# Implementing a Metadata Database for an Environmental Information System

Andréia da Silva Fagundes   Claudia Bauzer Medeiros
andreiaf@cpqd.com.br   cmbm@dcc.unicamp.br
IC - UNICAMP - CP 6176
13081-970 Campinas SP Brasil

**Abstract**

*This paper presents the modeling and implementation aspects of a metadata database for the information system of the BIOTA/FAPESP research program. This program's goal is the cooperation among biodiversity researchers in the State of Sao Paulo, Brazil, thus helping to maintain and create environmental protection programs within the State. The information system, under development, shall integrate data from the different research groups and foster the dissemination of their work. This information system is unique in several aspects, including the diversity of data managed, and the spectrum of users. The metadata database is the system's component responsible for the high level description of the various biodiversity data gathered by the researchers. This paper concentrates on the metadata standard developed for this information system, and on the database implementation aspects.*

## 1 Introduction

Environmental information systems are concerned with the environment and natural resources to help manage data about soil, water, air and the various plant and animal species [16]. According to [16], the management of this type of data is considered very complex since (1) the volume of data that need to be processed in a query is normally large; (2) data to be managed are heterogeneous; (3) user needs and work profiles vary widely; and (4) the operations to be performed depend on the kind of living organism under study, its geographic location and the goals of each application. Besides being heterogeneous in nature (encompassing flora and fauna and the geophysical description of their habitats), these data are also heterogeneous as to storage media, and characterized by total absence of standardization.

One of the means to help manage data with such heterogeneity and distribution characteristics is the use of *metadata*. Metadata can be defined as "data about data", describing the contents of a data set, their measurement units, quality and purpose [14]. They help standardize the description, the processing and the integration of heterogeneous data, furthermore speeding up their access and updates.

There are several country and world-wide initiatives for the development of environmental information systems. In Brazil, a pioneer effort is the Biodiversity Information System for the State of São Paulo, financed by FAPESP (the State's research agency). This system, from now on referred to as BIOTA-IS (for BIOTA Information System), aims to integrate data which allows mapping the State's biodiversity, supporting the integration of distinct research efforts and providing input to environmental protection programs (both in terms of monitoring a given area and establishing new programs). The data being integrated consist of field observation records, natural history museum collections, and laboratory research measurements. These data are produced by distinct groups of scientists – biologists, ecologists, geographers and others. The system's end-users vary from an expert biologist to school children to policy makers, and therefore the system must provide different types of interfaces and functions to cater to such a varied public.

Similar to other multipurpose environmental information systems (e.g., the Australian ERIN [2]), BIOTA-IS is not limited to species description but encompasses their spatial distribution. Data are therefore georeferenced (i.e., associated to their location on Earth) and are managed by integrating a relational DBMS and a GIS (Geographic Information System). BIOTA-IS nevertheless differs from other environmental information systems for two reasons: it does not limit the set of species to be considered, and it is geared towards a broad user spectrum, ranging from biologists to laypeople. All other systems are either specialized in a specific set of organisms, or restrict their use to biologists, or do not allow correlating distinct species, which are maintained in isolated non-integrated databases.

In the design of BIOTA-IS, it was decided to create a metadata database to speed up access and integration of the different data types concerned. However, there is no metadata standard for the kind of application envisaged, especially if the broad spectrum of user profiles and species diversity is considered. This paper presents the solution developed for this problem, which consists of a new metadata standard, that can be used not only by BIOTA-IS, but also by other systems with similar characteristics. This standard was used to implement the BIOTA-IS metadata database. A first prototype was developed in the LIS laboratory at UNICAMP. The information system proper is under development at Fundação André Tosello, and uses the same standard.

The rest of this paper is organized as follows. Section 2 presents a theoretical foundation for environmental applications and their use of metadata. Section 3 presents the main characteristics of BIOTA-IS. Sections 4 and 5 describe the metadata standard and some implementation features. Finally, section 6 presents conclusions and ongoing work.

## 2 Environmental applications and their metadata

The growing concern with environmental conservation is motivating the amassing of a large amount of data about the environment and development of tools to manage these data. These tools have the generic name of *environmental information systems*, and vary widely in scope, data types and goals [16, 10].

Many of these systems are available on the Web. Some are limited to providing, via metadata, access to organizations and their documents – e.g., [3, 5]. Others support complex data manipulation functions – [2, 17]. In spite of the differences among these systems, all share the property of using metadata files or databases to speed up data access. Metadata for these systems usually document content, context, structure and access modes, and a limited amount of information on data quality [7, 9]

Many initiatives have taken advantage of facilities provided by metadata to describe standards which are appropriate to particular domains. The best known standard is Dublin Core [20], widely used in digital libraries. Environmental applications usually adopt FGDC [6], which is geared towards maps and spatial data, and the standard of EOSDIS [12] (for applications involving global changes).

