



Management of the GenerationCP Central Registry (SP4)

Collaborators: Tom Hazekamp (Bioversity International), Guy Davenport (CIMMYT),
Samy Gaiji (Bioversity International), Raj Sood (Bioversity International),
Milko Skofic (Bioversity International)

The Central Registry provides the GCP with a central facility where datasets can be registered, stored and shared with the entire GCP community and beyond. GCP partners are obliged to register datasets resulting from GCP funded research at the Central Registry.

[HTTP://DATA.GENERATIONCP.ORG](http://data.generationcp.org)

The Central Registry can be used:

- To register and upload datasets
- To search for datasets from a central point
- To download files and access other on-line resources

DATASET REGISTRATION AND FILE UPLOAD

The GCP Central Registry has been operational since November 2005.

Registration

As of 15 August 2007, there were 106 datasets registered at the Central Registry (see top graph).

Datasets can be registered BEFORE the actual data files exist! The registration requires only a description of the dataset including its availability and who to contact.

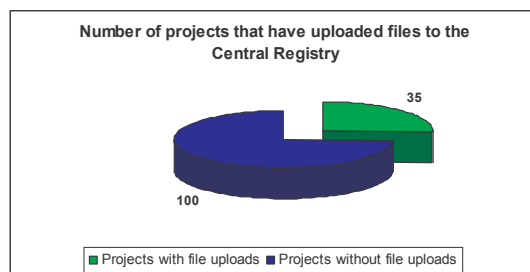
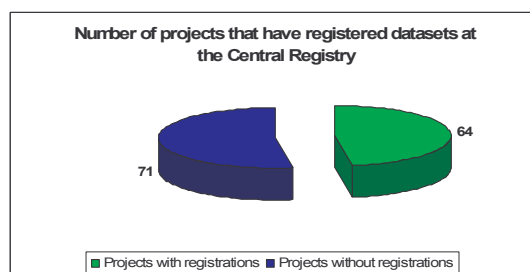
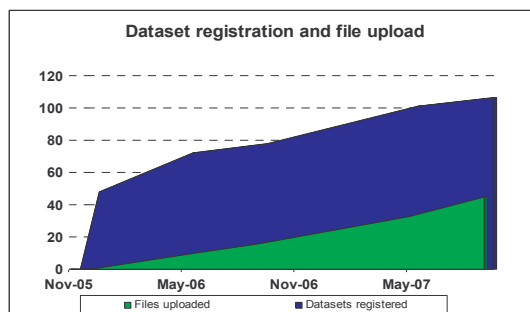
The 106 datasets that were registered originate from 64 different projects. These account for almost half (47%) of the 135 GCP projects that are expected to generate research data (see middle graph).

File uploads

There were 45 files uploaded to the Central Registry as of 15 August 2007 (see top graph). These files originate from 35 projects. This corresponds to a quarter of the all projects that are expected to generate data (see bottom graph).

There has been a steady increase in the number of datasets registered and files uploaded. Many projects have yet to submit their datasets, however.

Please check-out the Central Registry on-line!



Further developments for 2007

In collaboration with the GCP Data template project, led by CIMMYT, data validation tools are developed to enhance the quality control on submitted data

Those datasets which are available as file downloads will be made available as Webservices. This will allow a seamless integration of the data stored at the Central Registry with the tools of the bio-informatics platform developed by the GCP