The Computer catalogue of the Kunstkammer museum collections and perspectives of an Internet-shared anthropological database

Yuri Osintsov
Peter the Great Museum of Anthropology and Ethnography (Kunstkammer)
3 University Emb.
St.-Petersburg, 199034, Russia
Phone: +7-812-3280025 - Fax: +7-812-3280811 - E-mail: Y.V.Osintsov@mae.nw.ru

Yuri Chistov
Peter the Great Museum of Anthropology and Ethnography (Kunstkammer)
3 University Emb.
St.-Petersburg, 199034, Russia
Phone: +7-812-3232719 - Fax: +7-812-3280811 - E-mail Y.Chistov@pobox.spbu.ru

Dmitriy Gerasimov
Peter the Great Museum of Anthropology and Ethnography (Kunstkammer)
3 University Emb.
St.-Petersburg, 199034, Russia
Phone: +7-812-3280025 - Fax: +7-812-3280811 - E-mail dger@dg6388.spb.edu

Abstract: The Computer catalogue of the Kunstkammer Museum is part of a large museum database that has been in development in the Museum of Anthropology and Ethnography (Kunstkammer) since 1996. The two major purposes of the project are: 1) organizing of museum storage routine and museum documentation processing on an up-to-date level and 2) providing free access via Internet to scientific and inquiry information.

We provide Internet users with information about collections in cultural and physical anthropology stored in the Museum. Attributive information for the materials, such as archaeological sites and cultures, historical periods, dating, geographical locality is also available.

The web part of the computer catalogue is based on dynamic web pages. Multi-aims and multi-conditions search through the database is available. The purposes of creating a museum computer catalogue, project achievements and perspectives of its development, as well as system architecture and data layout are special topics for discussion in this paper.

At present information on approximately 400 archaeological, 571 anthropological, and 200 ethnographical collections and subcollections is available for database users. The Russian version of the Kunstkammer database is accessible in Internet on http://www.mae.nw.ru/maedb or from the Museum WEB-site: http://www.kunstkamera.ru.

Key words: Museum computer catalogue, data base, internet, anthropology, archaeology.

Introduction

Collections kept in the Peter the Great Museum of Anthropology and Ethnography (Kunstkammer) present ethnological, archaeological and physical anthropological materials from all over the world. Museum storage consists of about 6,5 thousand collections that contain about 700 thousand pieces. These collections are important for researches in cultural and physical anthropology and are also very attractive for amateurs and everyone interested in human origin and culture.

The developing of a computer technology and Global Computer Network allows the organisation of information about the Museum's collections on an up-to-date level and to give access to this information to anyone interested in it via Internet. Thus, researchers in all branches of anthropology gain an opportunity to check if there are any materials for their research, to get some information about them to decide if it is necessary to work with these collections personally and to get some reference information or even some characteristic of the material that they can use in scientific work. The Russian version of the Kunstkammer Database is now available at http://www.mae.nw.ru/maedb. This Internet source is also very useful for students in anthropology and anyone interested in the origin and history of mankind.

Purposes

There are three purposes of creating the Database of Kunst-

kammer collections – research tool, storage system and educational database of cultural and physical anthropology.

Science

As a research tool the Database provides opportunity to search and select materials for certain research purposes and therefore it should consist of information important for researchers, including description of specific features for each kind of materials, such as technological and typological features of flint objects and ceramic, measurements of skulls, etc. It has already been mentioned that Kunstkammer collections present materials in three branches of anthropology such as archaeology, ethnology and physical anthropology, and the Database contains specific descriptions of any kind of collection.

It will be also possible to make use statistic, such as measuring of technological indexes of prehistoric flint industries or to use the data for multivariate statistical analyses in physical anthropology.

Storage

Since the Kunstkammer museum was founded in 1714, the organization of its storage system is quite complicated – materials from different places can be kept in the same collection or different collections can consist of materials from the same place (e.g. from the same archaeological site). Thus, it was necessary to reflect this complicated structure in the Database.