Examples of environmental systems that adopt or propose metadata standards are CDS - *Catalogue of Data Source* [3] (the CDS standard), *Earth Pages* [13] (the EOSDIS standard), EOSDIS - *Earth Observing System Data and Information System* [12] (the EOSDIS standard), GCMD - *Global Change Master Directory* [11] (the DIF standard), *Geospatial Data Clearinghouse* [6] (the FGDC standard), GILS - *Government Information Locator Service* [19] (the GILS standard), NOAA - *National Oceanic and Atmospheric Administration* [18] (modified FGDC standard), and UDK - *Environmental Data Catalogue* [5] (the UDK standard).

Environmental systems that contemplate management of biodiversity data are relatively scarce and generally geared towards expert end users (mostly biologists and researchers on environmental studies). These systems differ from the ones mentioned in the previous paragraph because, besides environmental data, they also publish data on flora and/or fauna. The oldest and best known is ERIN [2], which basically adopts FGDC's standard. Though it manages a large data collection (roughly 180Tbytes), these data are fragmented across independent databases and systems, and cannot be correlated.

Existing environmental metadata standards are not sufficient for biodiversity studies, because they ignore species' characterization and their habitats. For this reason, it was deemed necessary to develop a new standard for BIOTA-IS, which adopts (and adapts) relevant sections of existing standards, in special FGDC. The sections that follow describe this new standard and the database built to manage the associated data.

# 3 The BIOTA/FAPESP information system

BIOTA/FAPESP [1] is a multi-disciplinary program geared towards supporting research in biodiversity preservation for the State of São Paulo, Brazil. Encompassing several research fields and projects in the broad area of bio-systems, it was officially launched in 1999 by FAPESP (the State's research funding agency) with the goal of enhancing and disseminating knowledge about the State's biodiversity. All projects within this program are dedicated to studying distinct biodiversity factors, developing methodologies and referentials for studies of environmental impact. This will allow estimating biodiversity loss in distinct spatial and temporal scales and identifying areas where conservation must become a priority. The BIOTA-IS information system is being developed to manage the data collected by all these projects.

## 3.1 General architecture

From a macro perspective, one can say that the BIOTA-IS manages two kinds of data, both of which have associated metadata: *species observation records* and *habitat* data; and *digital maps.* Observation records are linked to the maps, thereby allowing associating organisms and habitats via the location where the observation took place, for a given time frame – i.e., georeferencing the observations. Besides being associated with the place and time frame in which organisms were observed, an observation record ideally contains the research method employed to perform the experiment (e.g., instrumentation).

Speaking from a high level perspective, observation/collection work involves field trips and data gathering and registering and observation procedures which are particular to the organisms (flora or fauna), to the scientists conducting the observation, and to the goal of the experiments being conducted. Nowadays, for trips conducted within BIOTA projects, each field group takes a GPS along to georeference collection data, thus standardizing the association of field observations to their geographic location.

The notions of *observation* and *collection* are central to the information system. The term "collection" is often associated with museum records, whereas "observations" may be annotations made from the observation of species at some location. The main difference is that a collection item is "physical", in the sense that the organism was actually collected from the field and preserved in a curated collection [17]. Collected specimens may be prepared in different ways, and thus there may be several collection records for one single organism. One collection record may also refer to several species, if different organisms were observed at a certain geo-spatial location (e.g., when insects are collected by means of a trap). BIOTA-IS will manage both observation records and collection data. From now on, this paper will refer to these two kinds of records generically as "collection records", or simply "collections".

Collection records can be found under different guises: museum catalogs, scientific papers, scientific expedition reports, spreadsheets and database records. Frequently, the indication of place is imprecise (e.g., "Iguaçu Falls", "Tijuca Forest").

In BIOTA-IS, every collection record is characterized by two sets of data: basic set (fixed format and compulsory) and complementary set (which varies with the observed species). The basic set characterizes a record according to the following items: *who* performed it (author), *where* it was conducted (geo-location) and *when*. The complementary set describes *what* was observed in that space-time frame, by that author (list of species, specified according to a given taxonomy). In BIOTA-IS, thanks to georeferencing, all collection records will be linked to digital maps.

Figure 1 schematically shows the system's architecture. It is composed by three main functional blocks: user interface; data and metadata managers; and data repositories. This paper is centered in the description of the part concerning DBMS-managed data (collections and metadata) – see outlined boxes in the figure. The rest of the system is under construction right now (July 2000).

We point out that this architecture presents no novelty at this level of detail, but has been included in the paper in order to allow explaining how the system works.

