Newsletter

Biodiversity Heritage Library for Europe

BHL

Biodiversity Heritage Library

EUROPE

www.bhl-europe.eu

CONTENT

Editorial

The project Biodiversity Heritage Library for Europe (BHL-Europe) is intended to help scientists, teachers, artists and the general public to access a wide variety of rare and common, old and new taxonomic publications, as well as other biodiversity publications, supplemented with illustrations often representing masterpieces of art.

For dissemination of the idea we have launched the BHL-Europe Newsletter, providing more detailed information about the project and associated activities. These are largely digitizing activities centered on taxonomical literature and biodiversity. The main purpose of the newsletter is to provide up-to-date information about the project and new project outcomes, meetings and other information which could be helpful to a variety of users. Additionally, we will also publish news about other projects and initiatives dealing with digitalization and biodiversity. We look forward to the work and hope to provide you with useful information focused on digitalization and biodiversity.

Jiří Kvaček
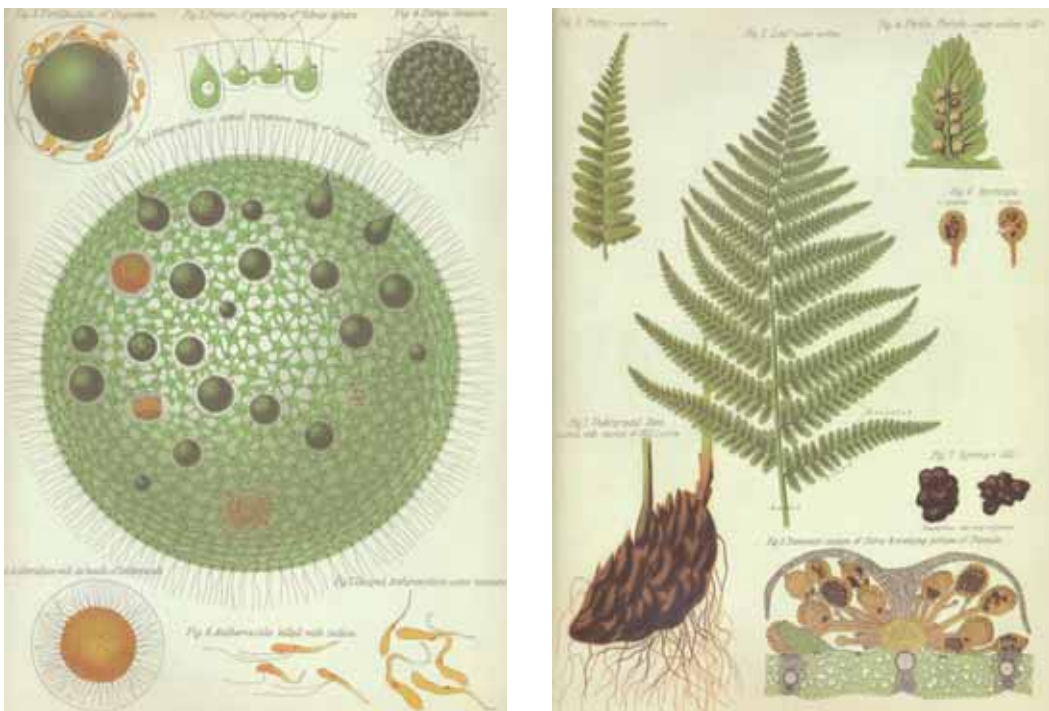
News

Main achievements of BHL-Europe up until 15 August 2009

Kick-off Meeting in Berlin [M1.1]
The Kick-off Meeting of BHL-Europe including two initial workshops was held in Berlin (Museum für Naturkunde) from 10 to 15 May 2009. The main points of that meeting included information about BHL and the BHL-Europe objectives, the introduction of partner institutions and their scanning operations, the constitution of two working groups on the project including the initial workshops for these groups. All our partners sent representatives and altogether up to 60 people attended the meeting.

Catalogue of content holder requirements (quality, quantity, accessibility, standards and specifications of content and metadata) [D2.1]
The first important document leading towards the development of a live BHL-Europe system was the content holder requirements. This document provides a minimum specification for the supply of metadata and digital image files sent to the BHL-Europe project. It will be the base for the content delivery to BHL-Europe agreed between all partners.

Metadata and content harmonisation
The harmonisation of metadata is a very important aspect of the project as every partner is using different standards, different interfaces or different content type. The first test sets of metadata from four consortium partners were imported to a relational database. In addition, the input of a book accompanied by bibliographic metadata was tested using an interim solution. We are now ready to ask for more metadata from other partners to elaborate further on the mapping procedures.

Deliver composition of Technology Management Board [D3.1]
The composition of the Technology Management Board (TMB) was made up from a number of institutions and companies within BHL-Europe, the members include: Kai Stalmann (Museum für Naturkunde Berlin, Leader Work package 2), Adrian Smales (Natural History Museum London, Leader Work package 3, Chair TMB), Bernard Scaife (Natural History Museum London), Walter Koch (Angewandte Informationstechnik Forschungsgesellschaft mbH), Christopher Freeland (Missouri Botanical Garden), Julie Verleyen (European Digital Library Foundation, Europeana), Anton Güntsch (Botanical Garden Berlin, EDIT). The board will have to find a suitable replacement for Bernard Scaife, as he will be leaving the NHM and the BHL project as of 14th August 2009.

BHL-Europe Architecture
The overall BHL-Europe system architecture has been elaborated. As a tool for repository management, FedoraCommons was selected. Emphasis is put into three components: pre-ingest (metadata harmonisation and creation, semantic enrichment using ontology WebServices), storage (integrating cloud technology) and access (providing multilinguality and ontology support for query expansion).

Web site, including multimedia presentation [D5.1]
The Web site of BHL-Europe has been functional since 31 July 2009 under http://www.bhl-europe.eu . It aims to inform the target users about the project and it's progress, the partners and team members, related projects and important events. It will also link to various other resources including BHL and

Europeana. The Web site has a multimedia presentation for download with additional information about the project. Both Web site and multimedia presentation will be updated continuously.

BHL-Europe newsletter and mailing list [D5.2]

The BHL-Europe newsletter was set up to publish important news of the project and report back project results to the public. In addition, articles on scanning activities of project partners and on activities from related projects (e.g. BHL, Europeana, EDIT, EOL) will be included.

Database of relevant conferences/events [D5.3]

BHL-Europe needs to be and will be present at various conferences in the next three years covering (digital) library, biodiversity and information technology conferences and workshops. This will help to make the target users aware of the project, allow active discussion of the project outcomes with the target users, request user feedback and improve the networking capacity of the project. Three papers have already been presented in the

first months of the project and three more papers are already accepted for other conferences.

Deliver composition of Communications Working Group and 1st dissemination plan [D5.4]

The dissemination plan aims to explain the procedures for sharing project outcomes with the target users. We need to identify the key audiences for dissemination, the activities to be carried out and the key messages to be communicated. Dissemination of BHL-Europe will be continuously supported by the Communications Working Group of the project consisting of five full members.

BHL-Europe Summer Meeting

From 9-12 August the second BHL-Europe general and working group meeting was held in Naturalis, Leiden. Thirty team members from 20 of the 28 partners came together to discuss recent results and achievements. The dissemination activities of the project were successfully revised. We also discussed the BHL-Europe hardware infrastructure and architecture. We prepared the next deliverables and a detailed workplan.

http://www.europeana.eu/ is a portal that brings together cultural and scientific material from Europe's libraries, museums, archives and audio-visual archives. Highlights include the Rijksmuseum in Amsterdam, the British Library in London, and the Louvre in Paris. Currently Europeana.eu links you to 4.5 million digital items. These are images, texts, sounds and videos.

