

Euro-VO Data Centre Alliance



Mark ALLEN

DCA Project Scientist

Centre de Données de Strasbourg (CDS), FRANCE

allen@astro.u-strasbg.fr

Outline

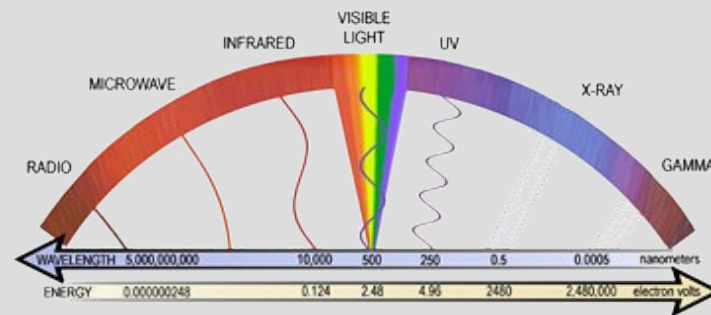
- Providing and using data *'in the VO'*
- Coordination of Euro data centres
- Scientific guidance

Providing and using data in the VO

Need driven by **science**, and **technical necessity** of the data volumes

Science:

Multi- λ Astronomy

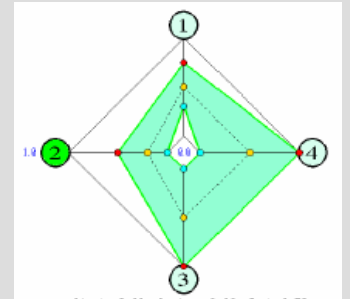


- **↑** % publ. science is from several instruments
- Scientific combination of data (e.g. SEDs)
- Data mining of federated multi- λ archives

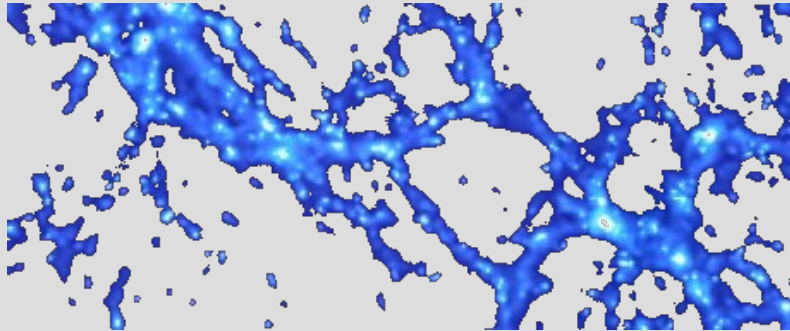
Science:

VO as a discovery tool

- Statistical analyses on large numbers of sources in multi-dimensional data sets