### 3.1.1 Data repositories

The repositories store data pertaining to various biodiversity projects (field observations, bibliographic databases, papers, reports), maps and metadata. Repository data are linked to allow georeferencing collections via the maps, thus enabling correlating data on distinct species, as well as performing different types of spatial analyses on geographic characteristics against organisms observed.

Each collection entry corresponds to one "record" (in the physical sense) containing basic data (who, when, where and taxonomic groups) and several complementary records (describing the observed organisms and their habitats).

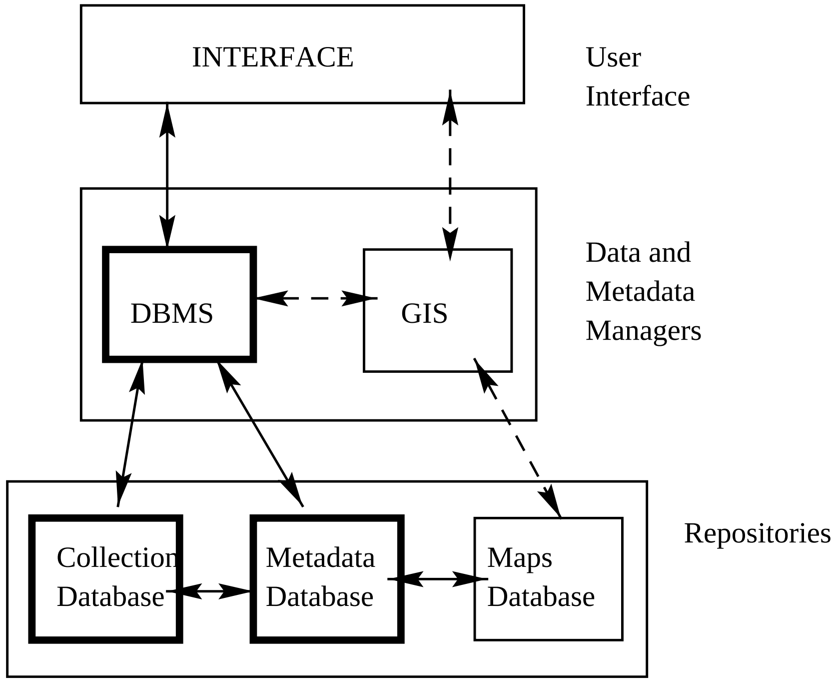


Figure 1: Generic architecture of BIOTA/FAPESP's information system.

#### 3.1.2 Data and metadata management systems

Data and metadata are managed by the DBMS and by the GIS. There are two types of query profile: invoking predefined functions (similar to database views) and browsing. This second type of access allows collections and maps to be browsed directly (via DBMS and GIS). It is expected, however, that most queries will be processed having the Metadata database as a starting point, thus allowing a preliminary step of data filtering. The GIS manages the map repository, while the DBMS manages two repositories: Collection and Metadata, respectively describing collection records and digital maps.

#### 3.1.3 User interface

Basically, the interface module provides two functionalities: (1) data entry, with functions that allow biologists to insert and update collection data online; and (2) query and visualization of stored data, which are processed by the data management systems (DBMS and/or GIS - a Geographic Information System). Queries may be both textual (using predefined forms) and interactive, directly manipulating maps.

### 3.2 The Collection database

This section discusses some of the issues in modeling the data on collections – the so-called *Collection database*. Figure 2 shows the first level model for this database, while section 4 presents details on metadata.

The modeling of the Collection database was achieved after countless meetings with biologists, with analysis of users' profiles and their work procedures. Another factor considered was that of legacy collection records, which must eventually be inserted into the database. Again the problem faced was the diversity of users and data sources, as well as the variety of research methods employed for observing/collecting each species.