This website is a prototype. The next step to develop the Europeana prototype into a fully functional and operational website is being carried out by a new project, Europeana version 1.0. This will allow users to access over 10 million digital items by 2010.

Europeana version 1.0 focuses on users' involvement in the development of Europeana, so that they can themselves determine the services they want. Europeana content will be easy to use on mobile devices. It will be also brought to the users' virtual space. These could be college and school websites, special interest sites, virtual learning and research environments, blogs and social networks.

Europeana's interface is already multilingual; the next step is to start making the search system work across languages, so a word entered in one language is automatically translated and results are retrieved in other languages. Various related projects are involved in developing technical solutions for multilingual search that will be contributed to Europeana.

Europeana version 1.0 will also address copyright and intellectual property rights (IPR). One of the objectives is to make it possible for people to re-use and re-purpose out-of-copyright or rights-cleared content in Europeana and will work with rights holders to achieve this.

Europeana version 1.0 is a 30-month Thematic Network project that began in February 2009. It is coordinated by the EDL Foundation. The Foundation was set up to run Europeana and act as a partner and coordinator in a range of projects that will develop the portal and related services in the coming years. These projects are run by different cultural heritage institutions and comprise the Europeana Group of projects.

Europeana is organising a plenary conference, Creation Collaboration & Copyright, on 14-16 September 2009 in The Hague. More information can be found at http://group.europeana.eu/web/guest/plenary/.

For more information, please contact Aisulu Aldasheva (Aisulu.aldasheva@kb.nl).

Meetings

Forthcoming meetings

**Forthcoming meetings
Workshop on Advanced Technologies
for Digital Libraries, Trento, Italy
8 September 2009**
deadline for registration August 25, 2009
Link: http://www.cacaoproject.eu/at4dl
The workshop aims to bring together the various stakeholders in order to present to them an overview of the state-of-the-art systems in this field and identify open research problems that require further work.

**14th international congress
"Cultural heritage and new Technologies"
Vienna, Austria 16-18 November 2009**
Link: http://www.stadtarchaeologie.at/
Main topic: Archiving - building an information system.

**TDWG 2009 Annual Conference,
Montpellier, France
9-13 November 2009**
http://www.tdwg.org/conference2009/
Biodiversity Information Standards (TDWG) was previously known as the Taxonomic Database Working Group.
Information Standards (TDWG) now focuses on the development of standards for the exchange of biological/biodiversity data.

Meetings attended

**Systematics 2009, Leiden
August 10-14 2009**
Link http://www.biosyst.eu/
Systematics 2009 was the first meeting of Biosyst EU and the 7th Biennial Conference of the Systematics Association. Several formal symposia and numerous contributed paper sessions were scheduled to cover a large variety of topics.

**BHL-Europe at Systematics 2009
in Leiden**
BHL-Europe was invited to give a Keynote lecture during the Systematics 2009 conference. The presentation was given by the Project Coordinator, Henning Scholz, on Wednesday, 12 August 2009. The lecture aimed to introduce the project to the scientific community.
An overview was given outlining past achievements and future activities. A call for participation was made during the lecture which was intended to encourage the audience to help us in improving BHL in the future. This hopefully will lead to contact between potential and future content providers and also to get feedback for future improvements of the BHL-Europe system.

Focus

Welcome to the Biodiversity Heritage Library for Europe

The Biodiversity Heritage Library for Europe (BHL-Europe) is a project involving 26 major European natural history museum libraries and botanical libraries and two libraries from the United States. It is a part of the global Biodiversity Heritage Library network and Europeana – European digital library.

Biological information on animals and plants, often found in scientific books and journals from past centuries, is essential to the implementation of the Convention on Biological Diversity (CBD) of the United Nations. At present, the only way to access this knowledge is to visit a number of different libraries which complicates much of the fundamental research in biological science. Since 2007, the 'Biodiversity Heritage Library' has been systematically removing this barrier by making this literature available online.

BHL-Europe will further develop, expand, and enhance the Biodiversity Heritage Library by bringing together the extensive collections of biodiversity literature held in major European natural history, botanical, and research libraries. As part of the project a European site and a multi-lingual software interface will be developed. These and other enhancements will make the project a global initiative.

The wider public, citizen scientists, and decision makers will have unlimited access to these important sources of information. Everybody will be able to get first-hand information on animals and plants study rare original works of important scientists like Charles Darwin or Alexander von Humboldt, and admire artwork in publications from the 17th or 18th century.

Our ambitious plan is to produce a robust facility, which will be the main source of knowledge for future generations.

The libraries of the European natural history museums and botanical gardens collectively hold the majority of the world's published knowledge on the discovery and subsequent description of biological diversity. As yet this wealth of knowledge is only currently available to those few people who can gain direct access to these collections. The body of biodiversity knowledge is thus effectively withheld from use for a wide range of scientific applications, which include research, education, taxonomic study, biodiversity conservation, protected area management, disease control, and maintenance of diverse ecosystems services. Much of the early published literature is rare or has limited global distribution and is available in only a very few libraries. From a research perspective, these collections are of exceptional value because the domain of systematic biology depends – more than any other natural science – upon historic literature. The cited "half-life" (period of relevance) of natural history literature is longer than that of any other scientific domain and its "decay-rate" (rate at which it becomes irrelevant) is much slower than in other fields (cf. biotechnology). In order to positively identify a rare specimen, a working biologist may still have to consult a 100 year-old text, because that was the last time the organism was found and described.

The lack of access to the published biodiversity literature is a major obstacle to efficient research and a broad range of other applications, including education, biodiversity conservation, protected area management, disease control, and maintenance of diverse ecosystems services. BHL-Europe aims to make the biodiversity knowledge available to everybody who is interested by improving the interoperability of European biodiversity digital libraries. BHL-Europe will review and test different approaches for such libraries based on the experiences of the partners involved in the project. The consortium will establish a best practice approach and promote the adoption of standards and specifications for the large-scale implementation in a real-life context. BHL-Europe will provide a multilingual access point for search and retrieval of digital content through EUROPEANA. In addition, it will provide a robust multilingual portal with sophisticated search tools to facilitate the search for taxon-specific biodiversity information. The project will also develop operational strategies and processes for long-term preservation and sustainability of the data produced by national biodiversity digitisation programmes. BHL-Europe will generate activities to raise awareness and to ensure that the project outputs are known and used by the target users and that the proposed approach directly addresses user needs. BHL-Europe experience and best practice will be shared with the wider digital library community.

Once the collections of biodiversity literature are freely available on the Internet, this will be of great value to scientists, and also to a much wider public. Amateur naturalists (citizen scientists) who lack affiliation with major research institutions will be able to search, read, download, and print articles that were previously unavailable to them.

In addition to the scientific value of this literature, the taxonomic literature is also part of our cultural heritage. Taxonomists study and describe the organisms and biodiversity of particular areas. These areas commonly are cultural landscapes or parts of cultural landscapes. Cultural landscapes "represent the combined works of nature and of man" and reflect the evolution of human society over time in relationship to its ecological context. Thus cultural landscapes are part of the cultural heritage according to Article 1 of the UNESCO World Heritage Convention. As part of this cultural landscape, information on nature and biodiversity will be combined with information on archaeology and ethnology through the lead project of the European Digital Library Foundation, EUROPEANA. Descriptions and documentations of natural phenomena, of plants, and of animals, should be considered as part of the European cultural heritage.

The BHL is a global project and it is vital that Europe contributes its biodiversity literature to the project and that European users have access to the global BHL project. Some of the important English language literature of Europe is already part of the BHL. It is essential that the very significant amount of biodiversity literature held in other European languages – German, French, Dutch, Spanish, and others – is also integrated and becomes far more widely accessible to users. This requires a European effort to establish BHL-Europe as a dynamic component that will be both valuable in itself and also contributes to a global effort.