Education

As an educational database in anthropology it will consists of a large number of articles about archaeological periods, cultures, different ethnicities, etc., presented by collections in the Database.

Data layout

During the development of the project there were created separate schemas for description of archaeological, ethnological and physical anthropological collections. Then the schemas were merged and expanded, and the Museum storage system was integrated in the Database. As a result we tried to get the universal data layout for description of materials in all branches of anthropology (fig. 1).

Each record in the Database has mandatory fields: Identifier, Subordination identifier, Name and Additional info. These attributes uniquely identify any object in the database and allow to build a hierarchical structure tree. Indeed, a hierarchical tree reflects the structure of any typology anyone uses, i.e. a hierarchical tree is the core element of any scientific systematization. Such structure allows us to create procedures, libraries and client applications that can operate with any type of objects stored in the database.

All other record fields are optional in relation to data access, but serve to describe a certain object's peculiarity.

The whole scheme can be divided into the following parts or schemes: common information (common), documents (docs), information about any persons mentioned in the Database (persons), information about any organizations or organizational units (orgs), museum receipts information (receipts), museum storage system (storage), archaeological sites (sites), collections' and objects' data (sets).

Common scheme

Table common.chronology
There are two schemes currently presented in the table: relative and absolute chronology. Archaeological periods in the relative chronology scheme have no beginning and end dates because it is impossible to find universal dating for them. Absolute dating scheme consists of dating of events such as excavation dates or the existence period of an archaeological site. In the latter case period limits are the earliest and the latest possible dates of a site's existence.

Mandatory fields: Name – name of period or event; Info – description of a period or background for event dating.

Optional fields: Date_start; Date_end

Table common.locality
This table accumulates information about spatial location of object, such as an archaeological site.

Mandatory fields: Name – name of unit or site; Info – description of the characterized area

Optional fields: Northlimit – northern latitude; Eastlimit – eastern longitude; Southlimit – southern latitude; Westlimit – western latitude.

There are three schemes in this table – geographical regions, administrative units and absolute coordinates.

Table common.cultures
This table is factual about archaeological cultures or ethnicities.

Mandatory fields: Name – name of archaeological culture or ethnicity; Info – description of archaeological culture or ethnicity.

Table common.materials
Describes materials museum pieces objects are made of.

Mandatory fields: Name – material name; Info – material description

Persons' scheme

Table persons.list
This table consists information about all persons referred to in the Database.

Mandatory fields: Name – full name; Info – additional info about the person.

Optional fields: Surname – family name; Fstname – first name; Sndname – middle name; Dates_id – dates of life, reference to the table Common.chronology; Address – mail address; Phone – phone number; Fax – fax number; Email – e-mail address; Url – uniform resource locator, e.g. internet personal pages.

Orgs scheme

Table orgs.types
Types of organizations or units listed in the table.

Mandatory fields: Name – organization type; Info – type description.

Table orgs.profiles
Organization field of activity kept in the table.

Mandatory fields: Name – organization profile; Info – profile description.

Table orgs.tree
All organizations related to collections are kept in this table.

Mandatory fields: Name – name of the organization; Info – additional information about the organization.

Optional fields: Abbr_name – abbreviation of the organization's name; Manager_id – the manager of the organization's unit, reference to the table Persons.list; Address – address of the organization; Phone – phone numbers; Fax – fax numbers; Email – e-mail addresses; Url – uniform resource locator, e.g. internet site of the organization.

Docs scheme

Table docs.types
This table serves to control over search for any type of documents.

Mandatory fields: Name – document type; Info – description of the document type.

Optional fields: Viewer – reference to the viewer to see a document if there is a digital version; Editor – reference to the programme to edit a document if there is a digital version.

Table docs.tree
Information about all documents related to any Database objects, such as stored documents or bibliography, etc. is kept in this table.

