Ornithological Databases at some German Natural History Museums

Gerald Mayr, Goetz Rheinwald & Friederike Woog

Forschungsinstitut Senckenberg, Zoologisches Forschungsinstitut und Museum Alexander Koenig and Staatliches Museum für Naturkunde Stuttgart

Abstract. At the moment there is no cataloguing software worldwide that fits all demands for avian collections. Scientists working at three German Natural History Museums report on their cataloguing system and demonstrate their experience. This joint lecture was presented at the symposium to stimulate the discussion on the status and progress in bird collection cataloguing in Europe. Is there any advice that can be given to those that have not started cataloguing yet?

Key words. Cataloguing, avian collections, experience in ornithological cataloguing, Germany

1. INTRODUCTION

Until now there has been no joint initiative for all bird, animal or natural history collections in Germany to develop a cataloguing software for the computerisation of museums and institutes. The same is true on the international level. In this situation several natural history museums and institutes started to develop a cataloguing software on their own. In this joint paper we give an insight into the progress and failure of three large German natural history museums. After reports about the Research Institute Senckenberg, Frankfurt, the Zoological Research Institute and Museum Alexander Koenig, Bonn and the Staatliches Museum für Naturkunde, Stuttgart we discuss the possibility to access data from different databases by means of a common internet platform, i.e. using the software Species Analyst (http://speciesanalyst.net/) (TSA, see the paper of Peterson & Navarro-Sigüenza, this issue).

2. GERALD MAYR: DATABASES USED AT THE ORNITHOLOGICAL SECTION OF THE RESEARCH INSTITUTE SENCKENBERG

The ornithological section of the Research Institute Senckenberg in Frankfurt has so far entered about 16,000 data sets into the database system Microsoft ACCESS. The entire skeleton and spirit collection (about 4,000 specimens each), and about 8,000 of a total of more than 90,000 skins as well as all newly acquired specimens have been entered into the database. Due to limited personnel capacities, the current entry rate of earlier specimens (which are catalogued on index cards) is rather low: about 2,000 specimens a year. Only those data which are already on index cards are entered into the database, i.e. the name of the taxon, collecting data, locality, and the collector. The locality names are updated, but not verified against the actual specimen label; coordinates are normally not added. The taxonomy follows Peters (1931-62). Currently, collection data stored in ACCESS are not available on the internet.

In the course of 2002, a new database system was installed which is especially designed for the Research Institute Senckenberg and is named SESAM (SEnckenbergische SAMmlungen, i.e. Senckenbergian collections). The complete information currently stored in ACCESS will be transferred into SESAM, so that none of the current data sets is lost. SESAM will be used by all sections of the Research Institute Senckenberg. Therefore, a large variety of data fields will exist to meet the needs of all sections of the institute. It will then be possible to add digital scans of the original labels, additional fields for the identification history of a given specimen and all loan arrangements of a particular specimen.

SESAM is a relational data base, so that common data sets, like coordinates for a defined locality, will be available to all users and are entered by any person into SESAM only once. Some data, not yet specified, will later on be accessible for external users over the internet.

3. GOETZ RHEINWALD: CATALOGUING AT THE ORNITHOLOGICAL SECTION OF THE ZOOLOGICAL RESEARCH INSTITUTE AND MUSEUM ALEXANDER KOENIG, BONN

In the early 1980s we started with an inventory of the skin collection. Based on the inventory we prepared a card index for each taxon, i.e. to species or subspecies level. As far as time was available we then transferred the data of the labels onto the cards. The skeleton collection as well as the fluid collection have already been on index cards since the end of the 1960s. When about 50 % of all skins were entered onto these cards, the elec-

tronic age started. On the basis of PARADOX, the cataloguing software BIODAT (first version) was developed, and since the beginning of the 1990s cataloguing had to switch from cards to electronic databasing.

We learned to handle a rather difficult cataloguing system and until 1997 we entered another 20 % of the skin collection into BIODAT (first version). Meanwhile, the experts had developed another version of BIODAT which is much easier to handle. The new relational database is linked with several thesauri, one containing the avian system and another geography with coordinates for all localities in the background. It is possible to build up these thesauri by adding taxa and names of locations, but it is more desirable to import both tables from elsewhere.

There are two fundamental disadvantages in the new version of BIODAT. One is that all entries of the labels have to be verified in advance. That means firstly, that the taxon name has to be adjusted to the thesaurus stored in the background. At the moment BIODAT is not able to deal with synonyms. It also means that all localities have to be checked and changed according to the most recent geographical thesauri. It is this second point which does impend the transfer of those 20 %

computerized skins from the first version into the second one, so that all entries between the beginning of the 1990s and 1997 have been lost.

The second „mistake" is more serious. In the first version we had a master file number for every specimen that connected the different tables in the database. In the second version, the collection number connects the different tables. But for this task the collection numbers have to be in numeric format (numbers only). All labels from the days of Alexander Koenig – we have many thousands of skins from that time – have an alphanumeric format: a combination of Greek and Roman letters, with Arabic and Roman numerals. This means that we would have to change about one third of the collection numbers in the avian skin collection. It is obvious that this will never be done.