BHL-Europe is a project that will mobilise funding in individual EU nations in order to undertake and complete essential scanning work. Several of the partners have already indicated

that their government will promote national scanning initiatives once a large coordinating project like BHL-Europe is in place.

Benefits of BHL-Europe

The project aims to make Europe's biodiversity information, which is locked in many disparate libraries or scattered in many digital repositories, available for everybody with interest in biodiversity through a global portal (BHL) with specific biological functionality (e.g. taxonomic intelligence) and to a wide European cultural audience through EUROPEANA.

1) Enabling access to ~25 million pages of scientific literature on biodiversity to EUROPEANA will have a significant impact on the breadth and depth of European culture covered. Specifically, the action will:
· begin the fulfilment of the scientific dimension of the EUROPEANA cultural arena · demonstrate the global importance of European scientific endeavour in biological sciences · provide tools and information for the study of the history and sociology of European science · provide access to culturally-important documents – from Darwin, Linnaeus, von Humboldt, Wallace, Cuvier, Merian, etc. · provide access to many beautiful and culturally-important images – botanical drawings, zoological drawings, watercolours, etc.

2) Providing access to the biodiversity literature (images and text) using a common global portal with integrated and sophisticated search tools will produce a number of long-term benefits for the European and global biology communities. These outcomes include:
· improving the efficiency of research in the biology domain · improving access to biodiversity information for non-museum biologists · repatriation of species information in developing countries back to those countries via the Web · capacity building in the developing world

(reducing taxonomic knowledge gaps, supporting taxonomists training programmes)
· preservation of rare and fragile materials

3) The project mobilises 28 partners (23 of which may be literature providers) from 13 EU countries (of which 3 are New Member States). It is a consortium of museums, botanic gardens, universities, commercial companies, the BHL (represented by the Smithsonian Institution Library – SIL), the EDL Foundation, and other EU projects e.g. EDIT (European Distributed Institute of Taxonomy), SYNTHESYS (Synthesis of Systematic Resources), etc.

Expected results

The project will produce the following specific results:
(1) a robust biodiversity community portal with open, distributed architecture to provide multi-language access to the digital content,
(2) ~ 25 million pages of biodiversity literature from a large number of EU Member States for display through the EUROPEANA Portal,
(3) tested and validated best practice methods, standards and specifications for technology platforms, digitisation and image storage
(4) tested and validated methodology for content enrichment,
(5) tested and validated best practice workflow on implementing BHL-Europe architecture,
(6) sustainable and persistent digital curation of biodiversity heritage literature; preservation and conservation of rare and fragile material,
(7) the integration of Taxonomic Intelligence Web tools to facilitate the search for taxon-specific biodiversity information,
(8) improved efficiency of research in the biology domain; improving access to information to non-museum biologists; building public engagement, awareness and participation,
(9) permission from publishers to digitise previously published in-copyright content,

(10) a metadata repository and collection analysis tool for all the leading libraries involved, (11) strategies, plans and processes for long-term preservation and sustainability of the data produced by national biodiversity digitisation programmes as part of BHL-Europe.

Contribution to programme objectives

BHL-Europe focuses on Action 5.1. Best Practice Networks for interoperability of digital libraries with the objective of improving the interoperability of digital libraries currently held by 15 natural history museums and botanic gardens, two institutional archives, and libraries across 11 EU Member States and will be progressively extendable.

We will · establish standards-based interoperability between digitised documentation, text, metadata and the collections, thus ensuring the interoperability of the systems for o enabling the use of the multilingual and cross-cultural search and retrieval of the content o facilitating user access through EUROPEANA

· establish highly interoperable databases of the biodiversity-related content held by museums, archives and libraries · enable the actual content to be accessible and retrievable at item level by users across the European Research Area (ERA) and beyond in developing countries · take full account of the users and their needs and extend the range of users.

All these open access projects will use one of the common open access licensing approaches - the Creative Commons Share Alike (CC by-sa) licence. All digitised material will have associated metadata showing the intellectual property rights (IPR) status of the object and where it was digitised.

Multilingual and/or multicultural aspects

The biodiversity literature is available in many languages – predominately European in origin. This is particularly true for the old literature, which is the core of the current project. For instance, before 1829 the most common languages in biodiversity literature are Latin (28%), German (28%), French (14%), Swedish (9%), English (9%), Danish (6%), Dutch (3%), Italian (2%), and Spanish (1%).

BHL-Europe will have the ability to make this material available as widely as possible. The following key multilingual components are foreseen to make the literature retrievable in many European languages: Multilingual Web sites, Multi-language OCR, Multilingual indexes (Taxonomic Intelligence).

Recent target

Initial work will consist of evaluation of the technological solutions available, and assessment of the requirements of each content provider. We need to review the state-of-the-art technologies used by EUROPEANA and BHL for processing digital content, and we will include some of our experienced technology partners in that process. Simultaneously, each content provider will provide their specific requirements: what do they expect from BHL-Europe; how the digital content should be handled (technically, scientifically, and legally); what data standards and specifications they use. This technology and content-related information will be analysed in order to get a comprehensive picture of the variety of implementation approaches, their advantages, and their disadvantages.

In this phase, we also start developing the prerequisites for the management of the content, i.e. the content analysis tools and bibliographic databases. These databases will show the current position of each of the national scanning initiatives and, by merging the metadata from each partner, create a virtual taxonomic library database.

Henning Scholz and Graham Higley

Institutions cooperating in the project BHL-Europe

Museum für Naturkunde - Leibniz-Institut für Evolutions- und Biodiversitätsforschung an der Humboldt-Universität zu Berlin (MfN)
With more than 30,000,000 objects, the MfN is the largest natural history museum in Germany. The library of the MfN holds 175,000 items and currently subscribes to about 1,000 journal titles in the fields of zoology, palaeontology, and mineralogy. The museum is involved in the management of the EU I3 SYNTHESYS and will soon be leading WP7 of EDIT. In addition, the museum is involved in the development of numerous national and international projects funded by the EU, BMBF (Biota, GBIF), BMU and DFG (e.g. Graduate Research Program 503).

Natural History Museum (NHM)
The Natural History Museum in London is one of the world's greatest museums, with over 3,900,000 visitors and 13,000,000 online visitors per year. The NHM is also an international leader in the scientific study of the natural world. NHM has a strong track-record in EC funded research and training. It currently leads the €13m EU I3 Programme SYNTHESYS project, which provides access and training in 20 natural history museums and herbaria. NHM is currently in negotiation on nine FPVII

projects across a wide range of research disciplines.
In addition, NHM has led training networks and infrastructure projects in the last three frameworks; currently leads a FP6 RTN project, ORIGINS and is a partner in a further 15 FP6-funded projects. The NHM Library has the largest collection of natural history material in the world, with over 1,000,000 books (from 1469 onwards), 25,000 journal titles and 600,000 works of art. The NHM is a leading participant in the BHL and an active player in the Encyclopedia of life Project.

Project. NHM has a large number of staff members with a good background in Library Management Systems and strong IT skills including programming and database creation.

Narodní muzeum (NMP)
The National Museum is the largest museum and most distinguished leading public scientific institution in the Czech Republic with about 400 thousand visitors each year. It systematically enriches its collections including areas of natural and historical sciences from all over the world, but with particular interest to the Czech Republic. It conducts research in various fields of natural and historical sciences which it actively exhibits . It consists of five professional institutions: Natural History

Museum, Historical Museum, The Náprstek Museum of Asian, African and American Cultures, Czech Museum of Musicand National Museum Library. At present the National Museum houses almost 20 million items from the area of natural history, history, archaeology, arts, music and librarianship.