Mandatory fields: Name – the name of the document, Info – the description of the document

Optional fields: Type_id – type of the document, such as book, manuscript, photography, audio-record etc., references to the table Docs.types; Author_id – the author of the document, references to the table persons.list; Publisher_id – publishers (if any), references to the table Orgs.tree; Date – date of publishing; Abstract – abstract of the document with hyperlinks to related documents; Path – place of the document's storage; Url – hyperlink to the Internet-version of the document.

Receipts scheme

Table receipts.types
In this table such types of collection acquiring as buying, presenting, expedition, etc. are described.

Mandatory fields: Name – receipt type; Info – definition of the receipt type.

Table Receipts.list.
Any set of objects acquired by the Museum at the same date, from the same person or organization receives a unique number in the Main Museum Incoming Book. This set could be gathered by different persons, e.g. the same incoming set could originate from different archaeological sites, but we mean that this set could be collected by some person or organization. As we have said, there are several ways of collections' acquisition (receipt types) – a collection could be presented, sold, replaced from a different organization, obtained by the Museum's expeditions, etc.

Mandatory fields: Name – the number in the Main Museum Incoming Book; Info – any information on a collection's acquisition.

Optional fields: Date – the acquisition date; Type_id – receipt type, reference to the table Reciepts.types; Org_id – the organization that materials came from, reference to the table Orgs.tree; Collector_id – the person who gathered these materials, a reference to the table Persons.list; Supplier_id – the person who delivered a set of objects to the Museum or who is responsible for moving materials from another organization, a reference to the table Persons.list.

Storage scheme

Table storage.types
This table contains information about types of places of the collection's storage. The hierarchical tree allows describing different types of storage as an exposition or museum fund.

Mandatory fields: Name – type of storage; Info – definition of the type.

Table storage.tree
This table enumerates museum storage system up to the shelf or box level.

Mandatory fields: Name – name of the storage place; Info – additional information.

Sites scheme

Table sites.types
This table contains information about types of sites. It is especially important in archaeological sites' description (e.g. tumulus, crypt, tomb, cave burial, etc.)

Mandatory fields: Name – type of a site; Info – definition of the type.

Table sites.tree
A site is a very important Database object. It's not simply an archaeological site. We use this term to describe common origin of a set of objects. It is possible to describe site as a dot in a three-dimension space where the first axis is location, the second – time, and the third – cultural or ethnical identity. Archaeological site is the clearest example of this conception – materials from a certain archaeological site reflect certain cultural tradition in a certain period of time from a certain place. In the case of a multi-layer site each occupation layer should be described as a separate site. Hierarchical structure also allows to describe the whole site in a separate record.

The conception of a site is also very useful when describing an ethnographical collection. In this case certain sites reflect materials on a certain ethnical or social group that existed or exists in a certain time in a certain area.

Spatial and temporal location of a site could be described in different schemas – geographical regions, administrative units, and geographical coordinates, absolute and relative chronology. For some objects' sets collected several ages ago our knowledge is limited. But if we know something about their origin we should create a new site record. We call it a "ghost-site". For example, in a collection there are objects from East Siberia – that is all we know about them. But we make a site record – with the Geographical region attribute – East Siberia.

Mandatory fields: Name – site name; Info – short description of a site with a hyper-reference to a hyper-text document with description of the site and some images, such as photographs, maps, cross-sections, etc. These documents contain information of educational significance.

Optional fields: Period_id – the period of a site's existence in the relative chronology schema, a reference to the table Common.chronology, scheme relative chronology; Absdate_id – the period of a site's existence in absolute dates from the earliest possible to the latest possible dates and the background of absolute dating, reference to the table Common.chronology, scheme absolute chronology; Geo_id – geographical region where the site is located, a reference to the table Common.location, geographical scheme; State_id – the administrative unit in which the site is located, a reference to the table Common.location, administrative scheme; Coord_id – geographical coordinates of a site and a short text description of the site's location, a reference to the table Common.location, scheme geographical coordinates.

Sets scheme

Table sets.types
This table keeps information about types of sets, e.g. collection, subcollection, composed object, etc.

Mandatory fields: Name – types of set; Info – definition of the type.

Table sets.profiles
This table contains information about profile of sets, including attribute definition for different types.