At the moment, all computerizing of the collection has been stopped and we are discussing the possibility to start anew with a rather simple system like ACCESS or EXCEL because it is always possible to transfer data from such a simple database to a more sophisticated database, the opposite way being more difficult. It is our feeling that BIODAT is an excellent tool for collection management but not for cataloguing.
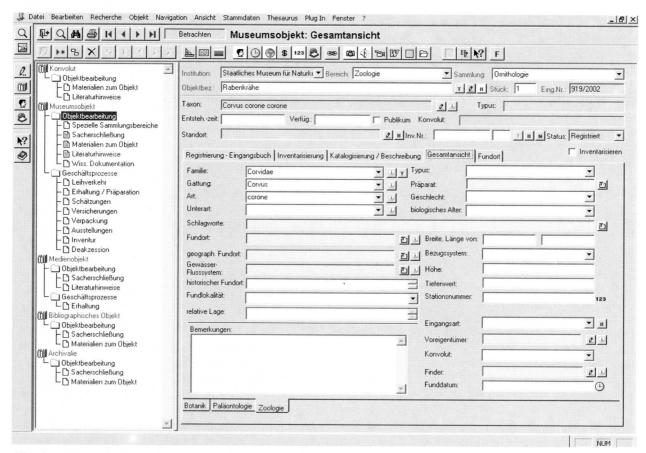

Fig. 1: Entry mask of IMDAS-Pro®

4. FRIEDERIKE WOOG: IMDAS PRO®
– A SCIENTIFIC DATABASE FOR MUSEUM COLLECTIONS

In 1997, as part of the MusIS (Museum Information System) Project, the State of Baden-Württemberg contracted Joanneum Research to adapt their IMDAS-Pro® inventory software for the use at museums. The aim was to develop a large-scale database which would serve all museums in Baden-Württemberg. To develop the database for the use in natural history collections, a pilot project was launched at the Staatliches Museum für Naturkunde Stuttgart (SMNS). Both, scientists and computer specialists (Angela JANDL and Markus GRABERT) worked out the structure and user interfaces of the database, and after various tests and improvements IMDAS-Pro® is now being used by various departments (i.e. in malacology, palaeontology and photographic archive). Numbered museum objects can be linked with data on i.e. sound, digital photographs, literature, preservation technique or habitat. So far, about 50,000 bird data sets have been stored in a dBase database, but the transfer of the complete dataset into IMDAS-Pro® is planned.

The IMDAS-Pro® database is a relational database for multi-users with ORACLE® or for single-users with WINDOWS®. It contains about 370 data fields with over 200 background tables and links with numerous thesauri (i.e. systematics, location, collector). The thesauri can be updated easily. Import and export of data from other programs is possible as well as user restrictions (i.e. for guests or volunteers).

Media objects can be inventoried, i.e. sound or film, with information on when and where a tape was recorded, under what conditions, duration etc. Digitised images (i.e. photographs, sonagrams) can be linked to the main objects. Simple, detailed, complex and even SQL searches are possible. For curators, this is probably the most important tool. Searches can be carried out using ,and', ,or' and ,and not' commands. Searches are not restricted to the bird database. It is possible to search objects within other databases at the museum (mammals, fossils). Within the systematics database, details about the taxonomic status of a species can be found (at the moment after HOWARD & MOORE 1998). Subspecies can be included, but a complete thesaurus is not yet available.

IMDAS-Pro® is more than a simple program for inventoring museum specimens. It is a complex database that allows linking data with multiple thesauri and from different collections. Thesauri on the systematics of molluscs and mammalian fossils as well as geographical regions i.e. Germany, Pacific islands and Ecuador are being developed or refined.

5. DISCUSSION

With their software „Species Analyst" (http://tsadev.speciesanalyst.net/), Townsend PETERSON and colleagues have shown that it is possible to access different kinds of databases on natural history collections worldwide from a common platform. Therefore, it is not necessary to use a particular database, but one that suits an institution best. The variety of databases available is confusing and the choice which one to use will be difficult for many. For more information on this point see in PETERSON & NAVARRO-SIGÜENZA, this issue.

In order to save time and money, institutions that have not yet started to computerize their collection are advised to use existing software, such as IMDAS Pro®, SESAM or Specify 4.0. Contact addresses for test versions are given below. The databases available vary in their complexity, and, more importantly, cost. A single user license for the sophisticated IMDAS Pro® is about 1050,– + tax, the SESAM software is still at a trial stage and Specify 4.0 is free (download at http://usobi.org/specify/). Small museums may not be able to afford a database like IMDAS Pro® and may use a less expensive or free option such as Specify 4.0, that fulfills most requirements.

If an institution decides to develop their own database anyway, we advice to follow the adapted guidelines of the DNFS (1997). The entry mask should contain the fields as in Fig.2, where we distinguish between obligatory fields (in bold), general fields and facultative fields (*).

