Strasbourg astronomical data centre (CDS)

42 years serving the community

Francoise Genova
An old (a very old) data centre!

- A 42-year old history
- Founded by INAG – now CNRS-INSU (in charge of French ground-based astronomy) in 1972, in agreement with the Université Louis Pasteur – now Université de Strasbourg
- A far-seeing vision
  - Collect ‘useful’ data on astronomical objects, in electronic form
  - Improve them by critical evaluation and combination
  - Distribute the results to the international community
  - Conduct research using the data
- Supported on astronomy money (mainly national) + contracts incl. EC
- Users ARE trusting us (~1 million queries/day)
- But external certification is useful
  - Labelled as “Research Infrastructure” at the national level (as well as telescopes...)
  - Regular member of the World Data System
  - Awarded the Data Seal of Approval
CDS role

• Support the international astronomical community in its research tasks (not only collect/preserve information), *science driven* (not technically driven)

  **Core task:** provide highly used value-added services (~1 million queries/day currently)

  **Keywords:** quality, scientific and technical relevance, collaborations, global networking of expertise and resources

  **Staff:** an integrated team of scientists, data librarians and computer engineers (~30 staff members, 1/3 on each profile)

• Close collaboration with academic journals, space and ground-based telescope archives, ADS, other data centres – a pioneer and an active member of the astronomy distributed data infrastructure

• Support to projects, including a major role in the Virtual Observatory (set of interoperability standards and tool enabling seamless access to on-line data)
CDS activities

• Reference, added-value services
  The basis of CDS service to the community

• Technological watch, R&D
  One key of long term sustainability
  The outcome: new services, new functionalities for the users and for ourselves

• Virtual Observatory
  Interoperability standards, tools (eg: the VO portal for images)

• Participation in projects
  Eg Gaia, RDA

• User support, knowledge dissemination, expertise

• Science
  Scientific staff are ACTIVE scientists
  Another key of long term sustainability
The CDS hub

Astronomical objects
identification, bibliography, data, measurements

Federation of tabular data – but not only + more and more non-tabular “long tail”
catalogues, published tables observation logs, very large surveys

Interactive sky atlas: Data discovery, integration, visualisation, manipulation
images, databases, catalogues, surveys, archives, user data
Among the challenges

<table>
<thead>
<tr>
<th>SIMBAD</th>
<th>2008</th>
<th>2011</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objects</td>
<td>3.900.000</td>
<td>5.400.000</td>
<td>7.750.000</td>
</tr>
<tr>
<td>Identifiers</td>
<td>11.750.000</td>
<td>15.124.000</td>
<td>18.551.617</td>
</tr>
<tr>
<td>References</td>
<td>216.000</td>
<td>253.000</td>
<td>294.000</td>
</tr>
<tr>
<td>Citations</td>
<td>5.750.000</td>
<td>8.180.000</td>
<td>10.740.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VizieR</th>
<th>2008</th>
<th>2011</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catalogues</td>
<td>6.500</td>
<td>9.100</td>
<td>12.650</td>
</tr>
</tbody>
</table>

• Growth in volume and complexity – not in staff -
• Semi-automated ingestion but QUALITY kept, ie human experts’ key role
Among the challenges

- The requirements can be somehow contradictory
- Rapid evolution of astronomy & technology
- Long time sustainability
- Operational constraints (24/24, 7/7, significant usage increase)
- Agile strategy needed, taking all aspects into account
- Staff members are experts with rare profiles – permanent staff
- Support from the community mandatory for sustainability

- Data bases ... **AND LINKS**
SIMBAD usage in ADS and telescope archives

Simbad interface

SIMBAD Name Resolver usage in ESO Archive

If you would like to query the Archive for instrument specific parameters, please use the dedicated. To search for reduced Data Products, please have a look at the ESO Data Products page.

The checkboxes on the right of the parameters define whether or not they will be displayed on the output.

in ESA ISOC Science Data Archive

in STScI MAST Archive
Integration of observational data

Images are built on-the-fly with data extracted from remote observatory archives, from the CDS reference image data sets, and with user provided data.
VO allows interoperability of data and tools
But also long tail, at least a kind of...

- CDS hosts data « attached » to academic publications: tables, images, etc, in VizieR which also hosts object catalogues from ground- and space based observatories (up to 2 billion rows)

- Agreement between CDS and the academic journals (started in 1993 with Astronomy & Astrophysics)
- Starting point: tabular data from publications
  
  Printed figures became usable data!

- More and more non-tabular data « attached to publications » (images, spectra, time series) - A&A: the scientific editor asks the authors to deposit their data

- More than 12 000 « catalogues »

- Standard description, agreed with the journals

- Hundreds of different quantities, well qualified by a UCD (VO standard)
Data validated by a publication Fully discoverable (OAI-PMH VO registry) and usable including numbers inside the tables

“Photometry viewer”: Spectral points extracted from the collection