Mandatory fields: Name – profile of sets; Info – definition of the profile.

Table sets.tree
The basic object in the database is collection – it's a set of objects kept in the museum with a certain museum registry number. As a "perfect" collection we imply a set of subjects from a certain site, which was obtained in a certain period of time by a certain person (e.g. materials of archaeological or ethnological expedition from one field season). As a rule materials with the same incoming number form the same collection. But sometimes materials with the same incoming number could be separated in several collections (e.g. physical anthropological and archaeological materials from the same archaeological site are kept in different funds).

At the same time, if materials from different archaeological sites come into the Museum at one time and receive the same incoming number, they could be stored under the same collection number. But the hierarchical tree allows making subcollection records – each one for different site. In such case link_id identifier of subcollection record refers to collection record. Collections from the same site or from a certain area of ethnological research also could be divided according to materials the items are made of. It also makes more sense as it is preferable to keep objects of different materials in different conditions, that means in different storage places.

Following the Database logic, each collection item with its own item number is described as a low-level subcollection, and sometimes there could be complex objects (e.g. composite tool) and it is necessary to describe each part of such object in the database as lower-level subcollection. It should also be mentioned that sometimes several subjects (e.g. several flakes or potshreds or different bones of one skeleton) could be registered under the same item number. Identifying each collection with certain organization or organization department allows to describe in the Database a collection not only from the Kunstkammer museum, but also from any museum or institution that possesses anthropological collections.

Mandatory fields: Name – collection number or any name of subcollection (things); Info – information about storage or any collection movement as participation in exhibitions or laboratory studies etc.

Optional fields: Receipt_id – incoming number of materials in collection, reference: to the table Receipts.list; Dep_id – Museum or organization where collection is stored or research department, which is responsible for the collection, reference to the table Orgs.tree; Date – the date of a collection's registration. Recorder_id – the person who registered the collection, reference to the table Persons.list; Keeper_id – the person who is responsible for the collection, reference to the table Persons.list; Subj_count – number of objects in the collection;

Item_count – number of items in the collection; Material_id – material that the object is made of, reference to the table Common.materials; Site_id – archaeological site or certain area where ethnographical or physical anthropology collection was obtained from, reference to the table Common.sites; Extractor_id – the person who obtained the collection (e.g. the chief of an expedition), reference to the table Persons.list; Extract_date – period when the collection was obtained, reference to the table Common.chronology; Type_id – type of the set; Profile_id – type of objects according to typological scheme or collection profile, reference to the table Sets.profile; Storage_id – place of collection or subcollection storage such as exhibition showcase or shelf, reference to the table Storage.tree.

Table sets.docs

One document can be related to many Database objects or one object can be related to many documents. This structure is reflected in the table Docs.tree.

Fields: Doc_id – Document name, reference to the reference table Docs.tree; Set_id – the object that a document is related to.

Table sets.attributes

Describes objects in collections according to typological schemes, given in Sets.profiles. A certain type of objects needs a certain description template. In the same collection objects from the same site could be of different profiles – for example, ceramic and flint in archaeological collections. In the Database we can support description of each type in its own scheme.

Mandatory fields: Name – name of the attribute; Info – additional information; Profile_id – reference to the attribute definition; Set_id – reference to the object in the table Sets.tree.

System architecture

For technical support of all of the declared purposes we chose a client-server system architecture. The server side is divided into two subsystems – the databank and the set of stored procedures for access. The client side consists of access libraries, granting high level interface to the server stored procedures, and the client applications themselves.

In order to keep strict control over access to the data, enforce data integrity and data access speedup direct manipulations in the databank were denied – only operations with data going through the stored procedures are allowed.

Client access library presents two levels. The low one is the extension of the ADO library. It consists of the following objects: "Connection" – database connection and interface for stored procedures; "Node", which represents the node of the data tree, and "Nodes", which is collection of "Node" objects; "Field", which represents attribute of any record with attribute description, and "Fields" – collection of "Field" objects, that is classic database record.