Comparison with models and simulation data



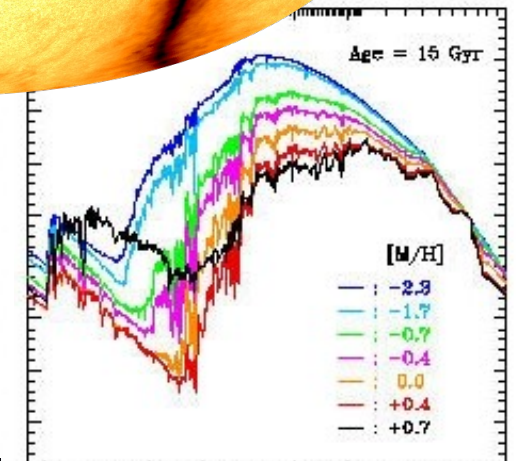
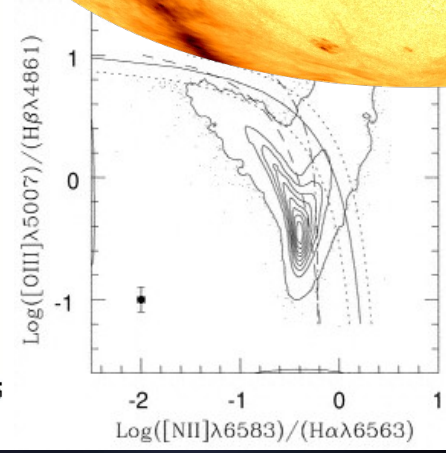
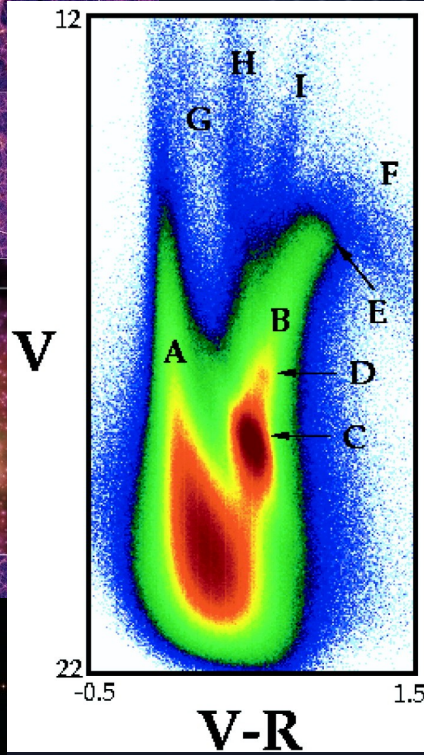
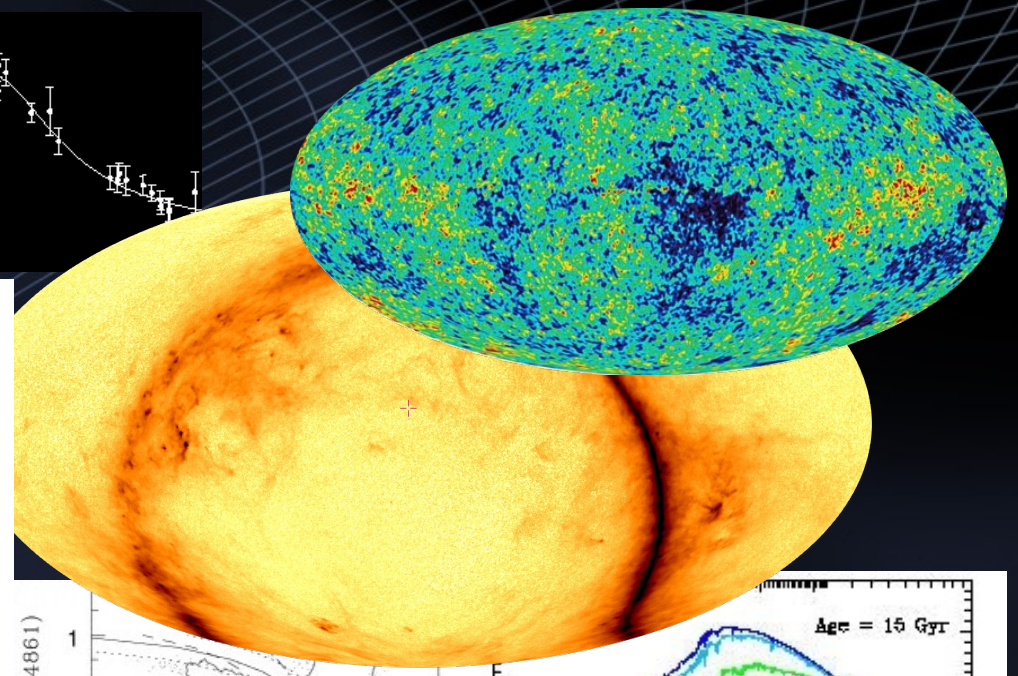
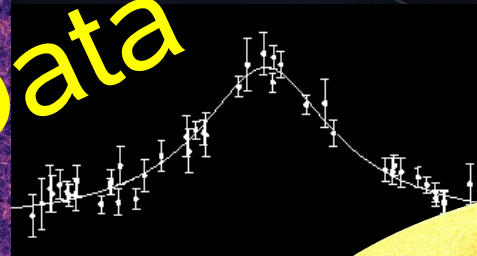
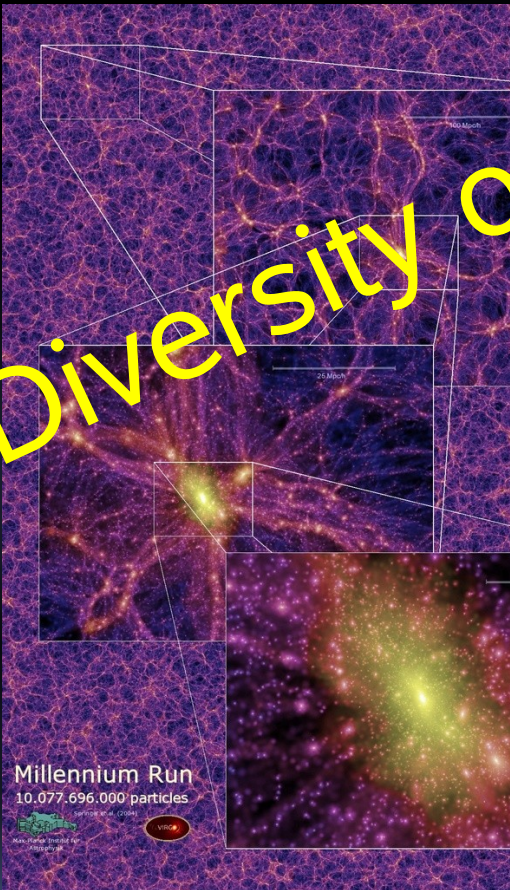
- '*Virtual observations*' of simulations

