Metadata - a Core Component of the GBIF Network

The Global Biodiversity Information Facility (GBIF) is designing its informatics infrastructure as a scalable, distributed architecture that adheres to international standards for data exchange formats and protocols thereby enabling the maximum degree of interoperability across heterogeneous, distributed data holdings and applications. In large distributed networks featuring numerous providers and consumers of data, the key activities of inventory, discovery and access must be well coordinated through provision of registries and metadata catalogues, and through generation of specific indexes. Metadata is thus a central component in an expanded GBIF network that would offer many types of web services for delivering and consuming data. The Initial Operating Capability (IOC) of the GBIF metadata infrastructure (Fig.1), planned for 2010, will feature a centralised, indexed cache of harvested metadata documents derived, through reciprocal sharing agreements, from i) the GBIF data cache of primary occurrence records, ii) GBIF Integrated Publishing Toolkit instances, iii) GBIF Participants’ metadata catalogues, iv) other participating networks’ catalogues. The GBIF catalogue, in turn, will link to clearing house portals such as the EuroGEOSS GeoPortal which will demonstrate multidisciplinary interoperability across the domains of biodiversity, forestry and drought, and contribute to the establishment of the Group on Earth Observations Biodiversity Observation Network (GEO BON)\(^2\). Beyond the IOC phase, there are plans, based on the recommendations of the GBIF Metadata Implementation Framework Task Group (MIFTG)\(^3\), to build a distributed system of regional metadata catalogues that support replication and harvesting.

1  www.eurogeoss.eu
2  www.earthobservations.org/geobon.shtml
3  Report of the GBIF Metadata Implementation Framework Task Group (MIFTG), Jones, M. et al. September 15, 2009 (draft available from eotuama@gbif.org)