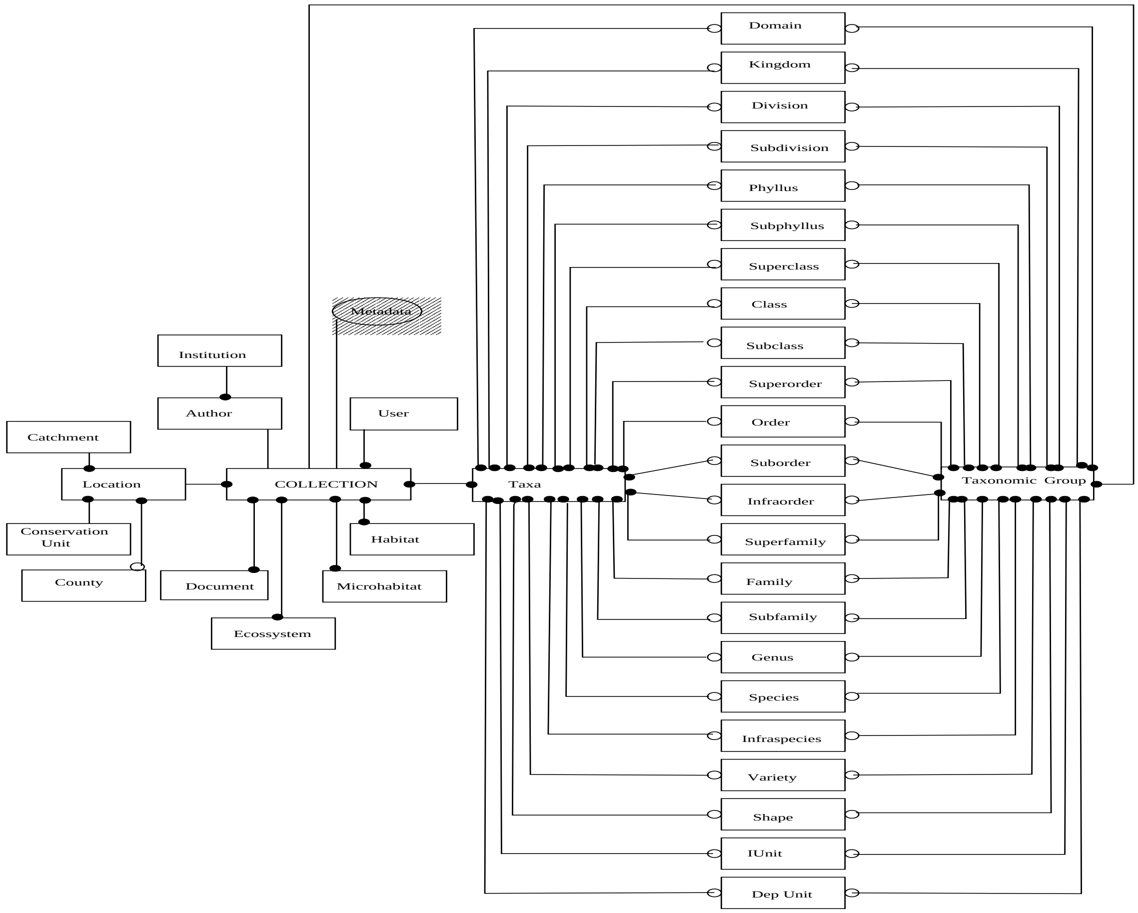


Figure 2: Collection database modeling for BIOTA/FAPESP.

The main entity of the model is the **Collection**, which encompasses the set of observations on (sets of) species for a given region. This entity is inserted by a **User**, and under responsibility of some **Author**, belonging to an **Institution**. The Author is the scientist responsible for the field observations, and provides information on the list of species observed. These species correspond to one or more **Taxa**, and are classified within some **Taxonomic Group**. Observations are associated to a given **Geographic location**. Data gathering is performed according to a certain **Method** which may vary from species to species and with author. A Collection may also be associated with sets of **Documents** (e.g., photos, video). Other important data include the ecological characterization of the environment where the observation was made, such as its period, in which **Ecossystem**, **Habitat** and **Microhabitat**.

Taxonomic classification gives margin to many scientific discussions. In the BIOTA/FAPESP program there are data curators and taxonomists responsible for defining the appropriate names and conventions. Taxonomic classification covers several fields, which range from **Domain** to **Shape**, identifying a species – see rightmost part of the figures. If an organism is physically kept in some natural history collection, it is associated with a **Depositary Unit** (e.g., a museum, and has an identifier **Id-Unit** within this institution. The naming of a species requires providing values to one or more of these fields, depending on their characteristics. For instance, plants are characterized by < Family, Genus, Species > fields, while micro-organisms need definition of < Subdivision, Family, Genus, Species and Infra-species >. A *taxon* is a group that is defined as a result of the process of systematically classifying organisms. Thus, instances of **Taxa** determine valid combinations of values for characterizing a given set of species.

Finally, the **Location** of a collection is described using a set of *geographic coordinates* (points) and sometimes the *areal extension*, a measurement which describes how a given region was traversed starting from the point. The extension attribute is often used in icthyology collection, due to collection methodology (e.g., using a trawler). The Location also includes the **County** and **Catchment**, and may also be associated with a **Conservation unit**, which is a place-name already protected by environmental laws. More details on the data model can be found in [4].

### 3.3 The Map and metadata databases

The Map database, under construction, will contain digital maps for the State. Maps are being produced in the scale 1:50.000, covering around 15 themes relevant to biodiversity characterization (e.g., relief, hydrography, vegetation and soil use). By the end of year 2000, this database will contain around 1000 maps, produced within the same set of standards and audited with a specific methodology, created for the project. Map metadata include information about the map creation process, theme, coordinates, etc.

The records within the Metadata database describe both collections and maps. In the future, metadata will also be associated with sundry documents. The next section describes this database in detail.