European Digital Library Foundation (EDL Foundation)

The Stichting European Digital Library (EDL Foundation) is a cross domain foundation, under Dutch law, set up for the purpose of fostering collaboration between Museums, Archives, Libraries and Audiovisual Collections in Europe. It aims to provide access to Europe's cultural heritage by facilitating formal agreement across museums, archives, audio-visual archives and libraries on how to cooperate in the delivery and sustainability of a joint portal. It also provides a legal framework for use by EU funded projects to bring their research or content into the EUROPEANA. Its current Board of Participants is made up of pan-European Associations from the 4 sectors which are able to represent and mobilise their members to contribute and form part of EU funded projects aimed at generating a European Digital Library. These are EURBICA, FIAT, ACE, EMF, ICOM Europe, CENL, CERL, LIBER, MICHAEL, the Koninklijke Bibliotheek, INA, the Bundesarchiv and the BnF.

Angewandte Informationstechnik Forschungsgesellschaft mbH (AIT)

AIT is an Austrian software and research company founded in 1979. Research work is done primarily in the field of information management (e.g. distributed databases, collection management and knowledge engineering). AIT has been involved with research projects such as MOSAIC (Museums Over States and virtual Culture; TEN-Telecom), COVAX (Contemporary Culture Virtual Archives in XML; IST-Programme), REGNET (Cultural Heritage in REGional NETworks), Media.Alp (Setting up an integrated communication platform for achieving a cultural community in the Alpine space; Interreg Alpine Space Programme) and DISMARC (DIScovering Music ARCHives; eContentplus). Within DISMARC, OAI technology and protocols are used to create a common catalogue of distributed archive metadata. At regional level, AIT is the technical provider for the DIS project, in which a virtual content catalogue for museums, archives, libraries and other institutions is being created using OAI technology and international standards (Dublin Core).

Atos Origin System Integration (ATOS)

Atos Origin is a leading international IT services provider. Atos Origin is the Worldwide Information Technology Partner for the Olympic Games. At Atos Origin, Systems Integration is not just about integrating new solutions, but includes getting the most out of legacy applications to prolong returns from existing IT investment. Successfully combining new solutions with established ones can transform the complete enterprise architecture into a single, seamless business system. Our extensive experience in integrating people, processes and technologies enables us to design, build and operate practical and robust solutions.

Freie Universität Berlin (FUB-BGBM)

The Botanic Garden and Botanical Museum Berlin-Dahlem (BGBM), with its extensive scientific collections of herbarium specimens (about 3.5 million) and living plants, is a centre of biodiversity research in Europe. It houses the most complete botanical library in Germany. The library holds a wide range of literature on plants from all over the world, in all printed languages and from five centuries, among them many precious and very rare books. BGBM has a separate department of Biodiversity Informatics with, at present, 20 staff members. Focal points of research and development activities are taxonomic information systems and networking of distributed primary biodiversity information.

Georg-August-Universität Göttingen Stiftung Öffentlichen Rechts (UGOE)

The EZOOLO/AnimalBase project is located at the Georg-August-Universität of Göttingen. It was initiated as a joint venture of Göttingen University Library (SUB) and the Zoological Institute of the university to provide free access to digitised versions of all taxonomically relevant early zoological work. In the first step (2003-2005) more than 100,000 pages were digitised from the earliest beginnings of scientific zoology in the 1550s until the year 1770

and 10,000 animal names were extracted and transferred to AnimalBase. In the second step, from 2008 onwards, literature until the 1820s will be covered, with approximately 50,000 animal names being extracted.

Naturhistorisches Museum Wien (NHMW)

The collections with more than 30 million specimens, including hundreds of thousands of types, are the basis for any taxonomic work. AT-TAF (SYNTHESYS) is one of the "first addresses" for taxonomic work. Additional a library with many historically important volumes is available. The library with c. 6000 scientific journals and tens of thousands of books complement the National Library of Austria and the University's libraries nearby.

Land Oberösterreich (Oberoesterreichisches Landesmuseen) (LANDOE)

The Biology Centre in Linz-Dornach, with more than 6 million objects, represents the 2nd largest in Austria. It currently publishes the series Stapfia and Denisia and three other journals. It holds the biodiversity database ZOBODAT, founded in 1972 as ZOODAT. The database today includes more than 3.3 million records, literature citations (more than 33,000), OCR scanned books (~150,000 pages) and until now bibliographies from about 4,000 biologists.

Hungarian Natural History Museum (HNHM)

HNHM holds more than 10 million natural history items. HNHM Library contains more than 300,000 volumes. The catalogue of the library is available on the Internet. The HNHM has published several natural history journals and books during its 200 years history. During recent years these have become available via the internet but there is a strong commitment by HNHM to digitise and provide free access to its own journals and books.

Museum and Institute of Zoology, Polish Academy of Sciences (MIZPAS)

The library collection is of national importance including literature on zoology, especially systematic and zoogeography, entomology and ornithology. Recently it has increased its collections of publications concerning molecular biology. At present the Library comprises 243,271 volumes and 5,378 archival items.

University of Copenhagen (The Natural History Museum of Denmark) (UCPH)

The museum holds an estimated 12 million specimens of animals, plants, books, archives, fossils, minerals, and other natural history related items. Part of the museum are three libraries, botanical, zoological and geological. The libraries hold more than 250,000 bibliographic entities (books, journals, reprints). The museum hosts the GBIF and the Danish GBIF node.

Stichting Nationaal Natuurhistorisch Museum Naturalis (NAT)

Naturalis was founded in 1820 and much of its collection dates back to the 19th and 20th century. Naturalis has a staff of about 160 people, which includes scientists, collection managers, exhibition designers, information officers, educators, etc. The collections of zoological, palaeontological and geological objects are estimated to total about 12 million objects.

With reference to the BHL-Europe project, Naturalis can rely on a strong and innovative department of information services, backed by natural history collections and archives which cover nearly 200 years of research and collecting.

National Botanic Garden of Belgium (NBGB)

NBGB is a 'complete' botanical garden, integrating a living collection ('Hortus'; 18,000 species in cultivation) and a large museum ('Herbarium'; more than 3 million plants incl. fungi). The library holds 50,000 monographs including 2,500 valuable historical books, 5,000 periodicals and 25,000 reprints. Historical literature on Central African flora was kept in NBGB since 1890. Data repatriation to partners both in Africa and Latin America is a priority for NBGB. Drawings and colour paintings of flowering plants and fungi were digitised and linked to specimens kept in the BR herbarium. NBGB is an institutional member of CETAF and is active in the EC projects ENBI, EDIT and

SYNTHESYS. It also collaborates as a training provider for GTI (Global Taxonomy Initiative) capacity building activities with developing countries.

Royal Museum for Central Africa (RMCA)

RMCA is a multidisciplinary institution with a special focus on Sub-Saharan Africa. The museum manages collections of about 10 milli on specimens of animals and 56,000 wood specimens from 13,600 different botanical species. The RMCA maintains an extensive library on African biodiversity, including the top scientific journals, but also a unique collection of rare, old colonial publications. The institution has about 1.2 km of Archives. Information on biodiversity is repatriated to African partners, suitable training in taxonomy is also provided. RMCA is an institutional member of TDWG and is involved in initiatives such as GBIF, ENBI, EDIT, CETAF and SYNTHESYS. It collaborates in GTI capacity building activities with developing countries.

