GLOMYRIS and TYMUNAC: Myriapoda and Acari databases of the GBIF-D node invertebrates II

Roland R. Melzer, Stefan Friedrich, Marc Ritzerfeld, Jens Bohn & Jörg Spelda


The section Arthropoda varia at ZSM houses large collections of Myriapoda and Acari of high taxonomic importance, including those of K. W. Verhoeff, Graf H. Vitzthum, L. Kneissl, C. Willmann, and W. Hirschmann. Within the framework of the node invertebrates II of the German Biodiversity Information Facility (GBIF-D), internet accessible databases are established providing information on type material, synonymy, taxonomic position, references, type locality and various media files. In this article we review our past and current activities, give examples for data queries and types of biodiversity information included. Moreover, we give a brief summary of the activities of the node invertebrates II and the way these activities are embedded in the emerging field of cybertaxonomy. It is our medium term aim to establish comprehensive electronic type catalogues including detailed pictorial information, and our long term aim is to contribute to global information systems for Myriapoda and Acari.

Zoologische Staatssammlung, Münchhausenstr. 21, 81247 München, Germany; e-mail: melzer@zsm.mwn.de

On the trail to “cybertaxonomy”

Currently about 1.8 million species are described of an estimated number of between 10 and 100 million existing species (Wilson 2005). Based on the work of taxonomists for centuries, a wealth of information about these species has been gathered, and was made available in taxonomic publications as well as by natural history collections housing type and other material plus sample data. Moreover, in addition to α-taxonomic information, data on biogeography, ecology, climate, phylogeny and many more have been accumulated in related disciplines of biology.

Next to analysis of species-specific sets of characters, the process of describing new species is mainly based on a checking and rechecking system in which newly recovered, yet not described species, closely related already known species and the available biodiversity information are compared and integrated in a sophisticated system of verification and falsification of species definitions and delimitations (e.g. Mayr & Bock 2002, Melzer & Schmidt 2008). This process leads to acceleration of data capture, since as more species are known as more detailed the descriptions and differential diagnoses have to be made. This also requires redescription of many species studied earlier using modern standards of character analysis. Furthermore, for several years molecular taxonomy opened a second field of data acquisition that is independent of the morphological datasets resulting in yet another acceleration of data amassment (Tautz et al. 2002, Hebert et al. 2003).

Exchange of taxonomic information between specialists is a crucial element of this system, and for over two centuries this has been accomplished by using journal articles available at institutions and a worldwide loan-and-return procedure that allows specialists of a given taxon for studying type and other material from collections all over the place – a process of high quality, but very slow, and therefore leading among other things to what some scientists recently call a taxonomic impediment (Lipscomb et al. 2003).

Exchange of taxonomic information between specialists is a crucial element of this system, and for over two centuries this has been accomplished by using journal articles available at institutions and a worldwide loan-and-return procedure that allows specialists of a given taxon for studying type and other material from collections all over the place – a process of high quality, but very slow, and therefore leading among other things to what some scientists recently call a taxonomic impediment (Lipscomb et al. 2003).

Next to “publications and lots” essential for integration of data in classical taxonomy were therefore catalogues and the collections’ books of arrivals in
which information on the collected and stored material was meticulously registered, e.g. on species, sample localities and circumstances, collection and rack numbers etc. pp. Thus, taxonomic work includes two main elements, (i) the scientific analysis and information recovery, and (ii) the administration of data as well as information integration and retrieval. It is trivial that for decades the second part has been accomplished using databases rather than written catalogues. These databases were built on various computer programs, modified in a multitude of ways for different taxa, museum departments, localities etc., and therefore mainly useful for the direct user in natural history collections. Accessibility for the scientific community was thus limited in the same way as the access to stored material and/or depending on their publication in print media.

