ECP-2008-DILI-518001

BHL-Europe

Annual Report

1 May 2009 – 30 April 2010

www.bhl-europe.eu

Deliverable number/name	*D1.3*
Dissemination level	*Public*
Delivery date	*22 July 2010*
Status	*Final*
Author(s)	*Henning Scholz*

e*Content*plus

This project is funded under the *e*Content*plus* programme[1],
a multiannual Community programme to make digital content in Europe more accessible, usable and exploitable.

[1] OJ L 79, 24.3.2005, p. 1.

0 Document History

0.1 Contributors

Person	Partner
Henning Scholz	MfN

0.2 Revision History

Revision Date	Author	Version	Change Reference & Summary
07 Apr 2010	Henning Scholz	0.1	Initial draft version for external review
19 Apr 2010	Martin Gordon	0.2	External review
22 Apr 2010	Henning Scholz	0.3	Revised draft taking external review into account
28 Apr 2010	Henning Scholz	1.0	Final version
22 July 2010	Henning Scholz	2.0	Revised version taking comments during the project review into account.

0.3 Reviewers

This document requires the following reviews and approvals.

Name	Position	Date	Version
Martin Gordon	Project Manager, RBB	19 April 2010	0.1
Martin Gordon	Project Manager, RBB	23 April 2010	0.3
BHL-Europe consortium		26 April 2010	0.3

0.4 Distribution

This document has been distributed to:

Group	Date of issue	Version
Reviewer and BHL-Europe consortium	16 April 2010	0.1
Reviewer and BHL-Europe consortium	22 April 2010	0.3
BHL-Europe consortium	3 May 2010	1.0
EC and Reviewers	3 May 2010	1.0
BHL-Europe consortium	23 July 2010	2.0

1 Table of contents

2 Project Objectives

2.1 Background, problem addressed and project objectives

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 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.

Currently, numerous natural history institutions in the world are digitising this biodiversity literature in an open access manner. Since 2007, ten major biodiversity libraries have collaborated in making the biodiversity knowledge accessible on an open access Creative Commons basis to a wide spectrum of end-users via the Biodiversity Heritage Library (BHL) project. Two European institutions are participating in the BHL project: Natural History Museum (London, UK) and Royal Botanic Gardens, Kew (Richmond, UK). The BHL has scanned more than 40,000 volumes to date and the corpus of digital literature is continuously increasing.

In addition to BHL, a large number of projects are digitising biodiversity material in numerous institutions across the EU to make access more open, but the corpus will still be seriously fragmented. Two examples of European projects are Gallica www.gallica.bnf.fr and AnimalBase www.animalbase.uni-goettingen.de. All these scattered and disparate projects do not use common standards or interfaces and are not interoperable.

In contrast to BHL, BHL-Europe is not a digitisation project and is not funded to digitise literature. Having in mind the large but disparate corpus of digital biodiversity content in European countries, BHL-Europe aims to make biodiversity knowledge available to everybody who is interested by improving the interoperability of European biodiversity digital libraries. BHL-Europe reviews and tests different approaches for the establishment and management of digital multilingual biodiversity libraries based on the experiences of the partners involved in the project. The consortium has established a best practice approach and promotes the adoption of standards and specifications for the large-scale implementation in a real-life context. BHL-Europe focuses upon providing a multilingual access point for search and retrieval of digital content through Europeana.

In addition, it will establish 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 is already generating

activities to raise awareness and to ensure that the project outputs are known and used by the target users, and also that the proposed approach directly addresses user needs. BHL-Europe experience and best practice will be shared with the wider digital library community. BHL-Europe facilitates the open access to taxonomic literature for a large number of target users including the general public. It will also facilitate and enable the initiation of scanning initiatives in European countries not yet involved in digitisation programmes and will improve the infrastructure for digital libraries in all EU countries. This also includes the negotiation with Rights Holders to enable access to in-copyright content.