Royal Belgian Institute of Natural Sciences (RBINS)

The RBINS houses a diverse and exceptionally rich collection, comprising about 37,000,000 specimens. The scientific library of the RBINS is the biggest documentary resource of natural history in Belgium. It offers a vast range of books (695,368 volumes) and has very specialised, often unique scientific magazines. Its catalogue is available online. RBINS has begun digitisation of the library catalogue and more than 185,000 titles are online.

Bibliothèque nationale de France (BnF)

BnF is one of the largest public and research libraries in the world and holds more than 50,000 monographs and around 3,000 titles of periodicals published in the field of natural sciences between 1801 and 1920. The BnF offers access to its digital library Gallica, www.gallica.bnf.fr, created through the library's commitment to digitisation of selected items of its collections. In spring 2008, the BnF launched a new version of Gallica with new, modern functionalities, drawing upon the most recent Web 2.0 experience. At a European level, the BnF is a founding member of The European Library consortium and is involved in the TELplus and IMPACT projects and is also a member of the EDLnet thematic network.

Museum national d'histoire naturelle (MNHN)

The main activities of MNHN are research, education and training, enrichment of its collections (around 68 million specimens), providing expertise and diffusion of scientific knowledge. MNHN was, and still is, heavily involved in various European programmes related to biodiversity: The European Topic Centre on Biological Diversity of the European Environmental

Agency; the French GBIF, GTI and CBD nodes, the French Focal Point for the EU-funded BioCASE project; a leading partner in several past and on-going EU projects, including ParSyst, ColParSyst, Fauna Europaea, ENBI and EuroCAT/Sp2000 Europa, SYNTHESYS, EDIT, MarBEF, EUMon and PESI. The central Library Department holds the world's third largest collection of literature, original drawings and manuscripts relating to natural history. The print collections include 20,000 periodical titles and 600,000 books. Digitisation plans are already on-going in full collaboration with BnF.

Consejo Superior de Investigaciones Científicas (CSIC)

CSIC is the largest National Research Institution in Spain. The participant institute, Museo Nacional de Ciencias Naturales (MNCN) houses the biggest natural history collections, library and archives in Spain. The relevant national project for the proposal lead by MNCN is Fauna Ibérica. MNCN is a founding member of CETAF, partner of SYNTHESYS, EDIT, LifeWatch and coordinates the Spanish GBIF node. The public library of MNCN contains more than 62,000 volumes and more than 6,400 scientific journals as well as access to more than 9,000 electronic journals.

Università degli Studi di Firenze (Museo di Storia Naturale) (MSN)

With more than 10 million specimens, it is the most important natural history museum in Italy. The Museum houses specimens of extraordinary scientific and natural history value: XVI century herbaria, valuable XVIII century waxes, fossil elephant skeletons, brightly coloured butterflies, huge tourmaline crystals, Aztec artefacts, imposing wooden sculptures and the world's largest flower. A lot of ancient and rare

books are also preserved in the Library (Biblioteca di Scienze).

Royal Botanic Garden Edinburgh (RBGE)

RBGE is an internationally renowned centre for botanical research and conservation work and holds one of the largest collections of living plant species in Europe, together with large collections of preserved plant and fungal material. It has one of the most important botanical libraries in the UK.

Species 2000 (Sp2000)

Species 2000 is a Network organisation which is creating an index of the world's known organisms. The programme reached production scale as an EC scientific infrastructure under the FP5 EuroCAT project. Its Catalogue of Life is a global service (www.catalogueoflife.org) recognised by the UN Convention on Biological Diversity, and presently comprises a synonymic species checklist of 1.1 million plants, animals, fungi and micro-organisms, about 2.5 million names and a comprehensive taxonomic hierarchy. It contributes content and the taxonomic hierarchy used by the taxonomic intelligent tool.

John Wiley & Sons limited (Wiley)

Wiley is a global publishing company founded in 1807 which markets its products to professionals and consumers, students and instructors in higher education plus researchers and practitioners in scientific, technical, medical and scholarly fields. Through the 2007 acquisition of Blackwell Publishing, Wiley has gained the Blackwell Synergy platform, home to over 850 Blackwell journals. Also during 2007, the company completed an initiative to digitise its entire historical journal holdings, making 8.2 million pages of content, dating back to 1799, available on Wiley InterScience.

Smithsonian Institution (SIL)

The Smithsonian Natural History Museum is dedicated to inspiring curiosity, discovery and learning about the natural world through its unparalleled research, collections, libraries, exhibitions and education programs. At the centre of the Museum's exhibition and research programs are its expertly documented collections of more than 125 million natural science specimens and cultural artefacts. Over 3.5 million specimens are out on loan each year; over 15,000 visitor days are spent in the collections; and there are almost 600,000 additional visits to collection data bases available on the Web. The BHL is led from SIL.

Missouri Botanical Garden (MOBOT)

The mission of MOBOT is to discover and share knowledge about plants and their envi-ent in order to preserve and enrich life. Founded by Henry Shaw and opened to the public in 1859, the Garden is a National Historic Landmark and widely considered one of the top three botanical gardens in the world. MOBOT is a founding member of the BHL and is supporting the development of the system's infrastructure, application layers and interfaces.

Helsingin yliopisto, University of Helsinki, Viikki Science Library (UH-Viikki)

Viikki Science Library is the bioscience campus library of the University of Helsinki. The library is the largest resource library in Finland in bioscience, agriculture and forestry, pharmacy and veterinary medicine. The library also serves the Finnish Museum of Natural History. The library has digitised old Finnish forestry literature which is made available in the DViikki open research archive. The library has printed collections of roughly 12,000 shelf meters.

Humboldt-Universität zu Berlin (UBER)

The University Library and the Computer and Media Services of UBER have carried out several projects within the fields of e-publishing, digital preservation and the digital library. Both institutions have been involved in several EU projects during the last 5 years: Open Archives Forum (EU-IST Programme); reUSE (eContentplus project); Digitisation-on-Demand (eTEN project).

Articles

The Biodiversity Heritage Library

Thomas Garnett, Smithsonian Institution, Washington

The Biodiversity Heritage Library (BHL), the digitization component of the Encyclopedia of Life, is a consortium of 12 major natural history museum libraries, botanical libraries, and research institutions organized to digitize, serve, and preserve the legacy literature of biodiversity. The European Commission's eContentPlus program has recently funded the BHL-Europe project, with 28 institutions, to assemble the European language literature. In addition, a Memorandum Of Understanding with the Chinese Academy of Sciences will be signed soon for a Chinese BHL. Negotiations are being pursued with the Atlas of Living Australia and other projects to join the BHL consortium. These projects will work together to share content, protocols, services, and digital preservation practices. These welcome developments are leading to a new model for a global BHL sufficient to meet the challenges of assembling, structuring, and preserving the published record of biodiversity as a shared heritage of all humanity. The BHL is becoming international. It is moving from a US/UK focused project to a global presence that mirrors the global nature of biodiversity research and related disciplines.

Prior to digitization, the resources housed within each BHL institution have existed in isolation, available only to those with physical access to the collections. These collections are of exceptional value because the domain of systematic biology depends – more than any other science – upon historic literature. Consequently, the relative isolation of these collections presented an antiquated obstacle to further biodiversity investigation. This problem is particularly acute for the developing countries that are home to the majority of the world's biodiversity.

Legacy Taxonomic Literature

In 2003 the National Academy of Science issued a report by J.H. Reichman and Paul F. Uhlir titled, A Contractually Reconsidered Research Commons for Scientific Data in a Highly Protectionist Intellectual Property Environment. In this report the authors distinguish big science supported by governments and with norms and enforcement for open access with small science about which they wrote, There are few government-controlled, public domain data centers in this type of research. The data are thus disaggregated components of an incipient network that is only as effective as the individual transactions that put it together. Openness and sharing are not ignored, but they are not necessarily dominant, either. These values must compete with strategic considerations of self-interest, secrecy, and the logic of mutually beneficial exchange. In small science, what occurs is a delicate process of negotiation, in which data are traded as the result of informal compromises between private and public interests that are worked out on an ad hoc and continual basis. Small science thus depends on the flow of disaggregated data through many different hands, all of which collectively construct a fragile chain of semi-contractual relations in which secrecy and disclosure are pitted against a common need for access and use of these resources. In this sense, big science projects are more likely to be subject to formal data access regulations while small science research is more emblematic of informal data exchange practices.

