Task 26: Development of the GCP Central Repository

**Collaborators:** Richard Bruskiewich (IRRI), S. Chandra (ICRISAT), Brigitte Courtois (CIRAD), Guy Davenport (CIMMYT), Dag Endresen (IPGRI/NGB), Samy Gaiji (IPGRI), Francis Moonan (IITA), Reinhard Simon (CIP), Raj Sood (IPGRI), Milko Skofic (IPGRI), Tom Hazekamp (IPGRI)

The Central Repository provides the GCP with a central facility where digital resources can be stored and shared with the entire GCP community (and beyond)

The Central Repository can be used:
- **To upload** virtually any type of digital research resource
- **To query** these resources from a central point
- **To download** these resources

The Central Repository provides a secure environment in which these resources are managed. Storage at the Central Repository provides the originator of such material the added-security of having an off-site duplicate of the resource.

**How it works**

**Uploading resources**

- **Conventional upload via FTP/HTTP**
  - Any type of resource can be uploaded using FTP/HTTP!
  - The registration web page (HTTP) allows the upload of the resource file and the recording of the registration data.
  - For very big files or when slow connections are used, as an intermediate step, the files can be uploaded to an FTP-site before they are registered.
  - Resources that have been converted to XML (e.g. data templates) have a well defined structure and structured queries could be performed on them.

- **Distributed upload**
  - The Automated Datacapture maintains a list of data providers, which type resources they provide and where they are located.
  - Requests for data are sent to distributed providers on the Internet (BioCASE installments).
  - Response files are processed so that semi- or fully automatic registration in the Repository can occur.
  - The response files are in a well defined and structured format (XML). This allows for structured queries on the data contained.

**Current status**

- **Conventional upload**
  - FTP site for intermediate storage of resources is available
  - Parser to convert structured Excel/Text files to XML is under development

- **Distributed upload**
  - BioCASE package and passport schema have been distributed to GCP partners for installation
  - The Automated Datacapture feature has been developed

- **Central Repository application**
  - Fully functional version (i.e. upload, registration, query and download) of Central Repository has been developed

**October-December**

- **Conventional upload**
  - Parser to convert Excel/Text files to XML developed
  - Solicit/assist GCP partners to upload and register resources

- **Distributed upload**
  - BioCASE installations for harvesting of data operational

- **Central Repository application**
  - Validation rules for registration established
  - Central Repository activated and available on-line