## 4 Defining the metadata standard

This section describes the metadata standard developed for BIOTA-IS. The study of standards used in existing environmental information systems (e.g., see section 2) and the requirement analysis of BIOTA-IS showed the need for defining a new metadata standard. It must again be stressed that existing environmental systems are usually geared towards a limited set of species or user profiles, whereas the BIOTA program encompasses in principle all living organisms in the State. As pointed out by [8], although geospatial metadata standards have been developed and endorsed by geoscientists, such standards do not exist for ecological sciences.

In order to define a standard for BIOTA-IS, we first discarded all standards that do not consider characteristics needed in the environmental domain. Due to the species variety and scientific work methodologies to be supported, it was hard to define a basic consensual standard. Moreover, different classification schemes exist for some species and given taxonomic groups may require specific attributes. For instance. the set of attributes needed to identify insects is completely different from the ones used to identify plants, which again differ from algae. Even the way of recording bibliographic references vary with the kingdom or class.

The Metadata databse schema was ultimately derived from expected query profiles. For instance, habitat characterization is one very important issue in biodiversity. Users are frequently interested in recovering data according to habitat description (e.g., collections that contain taxonomic group "Hydrozoa" and that were found in the "coastal" ecossystem). This type of query can be further restricted to geographic location (e.g., restricting the previous query to the "county of Peruibe").

Figure 3 presents the conceptual model of the standard, using the OMT notation [15]. The item **Data** refers to the Collection database, previously described. The figure schematically shows which fields are taken from which existing standard, as well as the fields particular to BIOTA-IS, and which do not fit existing standards. For instance, **Temporal Reference** comes from FGDC, while **Descriptive** borrows from Dublin Core [20], with modifications.

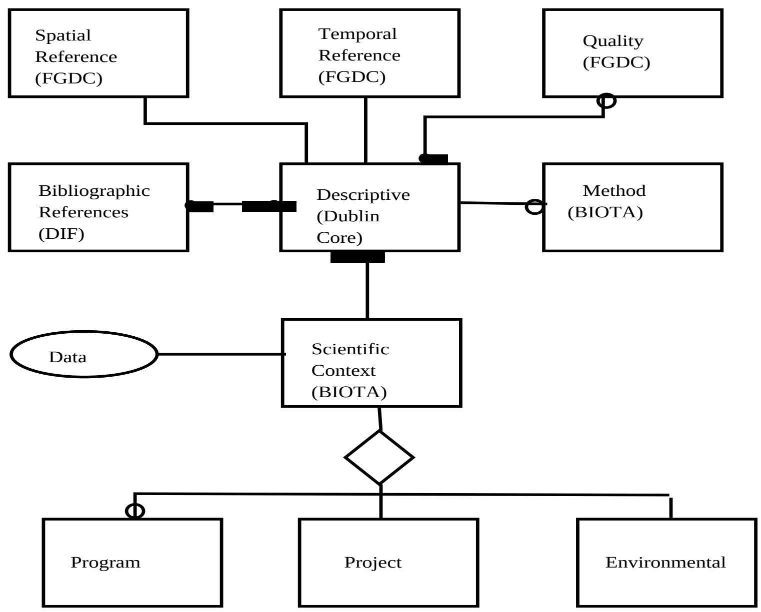


Figure 3: Metadata standard for BIOTA/FAPESP.

**Descriptive metadata.** Based on Dublin Core, concern conventional properties. It covers the following attributes: data set identifier, textual description, keywords, author and address; and pointers to additional data documentation. For maps, this includes furthermore map theme and category.

**Spatial reference metadata.** Describe spatial properties of the data, mainly to characterize geographic location, and borrows from FGDC. It has the following attributes: county where data collected, conservation unit, coordinates (GPS-given) and areal extension.

**Temporal reference metadata.** At present, just period in which collection was made.

**Quality metadata.** Describe aspects that allow evaluation of the quality of the data. Presently only applied to the map database, they are based on the FGDC quality information, and have been defined by the team responsible for the Map database. Fields are datum, scale and positional precision.

**Bibliographic reference metadata.** Based on DIF, contain information about references which describe the collection procedures.

**Methodology metadata.** Based on FGDC, describe the methodology adopted while observing/preparing specimens. Attributes are keywords concerning the methodology (e.g., "insect traps") and its textual description.

**Scientific context metadata.** describe the scientific context within which a collection was performed, and associated environmental information. They are based on a proposal by [8] for ecological projects.

These metadata are further divided into **Program** metadata (for high level description of the program), **Project** metadata (project within program) and **Environmental** metadata. Attributes of Project metadata include title, project keywords and related research areas, goals, coordinator and URL. Environmental metadata constitute one of the main differences between the proposed standard and others, describing the ecological characteristics of the area where a collection was made. Attributes are: taxonomic group, ecossystem, habitat and microhabitat.