Until recently this characterization of biodiversity sciences as a "disaggregated components of an incipient network" had some truth. However, in recent years a global network of biodiversity database, services, web sites, inventories, and research tools has emerged in response to the ecological crises and new web-enabled technologies. For example at the recent eBiosphere Conference held in London over 130 current

projects showcased ranging from international compendia such as the Encyclopaedia of Life (EOL) and the Global Biodiversity Information Facility (GBIF) to soil biodiversity surveys in Algeria, analysis of species density in Malaysian forests, and phytoplankton carbon sequestration analysis were viewed by 500 key personnel engaged in biodiversity informatics. The overwhelming focus was on weaving together a seamless web of biodiversity information in an effective way for scientists and researchers. A key component of this emerging network of content and services has been the published literature of biodiversity, which until recently has been the prime method of recording and disseminating the results of research since the time of Linnaeus.

The partner libraries collectively hold a substantial part of the world's published knowledge on biological diversity. Yet, this wealth of knowledge is available only to those few who can gain direct access to these collections. This body of biodiversity knowledge, in its current form, is unavailable to a broad range of applications including: research, education, taxonomic study, biodiversity conservation, protected area management, disease control, and maintenance of diverse ecosystems services.

Much of this published literature is rare or has limited global distribution. From a scholarly perspective, these collections are of exceptional value because the domain of systematic biology depends -- more than any other science -- upon historic literature. The "cited half-life" of natural history literature is longer than that of any other scientific domain. The so-called "decay-rate" of this literature is much slower than in other fields such as biotechnology. Mass digitization projects at large research libraries lack the discipline-specific focus of the Biodiversity Heritage Library Project. These other projects will fail to capture significant elements of legacy taxonomic literature.

Open Access

The BHL is striving to establish a major corpus of digitized publications on the Web drawn from the historical biodiversity literature. This material will be available for open access and responsible use as a part of a global Biodiversity Commons. We will work with the global taxonomic community, rights holders, and other interested parties to ensure that this legacy literature is available to all. "Science has no borders".

The BHL must be a multi-institutional project because no single natural history museum or botanical garden library holds the complete corpus of legacy literature, even within the individual sub-domains of taxonomy. However, taken together, the proposed consortium of collections represents a uniquely comprehensive assemblage of literature.

The BHL will immediately provide content for multiple bioinformatics initiatives and research. For the first time in history, the core of our natural history museum libraries and botanical garden library collections will be available to a truly global audience. Web-based access to these collections will provide a substantial benefit to all researchers, including those living and working in the developing world. BHL's open access and open source

approach will push the frontiers of biodiversity research.

Complementing the BHL Portal (http://www.biodiversitylibrary.org/), the BHL initiative will release CiteBank, a related repository of scientific citations and community-vetted bibliographies with services linking to other biodiversity projects. CiteBank allows users to upload and share bibliographies containing material related to their specific interests and upload files associated with these bibliographies, including PDFs of the articles and links to the books containing the articles within the BHL portal. As such, CiteBank is a crowd-sourced, user-dependant service.

CiteBank also serves as an interface through which users can network and collaborate, forming groups related to specific interests and subjects. A full release of this component is planned for October 2009. Please review latest updates at the BHL Blog http://biodiversitylibrary.blogspot.com

BHL Consortium Membership

Participating US/UK institutions:
- American Museum of Natural History (New York, NY)
- The Field Museum (Chicago, IL)
- Harvard University Botany Libraries (Cambridge, MA)
- Harvard University, Ernst Mayr Library of the Museum of Comparative Zoology (Cambridge, MA)
- Marine Biological Laboratory / Woods Hole Oceanographic Institution (Woods Hole, MA)
- Missouri Botanical Garden (St. Louis, MO)
- Natural History Museum (London, UK)
- The New York Botanical Garden (New York, NY)
- Royal Botanic Gardens, Kew (Richmond, UK)
- Smithsonian Institution Libraries (Washington, DC)
- Academy of Natural Science (Philadelphia, PA)
- California Academy of Science (San Francisco, CA)

Contributing Members (Libraries who have generously offered to share portions of their digitized collections with the BHL)

- The University Library of the University of Illinois at Urbana-Champaign

UIUC has agreed to participate in the BHL as a contributing member by digitizing important biodiversity journals originating in the state of Illinois.

- The California Digital Library has agreed to allow the BHL to serve selections of the texts from the University of California Libraries with significant biodiversity content.

BHL-Europe Membership

- Museum für Naturkunde - Leibniz-Institut für Evolutions- und Biodiversitätsforschung an der Humboldt-Universität zu Berlin
- Natural History Museum, UK
- Narodni muzeum NMP
- Angewandte Informationstechnik Forschungsgesellschaft mbH (AIT)
- ATOS Origin Integration
- Freie Universität Berlin
- Georg-August-Universität Göttingen Stiftung Öffentlichen Rechts
- Land Oberösterreich
- Naturhistorisches Museum Wien
- Hungarian Natural History Museum
- Museum and Institute of Zoology, Polish Academy of Sciences
- University of Copenhagen
- Stichting Nationaal Natuurhistorisch Museum, Naturalis
- National Botanic Garden of Belgium
- Royal Museum for Central Africa
- Royal Belgian Institute of Natural Sciences
- Bibliothèque nationale de France
- Museum national d'histoire naturelle
- Consejo Superior de Investigaciones Cientificas
- Università degli Studi di Firenze
- Royal Botanic Garden, Edinburgh
- Species 2000
- John Wiley & Sons limited
- Helsingin yliopisto UH-Viikki

Chinese Academy of Science (to come soon).

New Members

In May 2009, two new members were accepted into the Biodiversity Heritage Library:

· The Academy of Natural Sciences of Philadelphia
· The California Academy of Sciences (San Francisco)

These two important institutions will greatly contribute to the strength of the BHL.

For nearly 200 years, The Academy of Natural Sciences of Philadelphia has built one of the finest natural history libraries in the world. The Library has grown from five books and two maps in 1812 to its present size of over 68,000 printed titles (numbering over 200,000 volumes) and over 250,000 manuscripts and related pictorial materials. The manuscript collections include the Academy's archives as well as correspondence, photographs, journals, field notes, and original illustrations by and about American scientists from the first half of the eighteenth century to the present. It is the library of record for early accounts of American scientific expeditions.

The Academy Library, the Ewell Sale Stewart Library, is notable in the Americas for its holdings of historical works in every discipline of the natural sciences; the collection has particular strength in the history of science, evolution, early ecology, systematics, marine and freshwater biology, and geology. It was an early adopter of digitization as a method to make its holdings widely accessible. It began to scan digital images from books and archival materials and publish them on the web in 1999. Two early projects were funded by the IMLS in the form of a Leadership Grant, and one by the Getty Grant Program.