We recommend not to store too many data under „remarks" but to create additional, optional fields, i.e. for measurements, colours etc. Everyone should start to computerize their collection data as soon as possible, to save time when answering data requests and to be able to share the data worldwide. Even using a relatively simple Access or Excel table is possible, as export options within these programs will allow data transfer to a more complex database. Some scientists and collection managers hesitate to simply transfer data from bird labels or index cards onto the computer without thoroughly checking and reidentifying the bird in the hand. Although at some point this needs to be done, it takes a lot of time and, especially for large collections, it may not be possible to finish this task in the lifetime of one person. Therefore, we suggest computerizing data in a rough, first run. This can even be done by data inputters or students without any thorough knowledge in ornithology. When checking the input data, simple data queries will quickly show where data needs to be checked, changed or

Systematic	Locality	Object	Others
Phylum	Continent	**Coll. number**	**Collector**
Class	**Country**	Field coll.number	Coll. method
Order	Province	Conservation	Wing length
Family	Area	Age	Tail length
Genus	**Locality**	**Sex**	Bill length
Species	Latitude	Diff. parts***	Bill depth
Subspecies	Longitude	Room*, Closet*,	Tarsus-
Subclass*	Method to gain	Drawer*	Metatarsus length
Suborder*	coordinates**	**Coll. date**	Claw length
Subfamily*	Biogeogr. region		Eye colour
Subgenus*	Ecological data		Colour of bare
Identificator			parts
Classification			Size of gonads
following			Remarks

** = gained by GPS, Times Atlas, estimated value, Gazetteer or others

*** = from the same specimen parts exist in another collection (skeleton, fluid, tissue)

Fig. 2: Fields for the entry mask. Obligatory fields in bold, facultative fields with *.

Our natural history collections are the basis for thorough science on avian systematics, taxonomy, evolution, ecology, morphology, changes in bird distribution and contain tissues for environmental studies. Transferring this valuable information from dusty cabinets into a database is the first, most important step to safeguard collections in the future. Funding bodies should understand the importance of databasing these unique witnesses of natural history and provide sufficient funding.

Useful web-sites:

International Working Group on Taxonomic Databases – http://www.tdwg.org/

GBIF (Global Biodiversity Information Facility) – http://www.gbif.org/

improved. Once the data is in the database this can be easlily queried i.e. for certain specimens or specimens coming from a particular country. In a data use agreement the museum can state that data may be inaccurate and that researchers shall not use the data unchecked. Any serious scientist will visit a collection to check the provided data anyway. With the help of these visiting specialists data can be improved further and further. If there is ever such a thing as a final state of the database, it should contain coordinates so that the name and the spelling of a given locality is clear, the specimens should be identified to the subspecies level and the name of the taxon should agree with one of the accepted avian classifications. If possible, the collection date should be verified according to the life data of the collector and where possible checked against a diary of the collector. To simplify future data entry, thesauri on taxa, geographic information or collectors should be developed.

Regardless of the database in use, splitting up work on thesauri on i.e. different geographical regions or collectors between museums world-wide could avoid duplicate efforts. As most databases can be made compatible, a thesaurus from one museum could be linked or combined with the thesauri in use by another museum. A comparative thesaurus of bird species and subspecies as developed by Rémy BRUCKERT from MNHN, Paris (see this issue) could be shared with museums throughout Europe to avoid further duplicate efforts. An important step would be to set up a European thesaurus working group to coordinate these efforts.

Contact adresses for

Specify 4.0 http://usobi.org/specify/

IMDAS Pro® grabert.smns@naturkundemuseum-bw.de, or Markus GRABERT, Staatliches Museum für Naturkunde, Rosenstein 1, 70191 Stuttgart, Tel. +49/711/ 8936255

SESAM Lothar.Menner@senckenberg.de, or Lothar MENNER, Forschungsinstitut Senckenberg, Senckenberganlage 25, D-60325 Frankfurt/M, Tel. +49/69/ 7542363

Literature

Direktorenkonferenz naturwissenschaftlicher Forschungssammlungen Deutschlands (1997): Datenbanken und Datenbankanwendeungen zur Sammlungsverwaltung an Instituten der DNFS – ein Zwischenbericht. – EDV-Arbeitsgemeinschaft der DNFS.

Dr. Gerald MAYR, Forschungsinstitut Senckenberg, Senckenberganlage 25, D-60325 Frankfurt/M, email: Gmayr@sng.uni-frankfurt.de

Dr. Goetz RHEINWALD, Zoologisches Forschungsinstitut und Museum Alexander Koenig, Adenauerallee 160, 53113 Bonn, e-mail g.rheinwald.zfmk@uni-bonn.de

Dr. Friederike WOOG, Staatliches Museum für Naturkunde Stuttgart, Rosenstein 1, 70191 Stuttgart, e-mail: woog.smns@naturkundemuseum-bw.de