The system's requirements demand relatively complex metadata, which must describe environmental and geo-spatial aspects. These characteristics render this standard unique. Table 1 compares this proposal to some of the main metadata standards used by some environmental information systems.

The table also evidentiates the need for establishing this standard, since it incorporates new attributes (especially those which concern the ecological context of a collection) and does not need several other fields which are adopted by most standards (e.g., several of Dublin Core's and FGDC's fields). This difference is due to the fact that BIOTA-IS concerns (potentially all) living beings within their habitats, whereas other environmental information systems deal with either environmental characteristics or with a limited number of species. The table shows, for instance, that FGDC, NOAA and ERIN contemplate spatial data, whereas only the discussion in [8] is concerned with ecossystem metadata. We also point out FGDC's pioneer work in considering quality as part of metadata information.

The main contributions of our standard are: (1) coverage, since it is not restricted to collections but also to associated spatial data and maps, unlike all other standards considered, except for FGDC; (2) description of data generated by a wide spectrum of user profiles, covering various aspects of flora and fauna; (3) support of ecossystem characterization. Some aspects still require more working over, such as quality and access right constraints, which remain as future work, and will be defined based on actual system usage.

## 5 Implementation aspects of the metadata database

Part of the system described in section 3 has been implemented as a prototype in UNICAMP, using INFORMIX [4]. This implementation concentrated on metadata management, and relationships among Metadata, Collections and Map databases. The final system (www.biotasp.org.br/sia) is being implemented at Fundação André Tosello, with ORACLE and the ArcInfo GIS, using the same database model.

User access control was defined according to the following hierarchy:

- Administrator/curator. Privileged access to the basic data fields, and management of the Map database.

- BIOTA researcher. Access to biologists responsible for inserting, validating and updating collection data online via WWW.

- Other users. Query-only access, for all kinds of users.

There are furthermore different access levels within the first two groups, which do not concern this paper. We briefly point out distinctive characteristics of each kind of profile, basically distinguishing curators from other users.

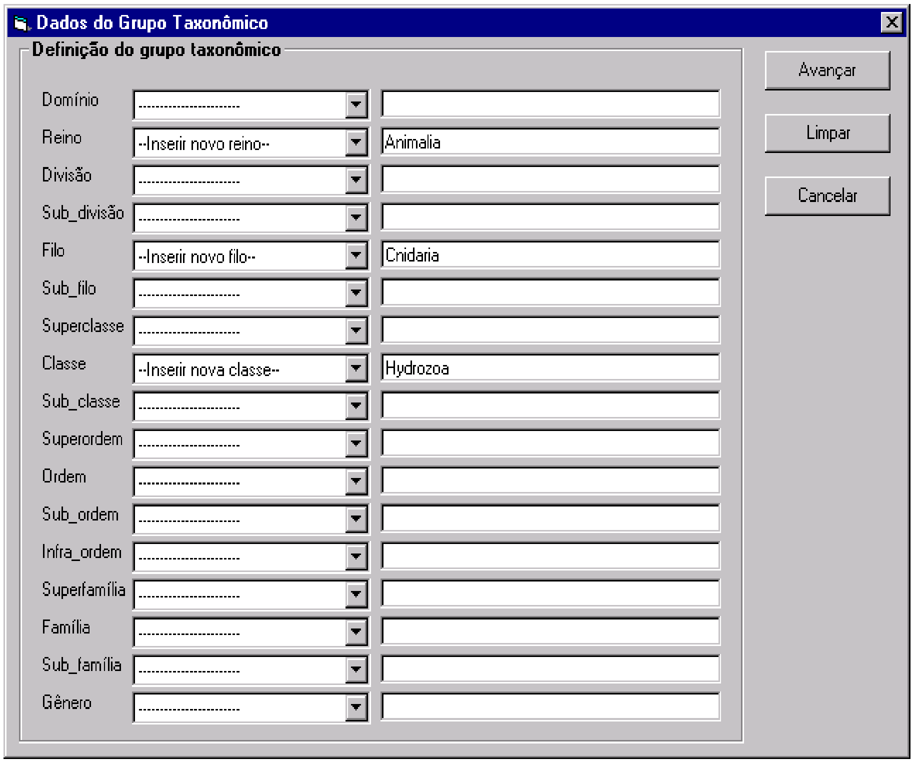


Figure 4: Prototype screen copy - insertion of data for Taxon Hydrozoa.