3 Project Results/Achievements

3.1 Performance Indicators

Indicators	Expected vs Actual					
	Year 1 (exp.)	**Year 1 (act.)**	**Year 2 (exp.)**	**Year 2 (act.)**	**Year 3 (exp.)**	**Year 3 (act.)**
Number of total accessible pages of biodiversity literature (in million)[1]	17	17.4	*21*		*25*	
Number of pages of biodiversity literature from European content providers to BHL-Europe (in million)	3.4	3.4	*4.1*		*4.8*	
Percentage of literature available through Europeana[2]	20%	0%	*50%*		*100%*	
Amount of metadata sets to be imported into the GRIB (in million)	0.15	0.15	*2.5*		*4.0*	
Number of interconnected repositories	7	14	*20*		*30*	
Number of content providers	20	21	*25*		*30*	
Number of portal languages[3]	1	1	*7*		*12*	
Page views through BHL Portal (in million) [3]	1.0	3.7	*2.0*		*2.5*	
Page views through BHL-Europe Portal	0.0	0.0	*0.5*		*2.5*	
Page views through Europeana Portal (in million) [2]	0.25	0	*1.0*		*3.0*	
Case studies of successful usage of the material by non-scientists	5	5	*10*		*15*	
Agreements with Rights Holders / Publishers	2	7	*4*		*6*	

[1] BHL currently has 29 million pages on display through www.biodiversitylibrary.org/. However, a significant amount of that content is currently not part of BHL-Europe. This content is from non-BHL partner libraries. Only the content of the BHL partner libraries and the BHL-Europe content providers is given herein.

[2] As the data harmonisation and harvest process was not finished before the end of the reporting period (30 April 2010), no BHL-Europe content is accessible via Europeana. The harvest process is finished 10 June 2010 so the first BHL-Europe data is available from that day via http://www.europeana.eu.

[3] BHL-Europe has no portal yet for accessing the biodiversity literature. This is in full accordance with the project plan. The first BHL-Europe prototype will be available in October 2010 (see also D3.5 of BHL-Europe for more details on the prototype). To measure usage of the digital biodiversity content, we refer to the BHL Portal statistics. Also the current portal language is the one implemented in the existing BHL Portal www.biodiversitylibrary.org/, as this is the only portal currently available for users.

3.2 Deliverables List

Deliverable No	Deliverable title	Delivery due date	Actual date of delivery
D2.1	Catalogue of content holder requirements (quality, quantity, accessibility, standards and specifications of content and metadata)	M 3	M 4, Aug 09
D3.1	Deliver composition of Technology Management Board and initial meeting	M 3	M 4, Aug 09
D5.1	Web site, including multimedia presentation	M 3	M 4, Aug 09
D5.2	BHL-Europe newsletter and mailing list	M 3	M 4, Aug 09
D5.3	Database of relevant conferences/events and ownership for BHL-Europe presentations	M 3	M 4, Aug 09
D5.4	Deliver composition of Communications Working Group and 1st dissemination plan	M 3	M 4, Aug 09
D1.1	Progress Report 1	M 6	M 7, Nov 09
D3.2	Document agreed standards, best practice & system components	M 6	M 7, Nov 09
D3.3	Plan for managing interoperability issues, data harmonisation and the integration of the content into BHL-Europe, EUROPEANA and the BHL	M 6	M 7, Nov 09
D5.5	BHL-Europe dissemination plan	M 6	M 7, Nov 09
D5.6	BHL-Europe promotion kit	M 6	M 7, Nov 09
D2.2	Prototypes of deduplication tool and bibliographic database system for monographs and serials	M 9	M 10, Feb 10
D3.4	Implement plans for all components in WP3, incl. data models, technology standards etc.	M 9	M 10, Feb 10
D4.1	Delivery of IPR working documents, including best practice guide, due diligence guide, pro forma agreements and process for formally agreeing rights management with rights holders. Complete agreement with EUROPEANA and BHL for reciprocal access and Rights metadata.	M 9	M 10, Feb 10
D1.2	Progress Report 2 including pre-financing request	M 12	M 15, July 10[1]
D1.3	Annual Report 1 including first ideas for BHL-Europe business plan	M 12	M 15, July 10[1]
D2.3	Prototype of Web-database for content management and collection analysis	M 12	M 13, May 10
D2.4	Content analysis and management status report 1 (metadata, page numbers, content providers)	M 12	M 15, July 10[1]
D3.5	Technical architecture status and progress report with particular focus on the development of the German prototype	M 12	M 12, Apr 10
D5.7	Online questionnaires for user survey	M 12	M 12, Apr 10

[1] These deliverables were initially submitted in M 12, but revised as a result of the BHL-Europe project review in May 2010.

3.3 Content Provider Requirements

We elaborated on the requirements of the content providers of BHL-Europe during the first year of the project. This work is still in progress, and requires several levels of detail towards the final best practice guidelines and standards, which will be published at the end of the project in April 2012. The first step was a library questionnaire for all participating libraries to better understand the library systems and workflows of the content providers. Parallel to this process, a catalogue of content holder requirements was compiled and discussed.