Other general scientific motivations include:

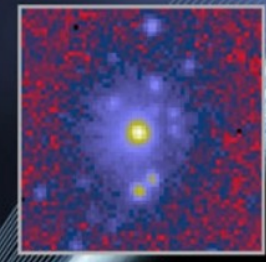
- Time domain astronomy - Multi-epoch/transient sources
- Fast response astronomy - e.g GRB follow-up
- Standard access to physical data e.g atomic line databases

Diversity of Data

Millennium Run
10,077,696,000 particles



Messier 81



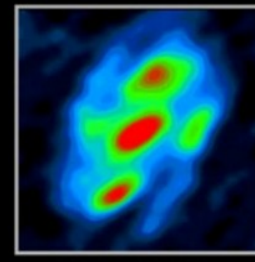
X-Ray (ROSAT)



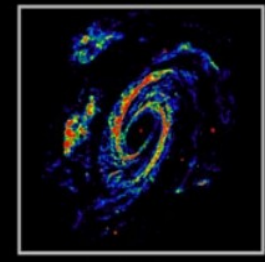
Ultraviolet (ASTRO-1)



Visible (Robert Gendler)



Infrared (IRAS)

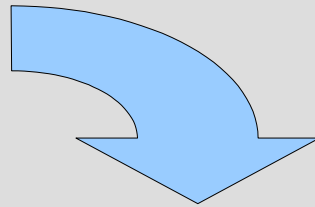


Radio (VLA)

Current issues for multi-archive use

At present, using data from:

- Different archives
- Different instruments
- On-line services



Requires:

- learning specific user interfaces
- access and download procedures
- data access tools
- making data scientifically compatible

Goals for providing data in the VO

- **Data Access**

- *Simultaneous across archives*
- *Transparent interfaces*

- **Federation of multi- λ archives**

- *possibility of applying analysis and statistical tools to entire data and service grid*
- *scientifically valid data aggregation*

Global standards for building the VO

- Standards & protocols for
 - publishing data to the VO
 - access mechanisms
 - data description
 - use of distributed resources



Coordination of European Data Centres



'Technology take-up and full VO compliant data and resource provision by astronomical data centres in Europe'