Looking at the mismatch between the limited speed of the old data exchange systems and the above described boost of data, it is a fortunate coincidence that since some years the establishment of the worldwide web enabled taxonomists to do a quantum leap in data availability, data exchange and data sharing by taking the opportunity to combine their classical work structures with modern information technology: Biodiversity informatics and/or “cybertaxonomy” uses common platforms to make morphology – as well as DNA-based – species descriptions directly available for everybody, to cross-link various categories of biodiversity data, and to integrate new technologies, e.g. modern imaging, image analysis techniques and GPS tracking (Döring & Behrendsohn 2007, Mayo et al. 2008, Zauner 2009), and therefore can cope with the data masses in a much more effective way than before. One example among numerous others for such a web-based project is the Global Biodiversity Information Facility (GBIF). The latter feeds a portal with primary biodiversity data accessible to everyone (http://www.gbif.org/).

Recently, a backlash of IT technology onto the basic taxonomy workflow has also been projected, like open access publication of species descriptions (Michel et al. 2009), long-term accessibility of descriptions and update of the zoological code (ICZN 2008) (http://www.iczn.org/), validation of species names by registration in Zoobank rather than publication (http://www.zoobank.org/), computer readable documents cross-linked to other types of resources (Agosti & Egloff 2009), and integration of descriptions and various categories of biodiversity information in “wikis” (www.species-id.net, Hendrich & Balke (2011); see also GBIF Task Group on a Data Publishing Framework (http://www.gbif.org/communications/).

In the framework of these recent developments, the GBIF portal with its actually more than 267 million internet accessible data records (May 2011) holds a key position in the worldwide exchange of biodiversity information. In addition to the provision of data via the internet, one important task of the numerous subprojects contributing to the portal is to advance the awareness of specialists of the available data and the possibilities of data retrieval for a given taxon. For example, a foreign Acari workgroup recently wanted to produce a catalogue of type specimens of a mite subtaxon stored at the ZSM not being aware that all the respective data were already made available via the internet by the TYMUNAC subproject of GBIF. In the present paper, we therefore give a brief overview of the activities of the node invertebrates II of GBIF Germany and review the data provided by two of the node’s subprojects dealt with by the Sektion Arthropoda varia of ZSM: GLOMYRIS, the Global Myriapod Information System, and TYMUNAC, Types of Munich Acari Collections.

The node invertebrates II of GBIF Germany

Since 2002 the node invertebrates II of GBIF Germany (http://www.gbif.de/evertebrata2, Haszprunar & Melzer 2003) has been established at the ZSM (http://www.zsm.mwn.de), representing the Mollusca (molluscs), Chelicerata (including spiders and mites) and Myriapoda (millipedes and centipedes) housed in German zoological collections (Bohn et al. 2003, 2004, Knebelsberger et al. 2005, Schrödl et al. 2004, Spelda et al. 2004, 2005, Spelda 2005, http://www.biologie.uni-ulm.de/cgi-bin/query_all-pl?lang=d&pr=gbif-e2). During the first funding period (German Federal Ministry of Education and Research, BMBF, grant 01LI0205),

Fig. 1. A. K. W. Verhoeff. B. One of many drawers containing Voerhoeff’s micro preparations (altogether 11,000 micro slides and lots in ethanol are stored at ZSM). Note ZSM sample and database numbers on red labels marking type specimens assigned during the GLOMYRIS project. Every object has to be cross-checked with the original descriptions. C. Glomeris malminoga (Verhoeff, 1912) (macrophoto: Jörg Spelda), original types at ZSM. D. Media file of gonopod of a still undescribed polydesmid first mentioned in the Verhoeoff manuscript of the early 1940ies shown in E. E. Scan of unpublished manuscript by Verhoeoff. F. Iconotypes of Mexican Polydesmida described by De Saussure in 1860; scan from original book formerly in possession of Verhoeff with his hand-written notes. Such iconotypes are included in the GLOMYRIS database and used to identify the original type specimens.
the node coordinated databasing of type specimens housed in German museums. A survey of the seven contributing subprojects, the number of type records made available through the GBIF and Systax portals between 2003 and 2006, and the databases used since this funding period is given in table 1. In addition, thousands of records of non-type material, media files, taxon literature connections etc. have been included in the databases since then.