The first important achievement of the project was the agreement of the consortium on the imaging requirements for the scanned pages images (resolution, file format, etc.), on metadata requirements (granularity, fields, etc.), and on file submission guidelines. In a next step, we agreed on the quality and quantity of content to be provided by every partner with a Memorandum of Understanding (MoU). This information is used as a baseline for the BHL-Europe content ingestion planning.

3.4 Global References Index to Biodiversity (GRIB)

BHL-Europe creates a bibliographic database containing information on monographs and serials which have been scanned in the past and which are available in the format defined in the MoU for integration into Europeana. The database, complete with content management and deduplication functionalities, is called the "Global References Index to Biodiversity" (GRIB; see Figure 1). The first prototype is ready for testing and usage for BHL-Europe consortium partners.

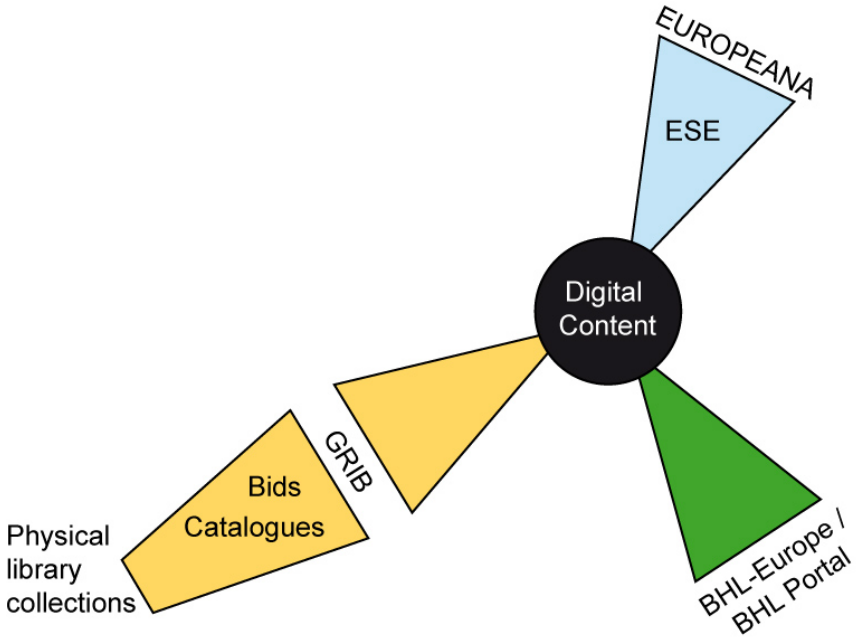

Figure 1. The Global References Index to Biodiversity is composed of the catalogue records of the physical collections, the index and the link to the page images (digital content). It is an important component of BHL-Europe and provides one access route for the digital image files in the repository. It is the major access route for the librarians managing the content hosted by BHL. The BHL-Europe users (taxonomists, general public) will mostly access the content either through the BHL Portal or Europeana (ESE = Europeana Semantic Elements).

Information will also be provided about which biodiversity literature is in the process of being digitised and which partner is responsible for scanning that material. The database will contain information on all relevant literature which should be scanned, and will identify the partner be responsible for providing the material. If there is no consortium partner with particular critical content, appropriate content holders will be identified and encouraged to join the network to provide this missing content. This system ensures that every content provider (and even potential content providers) can check before beginning digitisation to see if the material is already in the process of being digitised. The local digitisation processes can be planned accordingly and duplication is reduced to a minimum, ensuring the effective use of the local resources available in each partner institution.

3.5 BHL-Europe architecture

BHL-Europe is building an OAIS compliant system for archiving information (objects and metadata; see Figure 2). The OAIS (Open Archival Information System) framework is today the most relevant standard in digital preservation. During the first year of the project, an in-depth understanding of the OAIS components for BHL-Europe was created and the various components were described. Based on the OAIS component documentation, the BHL-Europe system will be built to ingest, manage, store, and disseminate the digital objects and metadata (see also D3.5 available on www.bhl-europe.eu soon).