## 5.1 Administrator/curatorship processing

Data administrators/curators are responsible for maintaining several basic metadata and data fields, which includes defining their (enumerated) domains – e.g., institution, ecossystem terms, scientific names. These attributes establish the context of a collection, and the basis for creating metadata records. For instance, the domain of the "habitat" metadata attribute is defined by the Brazilian Geographical Institute, IBGE. Other examples include county names, and basin/micro-basin characterization. When a new collection record is inserted, these data fields are extracted and inserted into metadata records.

Curatorship of biological data includes the very important issue of defining valid scientific names and taxonomic groups. Figure 4 shows a screen copy of the prototype, in which a new taxonomic group class = "Hydrozoa" is defined. In this group, keyword "Hydrozoa" automatically allows defining a species as belonging to Kingdom = "Animalia" and Phylo = "Cnidaria".

Other administrator tasks include map curatorship, auditing and management. Obviously, there are specific administration tasks which concern the information system itself but which are not within the scope of BIOTA-IS, but are rather standard information system management tasks (e.g., database tuning using the DBMS tools).

### 5.2 User processing

There are two kinds of activities performed by non-administrator users: collection record insertion and update (exclusive to BIOTA researchers) and queries (all users). Most metadata fields are extracted during online insertion via WWW. It must be stressed that online insertion and update are not available in other biodiversity systems, where data are inserted offline. At the moment (July 2000), there are roughly 25 biology research groups within the program, totalling around 300 researchers that can feed the Collection database online. Though this flexibility is good for supporting up-to-date query results, at the same time it brings problems in consistency and quality, which have yet to be evaluated. Online updates are one of the user requirements which complicated system design.

Query processing can be performed according to the following main execution strategies:

- Queries posed directly on the Collection database (via DBMS or Map database (via GIS) without considering metadata.
- Queries posed only on the Metadata database, via DBMS (e.g., to determine how many collections refer to a given taxonomic group).
- Queries on Collection/Map databases that go through an initial filtering step via the Metadata database access.

## 6 Conclusions and ongoing work

This paper described a new metadata standard for BIOTA-IS – the Biodiversity Information System for the State of São Paulo, Brazil, which is being implemented within the BIOTA/FAPESP program. The paper also gave an overview of the implementation of a prototype for the system's Metadata database. In this context, metadata were used with the main goal of helping data description (both collections and digital maps), in order to speed up query processing, as well as enabling users to better grasp query results.

The proposed standard integrates attributes from several metadata standards already in use, incorporating characteristics needed in the management of biodiversity ecological data, and which are not supported in other standards. This need for a new standard comes from the variety of living organisms contemplated by the program, as well as the wide spectrum of user profiles, with diversified work methodologies and data processing needs.

This proposal was validated through a partial implementation of of BIOTA-IS, in a prototype. This prototype is centered on metadata management, but its testing required implementing other databases. For more details on the development effort, and tests conducted, the reader is referred to [4].

There are many extensions to this work, some of which are right now in progress. More specifically, two basic extensions may be conceived for the Metadata database: modifications on the standard, prompted by actual usage monitoring; and implementation geared towards integration to other biodiversity systems.

Eventual metadata modifications may include, for instance, more information on data quality. The first large scale tests were conducted on december 1999, and usage is still too limited to allow drawing any conclusions as to changes.

Integration of BIOTA-IS to other systems is presently being examined (e.g., to the Species Analyst system [17]). The main integration vector will be the Metadata database. Thus, any changes in this database need to consider future usage not only of BIOTA-IS, but of the systems with which it will exchange data.

**Acknowledgements**

The work reported in this paper was partially financed by the BIOTA/FAPESP program, by CNPq, and by the projects PRONEX MCT/FINEP SAI (Advanced Information Systems) and CNPq/NSF (Interoperability in GIS). The authors also thank all members working in BIOTA-IS, as well as all the biologists who so patiently explained their needs, in special professor Carlos Joly, coordinator of BIOTA/FAPESP, for the countless meetings and discussions.

# References

[1] Biota. BIOTA – O Instituto Virtual da Biodiversidade. Web site http://www.biotasp.org.br, 04 2000.

[2] ERIN. Environment Australia Online. Web site http://www.environment.gov.au, 2000. Access on 04/2000.

[3] European Environmental Agency. *CDS - Catalogue of Data Source).* http://www.mu.-niedersachsen.de/system/cds (Access on 09/1999).

[4] A. S. Fagundes. Design and Implementation of a Metadata Database for the Biodiversity Information of São Paulo State. Master's thesis, IC-UNICAMP, dec 1999. In portuguese.

[5] Federal Environmental Agency – Austria. *UDK - Environmental Data Catalogue).* http://udk.bmu.gv.at (Access on 09/1999), year=1999.

[6] Federal Geographic Data Committee. *Geospatial Data Clearinghouse.* http://130.11.52.178/-FGDCgateway.html (Access on 09/1999).