Furthermore, the node invertebrates II is a co-operation partner of Edaphobase – GBIF-Database Soil Zoology (http://www.senckenberg.de/root/index.php?page_id=11082), and contributes records of centipedes and millipedes. Including a geographic information system (GIS), Edaphobase will provide cross-links between the species’ distributional records and ecological parameters, and make them accessible via the GBIF-portal.

Since 2010 the node Evertebrata II is funded again by the BMBF for the joint research project “GBIF-D, centres of excellence of innovative data mobilisation” (grant 01LI01001 B). Herewith the node will focus on innovative imaging techniques, e.g. high resolution and extended field of depth pictures of samples ranging from minute mites or millipede gonopod preparations on slides to the digitization of collection boxes each containing numerous specimens, and the mobilisation of these data. The second focus will be on data on non-type material, observational and geo tagged/geo referenced data.

Further cross links of the node’s activities are with 4D4Life (http://www.4d4life.eu/) where a global species database for Diplopoda, Pauropoda and Symphyla is in development and via the “DNA Barcoding Fauna Bavarica” project (funded by Bayerisches Staatsministerium für Wissenschaft, Forschung und Kunst; http://www.faunabavarica.de/) to the International Barcode of Life and Barcode of Life Database, respectively (IBOL and BOLD; http://ibol.org/, http://www.boldsystems.org). Herewith not only samples for barcoding, but also media files of the studied species are contributed. All these activities are aimed at the vision of establishing global information systems which should even

Table 1. Subprojects of GBIF-D node invertebrates II, databases used and recorded type material accessible via the GBIF- and Systax portals during first BMBF funding period grant (01LI0205).

<table>
<thead>
<tr>
<th>Eve II subprojects</th>
<th>databases</th>
<th>recorded types</th>
<th>workgroups and cooperating institutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>TYTEARBER</td>
<td>Systax</td>
<td>c. 2200</td>
<td>Anja Friederichs, Jason Dunlop</td>
</tr>
<tr>
<td>TYtypes of TErrestrial ARthropods at the BERlin Museum für Naturkunde</td>
<td>c. 2200</td>
<td>Museum für Naturkunde der Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin</td>
<td></td>
</tr>
<tr>
<td>TYTEARFIS</td>
<td>SeSam, Systax</td>
<td>c. 11,600</td>
<td>Peter Jäger, Forschungsinstitut und Naturmuseum Senckenberg, Senckenberanlage 25, D-60325 Frankfurt</td>
</tr>
<tr>
<td>TYtypes of TErrrestrial ARthropods at Forschungsinstitut Senckenberg</td>
<td>SeSam, Systax</td>
<td>c. 11,600</td>
<td>Jörg Spelda, Markus Unsöld, Roland Melzer, Forschungsinstitut und Naturmuseum Senckenberg, Senckenberanlage 25, D-60325 Frankfurt</td>
</tr>
<tr>
<td>GLOMYRIS</td>
<td>Specify, Systax</td>
<td>c. 9000</td>
<td>Jörg Spelda, Markus Unsöld, Roland Melzer, Forschungsinstitut und Naturmuseum Senckenberg, Senckenberanlage 25, D-60325 Frankfurt</td>
</tr>
<tr>
<td>GLOBal MYriapod Information System</td>
<td>Specify, Systax</td>
<td>c. 9000</td>
<td>Jörg Spelda, Markus Unsöld, Roland Melzer, Forschungsinstitut und Naturmuseum Senckenberg, Senckenberanlage 25, D-60325 Frankfurt</td>
</tr>
<tr>
<td>TYMUNAC</td>
<td>Specify, Systax</td>
<td>c. 1800</td>
<td>Jens Bohn, Maria Fernanda Montoya, Roland Melzer, Zoologische Staatssammlung München, Münchhausenstr. 21, D-81247 München</td>
</tr>
<tr>
<td>TYtypes of MUNich ACari Collections</td>
<td>Specify, Systax</td>
<td>c. 1800</td>
<td>Jens Bohn, Maria Fernanda Montoya, Roland Melzer, Zoologische Staatssammlung München, Münchhausenstr. 21, D-81247 München</td>
</tr>
<tr>
<td>MOTYBER</td>
<td>Systax</td>
<td>c. 6400</td>
<td>Matthias Glaubrecht, Frank Köhler, Museum für Naturkunde der Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin</td>
</tr>
<tr>
<td>MOlluscan TYtypes at the BERlin Museum für Naturkunde</td>
<td>Systax</td>
<td>c. 6400</td>
<td>Matthias Glaubrecht, Frank Köhler, Museum für Naturkunde der Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin</td>
</tr>
<tr>
<td>MOTYFIS</td>
<td>SeSam, Systax</td>
<td>c. 10,400</td>
<td>Ronald Janssen, Eike Neubert, Forschungsinstitut und Naturmuseum Senckenberg, Senckenberanlage 25, D-60325 Frankfurt</td>
</tr>
<tr>
<td>MOlluscan TYtypes at Forschungsinstitut Senckenberg</td>
<td>SeSam, Systax</td>
<td>c. 10,400</td>
<td>Ronald Janssen, Eike Neubert, Forschungsinstitut und Naturmuseum Senckenberg, Senckenberanlage 25, D-60325 Frankfurt</td>
</tr>
<tr>
<td>MOTYMUNHACIS</td>
<td>Specify, Systax</td>
<td>c. 1100</td>
<td>Bernhard Hausdorf, Thomas Knabelsberger, Michael Schrödl, Vollrath Wiese, Zoologische Staatssammlung München, Münchhausenstr. 21, D-81247 München, Zoologisches Institut und Museum Hamburg, Martin-Luther-King-Platz 3, D-20146 Hamburg, Haus der Natur – Cismar, Bäderstr. 26, D-23743 Cismar</td>
</tr>
</tbody>
</table>