Established in 1853, the California Academy of Sciences is the third oldest institution of its kind in the U.S. and the only one in the nation with a major research center, aquarium, planetarium, and natural history museum under one roof. All programs and exhibits at the Academy embrace its mission to explore, explain, and protect the natural world, focusing on the evolution of life, its diversity, and its sustainability. Scientific research is a key part of the Academy's mission and this work is supported by a collection of 20 million specimens, 60 research scientists, and a library with more than 230,000 volumes and 1200 current serial titles. Examples of current biodiversity research include the work being done by Brian Fisher on the ants of Madagascar, Jack Dumbacher on the birds of Papua New Guinea, Peter Roopnarine on the evolution of communities through the fossil record, and Healy Hamilton on the effects of climate change on species distributions and conservation planning efforts.

Below is a map showing the current locations of the global Biodiversity Heritage Library.

Digitization in the Czech Republic

Richard Šípek, National Museum, Prague

The first large scale digitization project in the Czech Republic has already started and the first CD with a scanned selection from medieval manuscripts was burnt between the years 1992 and 1993 as a contribution to the UNESCO initiative Memory of the World. The Czech National Program adopted the name Memoriae Mundi Series Bohemica and it was coordinated by The National Library of the Czech Republic.

The project was run in cooperation with a private company, Albertina icome Praha Ltd. (since 1995 known as AiP Beroun Ltd.) which provided (and still provides) technical services, i. e. digitization of books and ensuring accessibility of the digital images, which were at that time still limited, on CD-ROMs. Albertina icome Praha Ltd. developed its own scanning device which fortunately has minimal effect on the scanned materials as they were almost all exclusively of high historical and artistic value.

The first manuscripts scanned in extenso were the Antiphonarium Sedlecense and the Chronicon
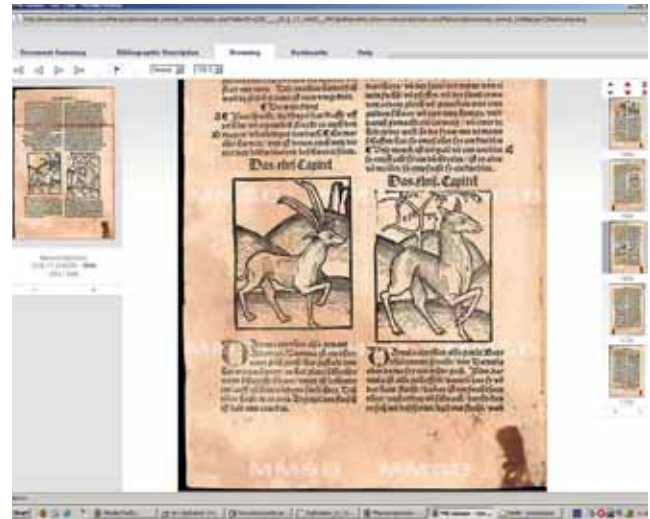
„The Webpage of „MS Viewer", the interface for display of digitized documents in Manuscriptorium. Digitized old print: Gart der Gesundheit, Strassburg, 1529."

Constantiniense in 1994 and 1995. Technological and description standards were developed during the following years, simultaneously with ongoing digitization of manuscripts. (By the year 2000, there were 39 fully scanned volumes on 48 CD-ROMS available in the study room of The National Library.) The standards applied by The National Library and AiP Beroun Ltd. were those recommended by the organization UNESCO to the Memory of the World contributors in 1999 and six years later, The National Library of the Czech Republic was awarded the UNESCO Jikji Prize in Korea.

In the year 2000, The Ministry of Culture of the Czech Republic began a subprogram within the framework of the grant program, Public Information Services in Libraries (VISK 6), to encourage digitization in as many institutions as possible, including libraries, museums, galleries and other similar establishments. The grant program VISK 6 still continues on an annual basis covering 70% of digitization expenses; 30% are covered by the financial resources of each institution in question.

A new project, Kramerius, was started at about the same time and focused on digitization of periodicals and other types of more recently printed books, dated mostly from the 19th century and

„The Webpage of „MS Viewer", the interface for display of digitized documents in Manuscriptorium. Digitized medieval manuscript: Jena Codex, Bohemia, 1490-1510."

„Homepage of the project Kramerius."

endangered by the acidity of the paper. The initiative was supported by the new subprogram of Public Information Services in Libraries, VISK 7, intended for digitization of those documents printed on acid paper.

Two online accessible digital libraries began three years later; Manuscriptorium (http://www.manuscriptorium.com), making accessible images digitized within the framework of the program Memoriae Mundi Series Bohemica and Kramerius (http://kramerius.nkp.cz), providing availability of the image data created by the a ponymous project. A year earlier, all the metadata attached to the digitized documents were converted to XML format which is still in use today.

The XML metadata description format MASTER was utilised. A new stage of the digital library Manuscriptorium was launched on December 1st 2007, represented by the project ENRICH which runs within the framework of the European digitization initiative eContentPlus. It aims to make Manuscriptorium open to other domestic, as well as foreign, digital libraries and with digitization initiatives functioning such as metalibrary and making accessible digital images created by the other institutions via the Manuscriptorium search mask and display interface which is optimized for contributors (description in the vernacular language of the contributing institution, different character set etc.). The image data may be stored in the home server of Manuscriptorium but if the HTML address/path to their actual storage is defined in the document metadata description, it is not mandatory. All the metadata created in MASTER were converted to the new format, TEI P5. At present , the digital library Manuscriptorium contains about 2 800 digitized documents (mostly manuscripts and old prints), i. e. more than 1 000 000 images. Moreover, 100 000 catalogue records are also available. In the Czech Republic there are other digitization initiatives and projects in addition to Manuscriptorium and Kramerius which have been launched and are run by libraries, museums and other relevant institutions where the digitization technology and services are delivered mostly by several private companies (Elsyst Ingeneering Vyškov, Microna, Ampaco). In a few cases, the home institution or library has the necessary digitization equipment (The Library of The Czech Academy of Sciences, The National Library of the Czech Republic). The scanned images are usually made accessible via web interface adopted, or tailor-made, for each institution.

