Global Taxonomy Initiatives and Taxonomic Authorities Discussion Group

Background

Species names represent the fundamental index for accessing information about biological organisms. For this reason, taxonomy and taxonomic lists are essential to biodiversity informatics efforts. Taxonomy is the science of discovering, describing, and classifying species or groups of species. The term systematics refers to not only the description and classification of organisms, but also the study of relationships among them. Taxonomy is an evolving science, and modern taxonomists are taking advantage of recent advances in technology, computer science, and molecular biology.

A growing concern about the lack of taxonomic expertise in many parts of the world has spurred the international community to create mechanisms to address the "taxonomic impediment." Chief among them is establishment of a Global Taxonomy Initiative (GTI) by the Convention on Biological Diversity. The GTI's work programme is designed to focus on supplying the taxonomic information needed to support the major work areas of the Convention, and to support capacity-building to ensure the ability of countries to undertake the priority taxonomic work required to implement the Convention. There is a strong emphasis on coordination and capacity building, with four operational objectives: 1) assess taxonomic needs and capacities; 2) help build capacity and infrastructure to obtain and curate specimens that are the basis for taxonomic knowledge; 3) improve access to taxonomic information with an emphasis on data repatriation; 4) integrate taxonomic objectives into the thematic and cross-cutting work programmes of the Convention.

Key Issues for Biodiversity Information Networks

Stability in Nomenclature versus Shifting Knowledge

Taxonomy must balance two sometimes contradictory goals: 1) the need to create stable nomenclatures that enable diverse audiences of nontaxonomists to consistently refer to and communicate about the same organisms; and 2) the need to continually refine and update classifications and nomenclatures to reflect current taxonomic understanding. There have been "lumpers" and "splitters" among taxonomists for as long as humans have classified and named organisms, leading to multiple sets of names for the same group of species. In recent decades, however, the widespread adoption of rigorous approaches for determining the relationships among organisms, or phylogenetic analysis, has compounded the problem of shifting nomenclatures. Such cladistic analyses focus on the reconstruction of family trees by emphasizing relationships based on ancestral rather than derived characteristics. While these approaches have greatly advanced our understanding of the relationships among species and higher taxonomic groups, they have also generated a large amount of proposed changes in nomenclatures. Some advocates of a "Phylocode" have even taken the step of suggesting that the Linnaean system of nomenclature be abandoned in favor of an ever shifting, tree-based approach. Adoption of such a proposal would undoubtedly create major difficulties for the biodiversity informatics community.

Standard Lists versus Multiple Taxonomies

A related issue has to do with the development of standard nomenclatural lists that can serve as the basis for cross-community collaborations versus an approach that emphasizes the maintenance of multiple alternative taxonomic views. In general, practicing taxonomists prefer the flexibility of multiple taxonomies, while practitioners prefer the stability afforded by standard lists. A key informatics need is the development of tools for managing and inter-relating multiple taxonomies in ways that can meet the needs of broader communities.

Infrastructure and Capacity Building

Ongoing management and development of the enormous volume of information involved in taxonomy (approximately 1.75 million species names) will require investments in infrastructure and capacity building. Inventories conducted at local, national and regional scales, or targeted at specific taxonomic groups, can be used as an opportunity to develop local expertise for accurate identification of specimens and to focus data repatriation efforts.

Sample of Existing Initiatives

Catalogue of Life

(http://www.sp2000.org/)
In June 2001, Species 2000 and the integrated Taxonomic Information System (ITIS) formed the Catalogue of Life Consortium, aimed at bringing together all of the naming conventions of the world in one, cross-referenced system. The goal is to provide a uniform and validated quality index of the names of all known species for use as a practical tool. Key outputs will include an electronic baseline species list for use in species inventory projects worldwide, an index to species databases on the Internet, a reference system for comparison between species data sets, and a comprehensive worldwide catalogue for checking the status, classification and naming of species. The current, online databases contain entries that represent perhaps 40% of the total known species. Databases are managed by a federation of organizations that specialize in each of the major groups of organisms.

Taxonomic Name Service – ECAT

(http://www.cria.org.br/eventos/tdbi/tdwg/presentations/Bjorn-TDWG_meeting.ppt)
The Global Biodiversity Information Facility (GBIF) works to make the world's biodiversity information freely and universally available. Technically, GBIF is evolving to be an interoperable network of biodiversity databases and information technology tools using web services and Grid technologies. In the near term, GBIF will provide a global metadata registry of the available biodiversity data with open interfaces. A major GBIF initiative is ECAT (the Electronic Catalog of Life). ECAT is internally linked to the databases that form the Catalogue of Life, and will provide a taxonomic naming service bridge between the Catalogue of Life and GBIF biodiversity information resources.

Taxonomic Database Working Group – TDWG

(http://www.tdwg.org/)
TDWG, an initiative of the International Union of Biological Sciences (IUBS), focuses on the development and promotion of standards to facilitate the exchange of taxonomic data. Active subgroups are focused on biological collections data (ABCD Schema), economic botany, geography, descriptive data (SDD Schema), and spatial data standards. Because of its inclusive and open nature, proposed TDWG standards and schema for structured data retrieval (DiGIR) have been adapted by leading bioinformatics initiatives such as REMIB (CONABIO), Species Analyst (University of Kansas), the Mammal Network Information System (MaNIS), and GBIF.

Guidance for the Discussion Group

Opportunities for IABIN / CHM Joint Work Plan

What would be the most urgent and highest impact activities that could be undertaken by the countries and non-governmental organizations of the Americas to address:

Scientific Cooperation

- Assessment of current taxonomic capacity
- Repatriation of specimen data to countries of origin
- Training of new taxonomic specialists and parataxonomists
- Developing content for taxonomic databases on species groups not yet covered by the Catalogue of Life

Technical Cooperation

- Linking specimen and nomenclatural databases, and making them freely available via the Internet
- Training institutions to implement emerging standards to increase interoperability, especially as developed by TDWG and promoted by GBIF

Institutional Cooperation

- Increased participation of smaller or more regional organizations in international initiatives such as the Catalogue of Life