European Grid of Solar Observations (EGSO)

Dr Robert Bentley
EGSO Project Coordinator
University College London
(Dept. Space and Climate Physics)

EGSO – European Grid of Solar Observations

- EGSO is a Grid test-bed addressing a particular application
  - Designed to improve access to solar data for the solar physics and other communities
  - Addresses the problem of a distributed heterogeneous data set and a scattered user community
- Funded under the EC’s Information Society Technologies (IST) thematic programme of Framework 5
  - Started March 2002; duration of 36 months
- Ten partners in Europe and the US, led by UCL-MSSL
  - 3 in UK, 2 in France, 2 in Italy, 1 in Switzerland, 2 in US
- Currently gathering user and science requirements and talking to data providers
  - Detailed design work starts in September
- Desirable to have EGSO and VSO collaborate as closely as possible
  - Can achieve more together
The Problem addressed

- Observations used to build up a picture of the plasma in multi-dimensional parameter space (incl. \(x, y, z, t, T\) & \(\mathbf{r}\))
  - Users need access to as many wavelengths as possible
  - For technical and practical reasons:
    - UV, EUV, X-rays and \(\gamma\)-rays observed from space
    - Radio and optical wavelengths observed from the ground (coverage)
  
- Data centres and observatories located around the world
  - Large and small facilities (with varying resources)
  - Aim is to make it easy to add new data sets
  - Increasing data volumes require new methodology
  
- Users scattered around the world
  - Do not need to know where the data is located
  - Capabilities of users computing vary greatly
  - Authentication issue needs serious consideration
  - Want to minimize how this affects the user

EGSO Search Engine

- Enhanced cataloguing used to select observations
  - Metadata versions of observing catalogues tie together heterogeneous data sets – data itself left untouched
  - New types of catalogues allow searches on events, features and phenomena rather than just date & time, pointing, etc…
  - Ancillary data (images, time series, etc.) provide additional search criteria and enhance the capabilities of the (Graphic) User Interface

- Input can be local to search engine (e.g. at “catalogue warehouses”) and/or remote sources
- Users able to temporarily include their own catalogues, etc. as search engine inputs
- Catalogue Registry allows hierarchical optimization

- Alternate entry point (to User Interface) allows access by researchers from other disciplines:
  - Astrophysics, Climate Physics, Space Weather…
Handling the data

- An objective is to dramatically enhance access to the data
  - User only needs to know observations exist, not where located
  - Addition of new sources made as simple as possible
  - System able to optimize use of sources (closest, least used...)

- Process as much data at source as possible
  - Solar data are usually stored in raw form
    - Extraction and calibration done on the fly
    - Software for processing defined by instrument team (IDL, C...)
  - Processing reduces volumes of data moved around
  - Simplifies requirements on user’s own system

- Standard (pipe-line) processing adequate for many users
  - Extract and calibrate the selected observations

- More complex problems require ability to uplink code
  - Used in analysis of extended data sets (helioseismology, etc)
  - System allocates resources (CPU, disk space, etc)
  - Security and integrity of source must be ensured

Who is involved

- University College London (UK) — Coordinator
  - Dept. of Space and Climate Physics (MSSL)
  - Dept. of Computer Science **
- Rutherford Appleton Laboratory (**) (UK)
- University of Bradford, Dept. of Cybernetics ** (UK)
- Institute d’Astrophysique Spatiale ** (France)
- Observatoire de Paris-Meudon (France)
- Istituto Nazionale di Astrofisica (Italy)
  - Includes Observatories of Turin, Florence, Naples and Trieste
- Politecnico di Torino, Dept. Automation & Informatics ** (Italy)
- University of Applied Sciences, Dept. of Computer Science ** (Switz.)
- Solar Data Analysis Center (NASA-GSFC) (USA)
- National Solar Observatory (USA)

Partners selected to include expertise in space and ground-based observations, and in computer science

Notes: ** Groups with IT expertise
Contact details

- EGSO homepage:
  - http://www.mssl.ucl.ac.uk/grid/egso

The web page pages are still being created and will shortly include things like:
- A bulletin board, Wiki, Forum
- Information on the project and how you can collaborate
- Etc...

Please use the e-mail address on the web page, or contact rdb@mssl.ucl.ac.uk
Handling the data

- An objective of EGSO is to dramatically enhance access to the data
  - User does not need to know where the data comes from
  - System able to optimize use of sources

- Process as much data at source as possible
  - Reduces volumes of data moved around
  - Simplifies requirements on user’s own system

- Standard (pipe-line) processing adequate for many users
  - Extract and calibrate the selected observations

- More complex problems need ability for user to uplink their own code to data source
  - Security and integrity of source must be ensured

Layers in the solar atmosphere
The appearance of the Sun changes dramatically with wavelength.

For a complete picture, scientists need to use as wide a range of observations as possible.

Identifying what observations were made and then retrieving them is a major obstacle.