[7] O. Gunther. *Environmental Information Systems.* Springer Verlag, 1998.

[8] W. Michener, J. Brunt, J. Helly, T.Kirchner, and S. Stafford. Nongeospatial Metadata for the Ecological Sciences. *Ecological Applications*, 7(1):330–342, 1997.

[9] W. K. Michener. *Ecological Metadata*, 1999. http://www.lternet.edu/ecoinformatics/-guide/michener.fv2.htm (Access on 10/1999).

[10] C. Murray and C. Daniel. *Environmental Information Systems*, 1999. http://essa.com/-services/eis/index.html (Access on 10/1999.

[11] National Aeronautics & Space Administration. *Directory Interchange Format (DIF) - Writer's Guide, Version 7*, 1999. http://gcmd.gsfc.nasa.gov/difguide (Access on 06/1999).

[12] National Aeronautics & Space Administration. *EOSDIS - Earth Observing System Data and Information)*, 1999. http://spsosun.gsfc.nasa.gov/NewEOSDIS.html (Access on 09/1999).

[13] National Aeronautics and Space Administration. *Earth Pages*, 1999. http://starsky.hitc.com/-earth/earth.html (Access on 09/1999).

[14] M. H. Pelkiki and G. J. Arthaud. *Metadata Standards for Forestry and Natural Resources*, 1999.

[15] J. Rumbaugh and et. al. *ObjectOriented Modeling and Design.* Prentice Hall, Inc, 1991.

[16] A. Sheth and W. Klas. *Multimedia Data Managment – Using metadata to integrate and apply digital media*, chapter Metadata in Geographic Environmental Data Management, pages 57–87. McGraw–Hill, 1998.

[17] System Analyst. System Analyst. Web site http://habanero.nhm.ukans.edu, 2000. Access on 06/2000.

[18] U. S. Department of Commerce – National Oceanic and Atmospheric Administration. *NOAA - National Oceanic and Atmospheric Administration)*, 1999. http://www.esdim.noaa.gov/NOAA–Catalog (Access on 09/1999).

[19] United States Environmental Protection Agency. *GILS - Government Information Locator Service)*, 1999. http://www.epa.gov/gils (Access on 09/1999).

[20] S. Weibel, J. Godby, E. Miller, and R. Daniel. OCLC/NCSA Metadata Workshop Report. Web site http://www.oclc.org:5046/oclc/research/conferences/metadata/ dublin_core_report.html, 1995.

| Field | BIOTA | CDS | GILS | UDK | FGDC | DIF | ERIN | EOSDIS | NOAA | [8] |
|---|---|---|---|---|---|---|---|---|---|---|
| Identifier | X | | X | X | X | X | | X | X | X |
| Description | X | X | X | X | X | X | X | X | X | |
| Keywords | X | | X | | X | X | X | | | X |
| Theme | X | | | | | | | | | |
| Category | X | | | | | | | | | |
| Author | X | X | X | | X | X | | | X | X |
| Address | X | X | | | X | X | X | X | X | X |
| County | X | | | | | X | | | X | |
| Conservation unit | X | | | | | X | | | | |
| Latitude | X | | X | | X | X | X | X | X | |
| Longitude | X | | X | | X | X | X | X | X | |
| Extension | X | | | | | | | | | |
| Period | X | | X | | X | X | X | X | X | |
| Precision | X | | | | X | X | X | | | |
| Datum | X | | | | | | | | | |
| Scale | X | | | | | | | | | |
| Reference | X | | | | | X | | | X | |
| Method | X | X | X | | | X | | X | X | X |
| Description | X | | | | | | | | | |
| Program | X | | X | X | | | | | X | |
| Program URL | X | | | | | | | | | |
| Project | X | X | | | | X | | | X | X |
| Proj. Keywords | X | | | | | | | | | |
| Proj. Goals | X | | | | | | | | | X |
| Coordinator | X | X | | | | | | | | X |
| Proj. URL | X | X | | | | X | | | | |
| Vulgar name | X | | X | | | | | | | |
| Taxonomic group | X | | | | | | | | | |
| Ecossystem | X | | | | | | | | | X |
| Habitat | X | | | | | | | | | X |
| Microhabitat | X | | | | | | | | | X |
| Associated documents | X | | | | | | | X | | X |
| Title | | X | X | | | X | X | | X | X |
| Modification date | | | X | X | X | X | X | | | |
| State | | | | | X | X | X | | | |
| Access constraints | | | X | | X | X | X | X | | |
| Use constraints | | | X | | X | X | | | | X |
| Consistency | | | | | X | | X | | | X |
| Completeness | | | | | X | | X | | | X |
| Lineage | | | | | X | | X | | | |
| Topic | | | | | | X | | | | |

Table 1: Comparison among some environmental metadata standards.