...Populate the system with data

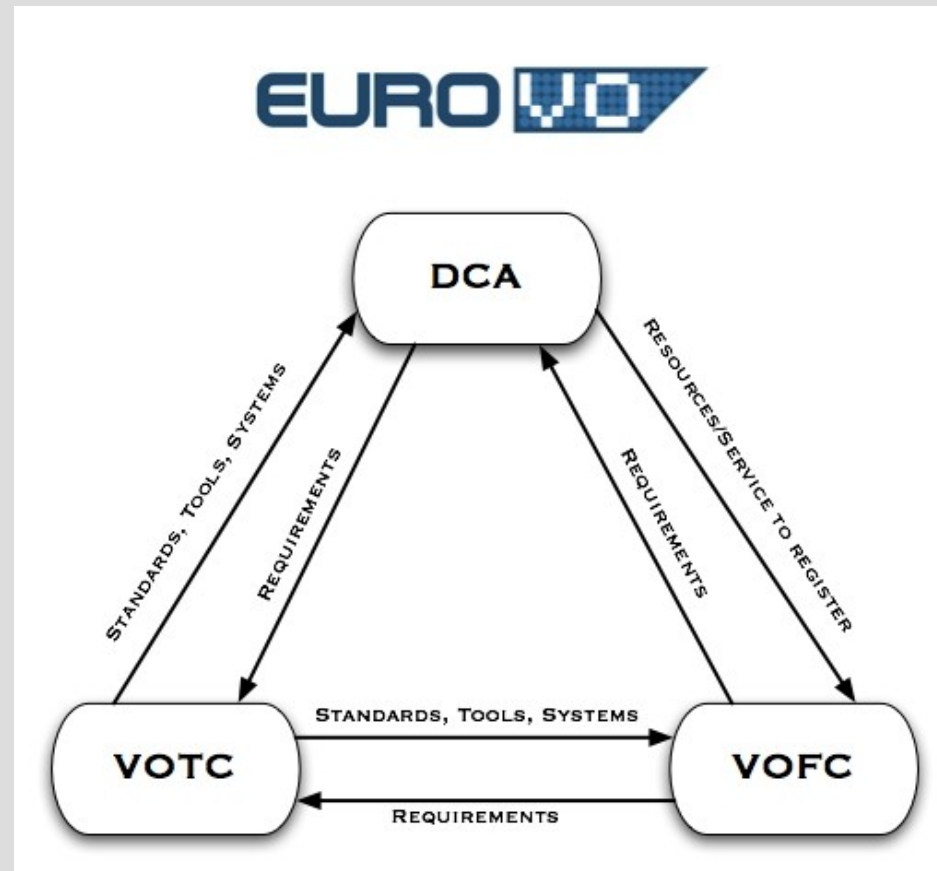
Funded EC Sixth Framework Program

- “Coordination Action”
- 8 partners across Europe



- Starting September 2006

DCA within EURO-VO



DCA Goals

- **Coordinate** European VO initiatives
 - define European DCA strategy
- **Disseminate knowledge** and good practice
 - support to uptake of interoperability standards
- **Feedback** from real implementations
 - VO-TECH & IVOA
- Prepare inclusion of theory in VO framework
- Coordinate with computational Grids

Document

Working Systems

Coordination activities

- **Data Centre Census**
- **Support to take-up of VO framework**
 - Support staff at partners - lead by example
 - Two EuroVO DCA Workshops
 - Detailed on-line tutorials and examples
 - Visits of experienced developers to meetings and sites
- **Theory in VO** – workshop, assessment
- **Computational Grids** – workshop, access to computational grids for data centres

Data Centres

- **Observation archives**
- **Services for data management, data processing, visualization**
- **Added-value databases**
- **‘Software instruments’ to tackle specific scientific questions**
- **‘Thematic portals’**

Data Centre Willingness

- **Initial limited surveys show strong desire to participate**
 - VOFrance – 41 responses from 24 labs
 - Wide range of data curators/services
- **Work to be done to encourage future large projects to be VO compliant**
 - DCA will engage
 - EuroVO SAC support extremely valuable

Scientific Guidance

- DCA Internal Science Team (to be formed)
 - scientists from partner teams
- Check global scientific coherence of DCA activities
 - interoperability between data centres
- Highlight scientific drivers to motivate adoption of DCA take-up
 - immediate scientific benefits
 - longer term benefits of VO science

Scientific coherence

- Data description is sufficient for
 - combining data scientifically
 - units ▪ calibrations ▪ photometric systems
 - uncertainties ▪ resolutions ▪ PSFs ▪ apertures
 - sampling ...
 - access and use with VO tools
- Guiding science cases
 - tests with real science
 - provide best examples and motivators
 - scientific opportunities

Challenges for scientific interoperability

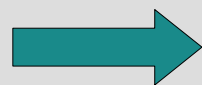
- Diversity of data, data providers
- Getting data centres to provide 'Science Ready' data
- Standards for describing photometric systems or calibrations – **things you commonly look up in manuals or in the details of papers**
- Interoperability of VO tools with astronomer's everyday tools and own programs

Summary



Coordinating data centres
in Europe for VO take-up

- Enable multi- λ , and cross-archive science
- Scientific as well as technical challenges to get things working



Your help needed

Reference sources

- Spectrum figure: Modified from <http://www.spacetoday.org/DeepSpace/Telescopes/GreatObservatories/Chandra/ChandraSpectrum.htm>
- N-body simulation figure: Modified from Astronomy magazine, Jan 26, 2005. Institute of Theoretical Physics, University of Zurich
- Multi-d data mining figure from: <http://www.roe.ac.uk/~rgm/sc4devo/9richard-holbrey-mdv.pdf>