ActiveX controls make a higher level with objects "Tree" and "Properties". The first is designed to navigate through the hierarchical tree of records and the second – to display and operate above the records' attributes.

Thus, the system architecture can be presented as a stack: ActiveX controls – ADO Extensions – Stored Procedures – Databank.

Using such system architecture with conjunction of the sets of the mandatory fields in any database record automatically gives us the following results: 1) the system is invariant for adding new objects and object attributes; 2) the total search through the database is available because all records have the set of standard attributes; 3) as data access is going through the stack, it is easy to convert inner protocols to standards ones, such as HTTP, LDAP.

Achievements

So far we have reached the following achievements: the server level is completed, a data management tool and a web page generator are created.

The data management tool looks like Windows Explorer and has a similar operating logic. It supports navigation through tree-organized records, records adding and editing.

The page generator is built using ASP technology and allows navigation and search. Three types of search are available: SEARCH BY ABSOLUTE DATES, total SEARCH BY NAME of the objects of any type and MULTIPURPOSE SEARCH, which reflects relations between objects. In advanced or MULTIPURPOSE SEARCH any type of data that has a name in the Name field of record, such as archaeological sites, cultures, periods, geographical, administrative regions, etc., could be the aim of a search. The user can choose the purpose of the search and then any kind and non-limited amount of search conditions. E.g., it is possible to find all archaeological cultures that existed on the territory of the European part of Russia in the Neolithic period, or geographical regions where a certain archaeological culture was spread. The only limit for such searches is that sites or regions to be found should be presented in materials stored in the Kunstkammer and described in the Database. It may well be true for the reason that all low-level database tables are related with each other through collection tables.

The created system is time shared and multi-user. According to the user's access permission it is possible for several users at the same time to read, add or edit data via local network or via Internet.

Some limits are still not overcome: at present all applications exist only in the Russian version, no access restriction is supported, and thus some fields with internal Museums' information are closed for web-users.

At present information about 400 archaeological, 571 anthropological, and 200 ethnographical collections and subcollections is available for database users.

Perspectives

The system described above was designed not only as a catalogue of Kunstkammer collections but also as a number of scientific, research and educational data bases connected with this catalogue. For instance, together with this project there was developed a project of creation of a Russian Craniological Data Bank that encompasses materials from almost all anthropological museums of Russia. Most probably, in the future there will be developed a number of other data bases connected with the catalogue of Kunstkammer collections (such as a bibliography of publications about the Museum's collections, etc.)

At present European and American anthropologists and archaeologists are working on organizing several global projects that cannot be realized without creation of big specialized databases. The developed technique, in our opinion, opens brilliant opportunities for the creation and administrating of such resources in the global computer network.

The architecture of the project also make it possible to create a powerful education and reference resource on anthropology and archaeology of the peoples of the world. The authors of the project did their best to provide the attribution of the Museum's collections with references, including gazetteer, reference books, illustrations, hyperlinks to other projects in the net, etc.

Speaking about the technical maintenance and upgrowth of our museum Internet-based computer catalogue, it should be mentioned that we have the following for the nearest future: 1) developing multi-language support; 2) integrating XML technology in Web application and access to the Museum directory by LDAP protocol.

Integrating an XML technology allows us to provide the created scheme of description of anthropological collections in standard XML-Schema terms. It also makes it possible to use convenient schemes granted by other providers. XML-documents in combination with style tables allow to use personal settings in web layout depending on the user's preferences, access tool and communication bit rate.

The ideology of data operating protocol, which is used in the present database is practically very close to LDAP ideas. In the nearest future we plan to remake and expand our protocol to conform to the LDAP standards. It will allow us to open access to the Kunstkammer Database for outer users.

As a result, we would be able to organize access to data and metadata by standard protocol, which supports references between directory servers and standard methods to operate with these references. In perspective, it will make it possible to organize a Common Museum Network. Using this protocol, any Network member will be able to create a personal client application layout for a certain (scientific or other) need. So, anyone interested in cooperation is welcome. We are looking for partners.

