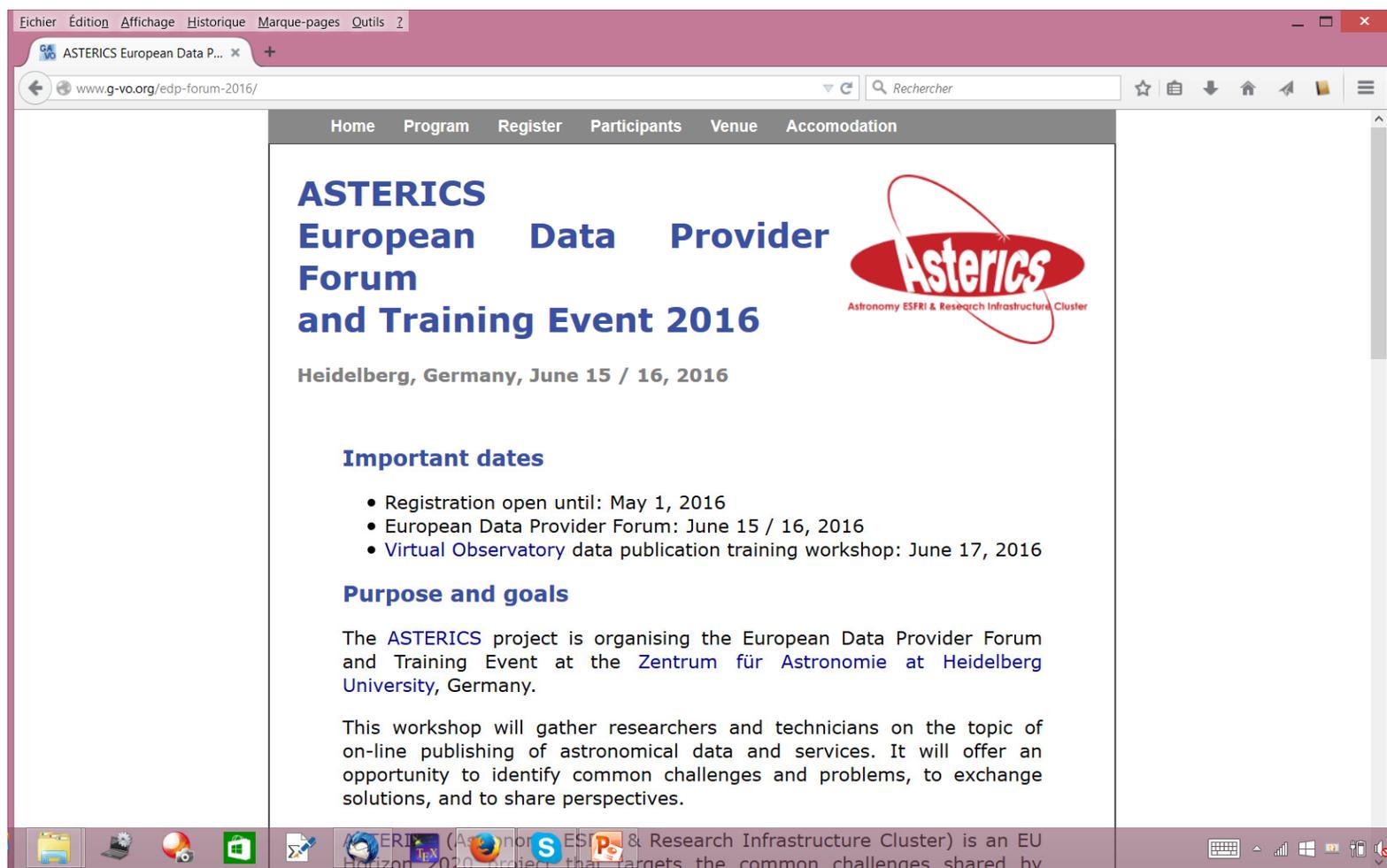
Among the other activities

- Technical work on Provenance (VO/CTA)
- Collaboration between CDS and EGO on Aladin customization
- ANTARES data in GAVO
- Autorisation&Authentication
- Collaboration with EUDAT (i.e. with the « generic » data infrastructure)

Next steps

- Discussion on specific support actions with the ESFRIs
- Cape Town IVOA meeting (May 2016)
 - Focus sessions on large projects, organised by M. Allen
 - VO standards and tools including multi-D and time domain
- European Data Provider Forum & Training Event, Heidelberg, 15-16 + 17 Juin 2016
- Second School, Strasbourg, November 2016 (TBC)

<http://www.g-vo.org/edp-forum-2016/>



The screenshot shows a web browser window displaying the website for the ASTERICS European Data Provider Forum and Training Event 2016. The browser's address bar shows the URL www.g-vo.org/edp-forum-2016/. The website has a navigation menu with links for Home, Program, Register, Participants, Venue, and Accommodation. The main content area features the event title, location, and dates, followed by a list of important dates and a section on the purpose and goals of the event.

Home Program Register Participants Venue Accommodation

ASTERICS European Data Provider Forum and Training Event 2016

Heidelberg, Germany, June 15 / 16, 2016

Important dates

- Registration open until: May 1, 2016
- European Data Provider Forum: June 15 / 16, 2016
- Virtual Observatory data publication training workshop: June 17, 2016

Purpose and goals

The ASTERICS project is organising the European Data Provider Forum and Training Event at the Zentrum für Astronomie at Heidelberg University, Germany.

This workshop will gather researchers and technicians on the topic of on-line publishing of astronomical data and services. It will offer an opportunity to identify common challenges and problems, to exchange solutions, and to share perspectives.

ASTERICS et l'ASOV

- Point sur les activités aux réunions plénières
- Support à des participations aux réunions IVOA et à certaines réunions techniques
- Formation des chercheurs
 - Les tutoriels peuvent être réutilisés
 - Participation à des formations locales ou nationales
 - Nouveaux sujets de tutoriels bienvenus (cf 'high Energy', produit en collaboration avec PCHE)
- Le Data Provider Forum & Training Event devrait être un évènement intéressant (cf la réunion CoSADIE en 2013)