Types recorded >40,000
provide the possibility of online determination in the future.

**The Global Myriapod Information System (GLOMYRIS)**

The general scope of the GLOMYRIS project (http://www.gbif.de/evertebrata2/glomyris) is to provide databases that facilitate taxonomic research on Myriapoda. Hence, its main idea is to establish a global myriapod information system providing data on, e.g. taxon names, their status, descriptions, revisions and geographical distribution. Within this framework GLOMYRIS activities developed in several directions.

(i) During the first GBIF phase funded by BMBF from 2002 to 2006 the main scope was the establishment of an online-available database for Myriapoda type specimens housed in German natural history museums, with special reference to the so-called “Verhoeff” collection. Karl Wilhelm Verhoeff (1867–1945) (Fig. 1A) was one of the most important myriapodologists of the past. He described about 2000 taxa and published circa 670 papers and books on Myriapoda and other arthropods (Verhoeff & Mauer-mayer 1962, see also overview in Verhoeff 1934). Shortcomings of this huge heritage are the distribution of the collection over several museums, with the main sections deposited at ZSM in Munich, MfN in Berlin and NHM in Vienna, chaotic type designation and labelling provoking numerous taxonomic and conservatory problems, e.g. type conflicts. Refurbishment and databasing of the material studied by Verhoeff deposited in Munich and Berlin was thus the core of this GLOMYRIS project phase (Table 2). According to the node concept of GBIF Germany, recording of material from German museums allows standardised data display for each myriapod taxon according to consistent taxon concepts. The latter has been achieved by introducing taxonomic backbones for all the contributing museums.

Following this concept, all of Verhoeff’s gonopod micro preparations stored at ZSM (2052 of Chilopoda, 3924 of Diplopoda) were recorded, including numerous object and label scans (Fig. 1B). Furthermore, the recording of the ethanol stored material including the specimens correlated with gonopod preparations has also been done at the level of lots. The latter, however, may contain numerous specimens which still have to be separated into different sample vials and database entries. Apart from this the old inventory catalogues of the MfN Berlin (3500 type specimens), and the Zoologisches Institut and Museum Hamburg (c. 500 type specimens) have been recorded as well (Moritz & Fischer 1973, Weidner 1960).