Figures

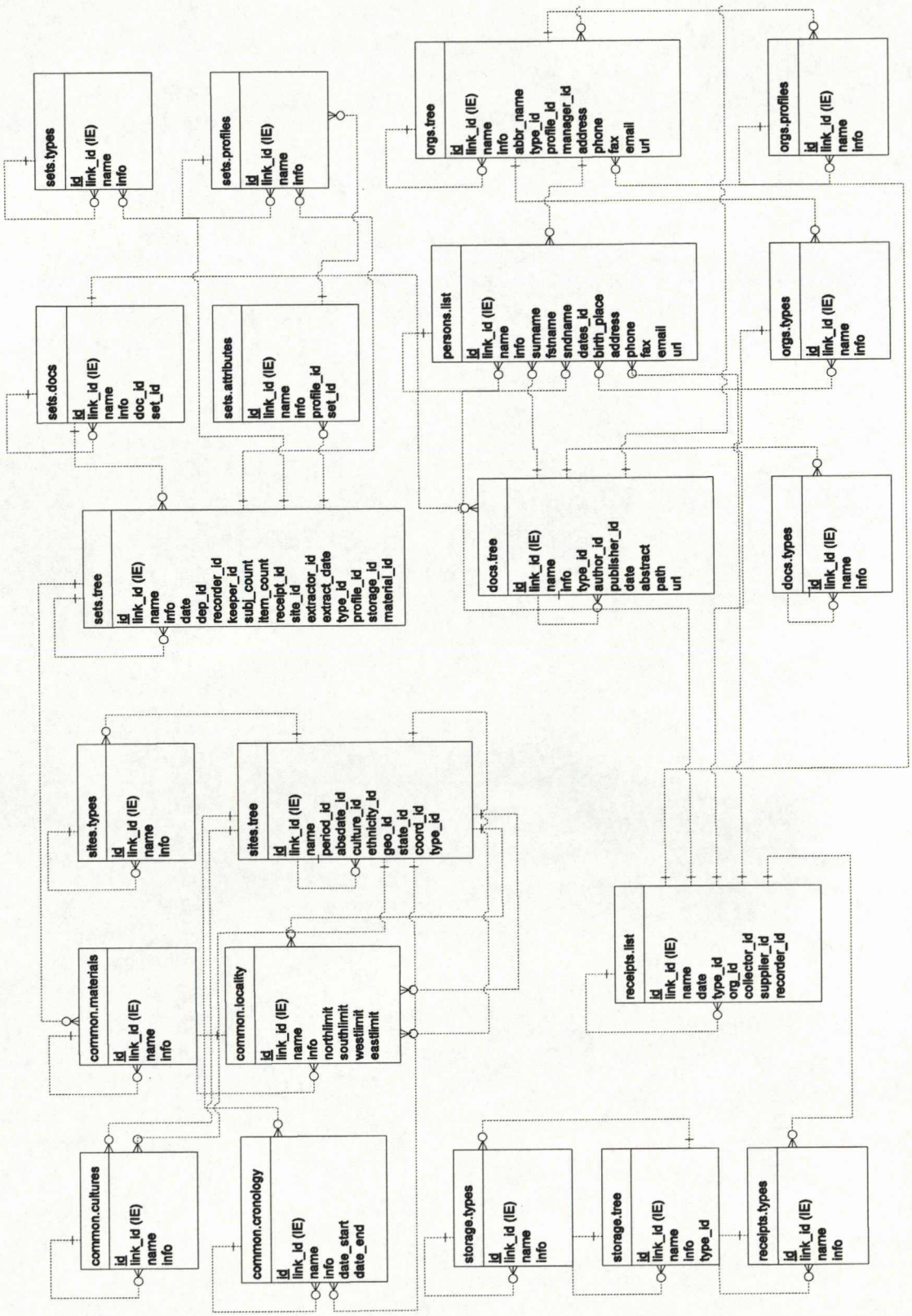

Figure 1. Data layout of the Kunstkammer Computer Catalogue.