South Bohemian Scientific Library maintains a digital collection of more than 1 000 historical maps (http://archiv.cbvk.cz/kramerius/Welcome.
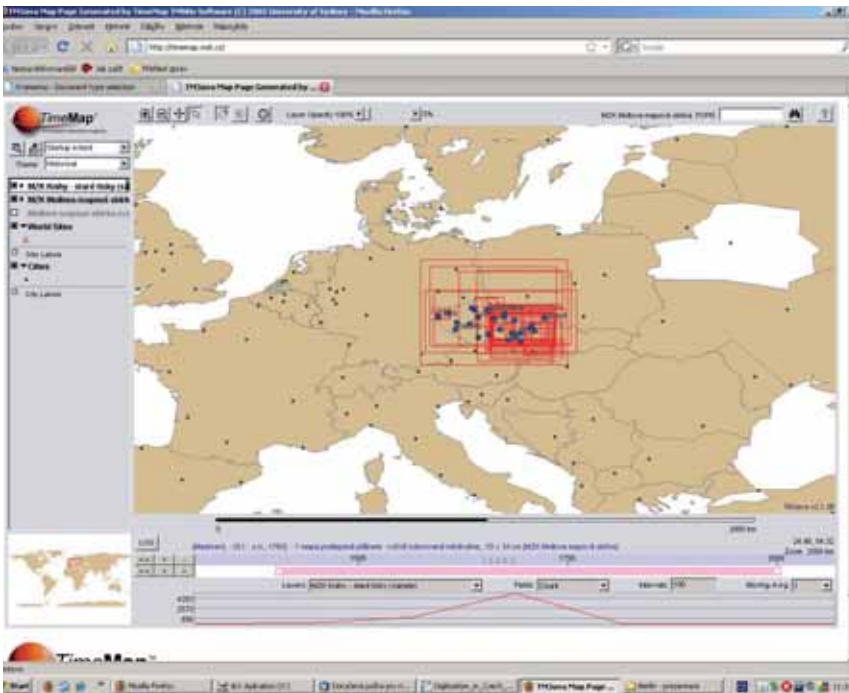
„Search and display interface of the digitized maps project run by the Moravian University."
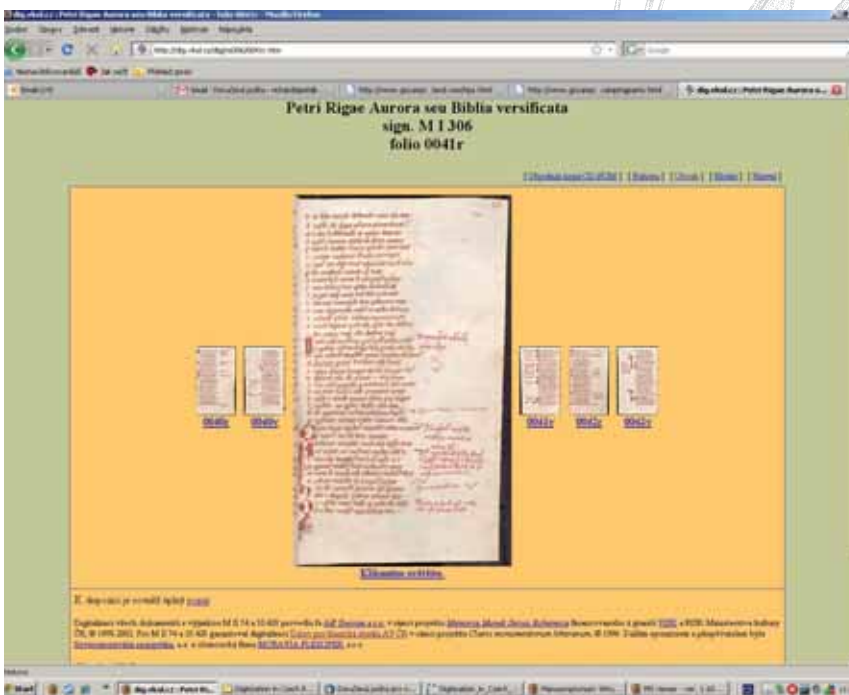
„Medieval manuscript Petri Rigae Aurora seu Biblia versificata. 1200-1250, digitized by AiP Beroun, accessible on the webpage of The Scientific Library Olomouc"

do?lang=en), 558 scanned maps from the collections of The Scientific Library Olomouc are accessible via the library's own homepage (http://mapy. vkol.cz/). The same library provides access via its homepage to 83 digitized manuscripts and incunabulas and 7 old prints from its collections (http://dig. vkol.cz/). Extremely interesting is the recently launched search and display interface providing access to the digitized maps of The Moravian Library, in particular for its sophisticated search system (http://timemap. mzk.cz/).

Digitization activities were from the very beginning focussed, naturally, on the scanning of historical book collections, prioritizing medieval manuscripts, incunabulas and old prints. However, this focus of attention changed comparatively quickly, and happily, turned to more recent books which are, nevertheless, in many respects more endangered by decay due to the fragility of the acid paper. As it was shown, the most important digitization projects in the Czech Republic are Manuscriptorium for the historical collections and Kramerius for the more recent ones. Nonetheless, the other digitized collections, displayed on the home pages of particular libraries, should not be omitted.

BHL Digitisation at the Natural History Museum, London.

Bernard Scaife, Digital Library Manager, Natural History Museum, London 2009-07-23

The Natural History Museum, London (NHML) has been involved in BHL since day one. Our involvement with the planning led to us being tasked with the creation of a Serials Mashup, designed to analyse our holdings, identify duplicate materials and divide the scanning responsibilities amongst the participating institutions. This is still running today, although hopefully will be greatly enhanced thanks to the BHL Europe project.

For our own contribution of content at NHML, we have been using a single Scribe scanner which is leased from the Internet Archive, the BHL's scanning partner. This is a manually operated machine, comprising a cradle operated by a foot pedal and two mounted Canon cameras. The first step is to retrieve the catalogue record associated with the item. This is achieved using Z39.50. The book is then scanned from cover to cover (two shots per page spread). An additional step to insert foldouts using a separate camera setup was introduced about eight months after we started scanning, reflecting the fact that a large quantity of our material includes foldouts of differing sizes.

Following the scanning, a Quality Assurance process is undertaken which is based upon ANSI z1.4 1993 Table 1, General Level 2. At this stage, physical page numbers are associated to the images in batch and an assertion made regarding the page type (e.g. cover, page).

The next stage is to ftp (or "pipe" in IA parlance) the book to the Internet Archive server. The raw files are further post processed to create surrogates in other formats such as pdf and the content is OCR'd to make it searchable. The package is stored on the IA servers and public access made available.

There is a whole other process whereby the content gets ingested to the BHL's website, but that will have to be the subject of another article.

Returning to our local processes at the NHML, there is a lot of preparation before the book ends up on a trolley ready to be scanned by the IA. First, we have to decide what we want to scan. This was heavily influenced by copyright issues in the early days. UK copyright legislation is quite restrictive, and we underwent a risk management process to conclude what date would be a sensible cut-off before which it was safe to digitise. The conclusion was pre-1860, so the majority of our materials are currently from that period. The other decision was to start with serials. The reason for this is that in many of our catalogues, the energy expended in de-duplicating serial titles was well worth it as de-duplicating one title could reveal many shelves of volumes of material which could be scanned, due to the common practice of not cataloguing serials at volume, but at title level.

The next stage is to ensure that it hasn't already been scanned (the Serials mashup is used for this purpose) and to place notice of our intention to scan alongside the title so other institutions will know not to scan it themselves.

There are some condition and size tests that have to be applied before the next stage can happen, but on the whole, this has not proved to be problematic despite the age of the material we are scanning.
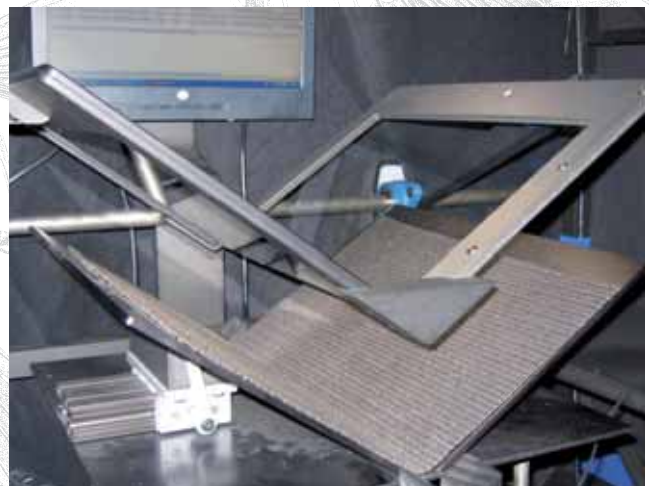
Next, the books are prepared for scanning using our in-house Scanning Management System (SMS). In technical terms, this is an addition to the serials mashup functionality that we use. Practically, the workflow is to ensure the title is on the system, confirmed as not already scanned, and is copyright clear. Then, a packing list is made. A packing list is a list of books to be supplied on one trolley. We aim to supply between 40 and 50 volumes per trolley. When we are building a packing list, we take the opportunity to not only link the volume to the correct catalogue system number (which the IA use to retrieve the record), but also to add some metadata about the volume number and date so that each volume is distinguishable from its neighbour on the website.

The trolley is then delivered to the IA. Using the same SMS, they simply enter the catalogue system number from the slip of paper in the book and this information populates their form with the bibliographic and volume level metadata. This process is called "Wonderfetch".

See below for some photos of the Scribe machine.

https://bhl.wikispaces.com/NHM+Photos