A central element of the system are cross links between taxon names and literature records that can be used for identifying type specimens (Spelda 2005). As a side effect of this procedure, 42,000 literature datasets were compiled from 2600 publications on myriapods that are also available via our portal. Furthermore, additional information has been gathered and included in the GLOMYRIS mediafiles (c. 5000), such as macro photos of species described by Verhoeff (Fig. 1C), photos of Verhoeff’s gonopod preparations (Fig. 1D), scans of original literature as well as of Verhoeff’s unpublished manuscripts (Fig. 1E) and iconotypes from old literature (Fig. 1F).

The GLOMYRIS project uses a local database system for primary data storage which contains features not yet implemented in the actual portals. The home databases were mirrored to SysTax, and are from there accessible via the internet from the Systax (http://www.biologie.uni-ulm.de/cgi-bin/query_all.pl?lang=d&pr=gbif-e2) and GBIF portals. Examples for Specify data entry masks and backbone taxonomies are depicted in Fig. 2F,G, and for a Systax data query in Fig. 2H–K for Acari types and can be carried out accordingly for our myriapod specimens.

(ii) Databasing Verhoeff’s collection revealed that as a second element of our activities type designation and type research programs are strongly needed. Due to missing or unclear type designations in many cases type “candidates” rather than actual type specimens had to be recorded. Furthermore, many micro preparations that appear as a single database entry contain several gonopod preparations or preparations made from a single object are distributed over several slides. In addition, various types have been recorded from the literature for which the deposition is unclear. For all these specimens, careful cross-checks with the original descriptions have to be made to identify the original type material and make clear designations. Since Verhoeff’s analyses in many aspects are superficial and cannot easily be corroborated with modern standards this makes redescriptions using modern techniques unavoidable. Such revisions are available for only a few of Verhoeff’s taxa (e.g. Djursvoll 2008, Shelley

<table>
<thead>
<tr>
<th>Table 2. Summary of results on ZSM material (from Spelda et al. 2005).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated type numbers</td>
</tr>
<tr>
<td>Material recorded</td>
</tr>
<tr>
<td>Media files</td>
</tr>
<tr>
<td>Taxon-literature connec-</td>
</tr>
<tr>
<td>Taxa in database</td>
</tr>
</tbody>
</table>
et al. 2005, Stoev 2009, Stoev & Enghoff 2008), and therefore we have started a type research program based on the GLOMYRIS database (Spelda 2001, 2008, Unsöld & Melzer 2003, Pilz et al. 2008. Here we use modern imaging techniques to obtain light microscopic pictures of the original material at high resolution at extended field of depth, as well as Scanning EM pictures.

(iii) In this connection one part of our recent BMBF funded GBIF project phase is the establishment of innovative imaging and image analysis techniques that will produce media files of high quality and connect them with older as well as new database records. This pictorial information ranging from macroscopic to microscopic magnifications, i.e. from habitus to high resolution gonopod photos and their provision via GBIF should therefore improve further the flow of information between housing institutions and external millipede taxonomists, and bring us one step further to online determination aids. Moreover, more myriapod collections including also non-type material plus relevant data from neighbouring fields such as ecology and biogeography will be incorporated. Already now, for some Verhoeff types geographical references to their sample locations using coordinates have been set and can be shown on maps via Systax. In sum, GLOMYRIS might contribute to global species catalogues within Species 2000, and to projects corroborating distributional records, like the European Invertebrate Survey (EIS; http://www.eis-international.org/index.php) and the Fauna Europaea project. Last but not least, with the contribution of GLOMYRIS to the International Barcode of Life cross references between GBIF data and DNA sequences of our species databased in BOLD (http://www.boldsystems.org/views/login.php) will be made.

**TYMUNAC – The Type Database of the Munich Acari Collections**

The “Arthropoda varia” section of ZSM houses five large Acari collections of high taxonomic importance, i.e. those of acarologists Graf H. Vitzthum (Fig. 2A), L. Kneissl (Fig. 2A), C. Willmann, W. Hirschmann (Fig. 2A) and E. Popp with altogether about 3700 type specimens, mainly in Hirschmann’s Uropodina. Within the TYMUNAC project a database was built providing thorough information on a big part of this type material including the locus typicus, citations, synonymy, taxonomic “backbone”, and media files accessible online via the GBIF-portal and SysTax, respectively (Fig. 2B–K) (Bohn et al. 2003, 2004).

The basic task of the TYMUNAC project was very similar to that of GLOMYRIS: Only the Vitzthum and Kneissl collections were at a good conservational state (Fig. 2C,D), while especially those of Willmann and Hirschmann suffer from bad conservation and unclear labeling (Fig. 2E). Together with the workgroup of C. Blaszak, University of Poznan, we have therefore begun refurbishing and partial re-embedding of the collection which was the basis for various catalogues of Acari subgroups published in the last years (Skoracki et al. 2002, Ferenc et al 2002, Bloszyk et al. 2003, Blaszak et al. 2003, 2004a–c, 2007, Gwiazdowicz et al. 2007, 2008a, 2008b, Olszanowski et al. 2007).

Next to this conservation and cataloguing work, the TYMUNAC workgroup has the task to establish a database containing the available information on Acari type material stored at ZSM. As the task is similar to that of GLOMYRIS, a similar work schedule is followed. As backbone taxonomy J. Hallan’s extensive Acari taxonomy was introduced (Fig. 2G; see also http://insects.tamu.edu/research/collection/hallan/acari/Acarina.htm). Hallan’s list includes a modern and complete mite system including all taxa down to the level of subspecies created by leading acarologists. Then, presumed type specimens are recorded using the information either from micro preparation labels or an older card-file inventory or sometimes even handwritten notes from the authors. This does not ensure that the correlation between type information and micro preparation is properly set. Since information from the labels and catalogues is often very scarce, the correct type status and the proper locality of the specimens needs consultation of the original species descriptions. Furthermore, additional specimen data useful for scientists using the database are included, e.g. media files of types and labels.

As with GLOMYRIS, data are stored in a local database used for recording (Specify 4 to 6, Fig. 2F,G), and provided via the Systax and GBIF-portals. Following this procedure, the Vitzthum
collection has been completely recorded during the 1st GBIF phase. A survey of the respective number of datasets and the types of information included is given in Table 3. Examples for the recorded data and possibilities of online data search are given in Fig. 2H–K depicting type information and media files. Recently, databasing of Popp’s collection (more than 1200 micro preparations) has been finished as well including label data and specimen scans, and hence our next task will be the completion of the Willman/Hirschmann records.

[Table 3]

TYMUNAC will use the same modern imaging techniques as described for GLOMYRIS established for our second GBIF funding phase. The main goal is thus to expand our database in a way that it fulfills the needs of an electronic type catalogue including high quality media files, viz. a complete pictorial atlas of the Acari types deposited in the ZSM.

Table 3. Summary of results of the TYMUNAC project *including overlap with oribatid taxa from the OBIF-project; **holotypes, syntypes, lectotypes and neotypes (from Bohn et al. 2004).

| Estimated number of types | 3770 |
| Material recorded | 2281 |
| Media files labels | 524 |
| Media files specimen | 681 |
| Taxon literature connections | 1412 |
| Taxa in database | 18,200* |
| Number of types | 1619 |
| Number of primary types | 938** |

Acknowledgements

The GBIF-D node invertebrates II was and is supported by the GBIF program of the German Federal Ministry of Education and Research, umbrella project IDs 01LI01005 (“GBIF-Germany: Knoten Evertrebra II (Mollusca, Chelicerata, Myriapoda)”), and 01LI01001 B (“GBIF-D, centres of excellence of innovative data mobilisation”). It is a cooperation partner of “GBIF-Informationssystem Bodenzoologie” supported by the German Federal Ministry of Education and Research, project ID 01LI01001A. Special thanks to Jürgen Hoppe and his team (University of Ulm) for hosting our databases in Systax and feeding the GBIF-portal.

References