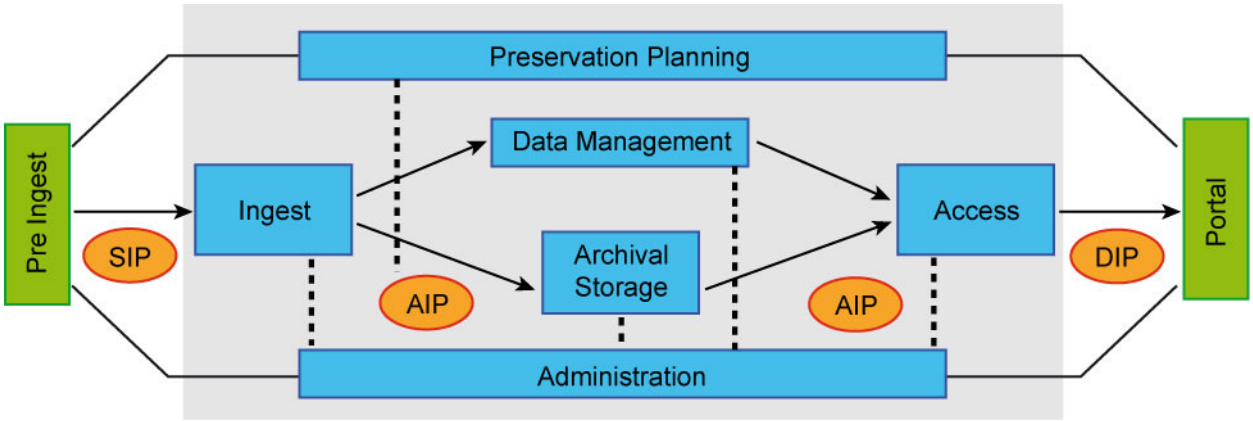

Figure 2. High-level overview of the OAIS components relevant for BHL-Europe (in blue). The OAIS reference model differentiates between three kinds of information objects. The SIP, Submission Information Package, is being sent in by the data producers (content providers), the AIP, Archive Information Package, is preserved in the archive, and the DIP, Dissemination Information Package, is provided for the consumers of the archive.

3.6 BHL-Europe content storage

The Biodiversity Heritage Library is currently still delivering the books and articles accessible via www.biodiversitylibrary.org through the Internet Archive. In order to establish a mirror of the BHL content in Europe and feed the content into BHL-Europe and Europeana, this content needs to be moved to Europe. As a first step to move the BHL content from the US to Europe, the Natural History Museum London (NHM) provided a hardware kit to BHL and set it up in the Marine Biological Laboratory in Woods Hole, Massachusetts, USA.

A second achievement was the establishment of a storage solution in the Natural History Museum of London designed to hold both the entire BHL content and the content of the BHL-

Europe partner libraries. This system is now in place and the first test of the data ingest will take place within the next months.

3.7 Harmonisation of metadata

A major part of the work of BHL-Europe and it's partner is the harmonisation of metadata in order to make the various repositories interoperable. At http://bhl.ait.co.at a DISMARC port was created for BHL-Europe with a first draft of an appropriate metadata set with the same base functionality. Metadata from various BHL-Europe content provider institutions was mapped against a MODS schema and inserted into the demo server. Last summer the test server contained 41,832 records from 14 partner institutions. Currently the DISMARC port is being used as a sandbox environment to further test content harmonisation and the mapping to the Europeana Semantic Elements (ESE) metadata schema.

3.8 BHL-Europe integration in Europeana

A major achievement of the first year of BHL-Europe is the submission of data to Europeana. With the Rhine release in summer 2010, Europeana will have the major part of the BHL corpus on display, i.e. more than 40,000 books of the biodiversity domain. In addition, two of the European partners of BHL-Europe, which are Naturalis (NAT) and the Oberoesterreichische Landesmuseen in Linz (LANDOE), will provide content to Europeana for the Rhine release.

In accordance with the decision to integrate these three initial BHL-Europe archives into the Europeana Rhine release an OAI provider for Europeana was set up. For first tests 10,799 records from BHL, 3,509 records from Naturalis (NAT) and 66 records from LANDOE have been loaded into the Europeana provider. The data was mapped to the current Europeana ESE profile and is being displayed in this format at (http://bhl.ait.co.at/). This allows the partners to give feedback to the mapping. Thereafter the data was inserted into the Europeana content checker tool where the data can be viewed in the official Europeana.eu style. The links to their data in Europeana-style was sent to the partners for further feedback.

The mapping of the BHL data is scheduled for completion at the end of April for harvesting by Europeana. The mapping of the two other datasets (NAT, LANDOE) is scheduled for completion in May, to be harvested by Europeana in June. The ingest plan of the other BHL-Europe partner libraries was discussed and agreed in order to be prepared for the re-harvest of data by Europeana in the months following the Rhine release.

3.9 Awareness and Dissemination

BHL-Europe partners were very active in the first year of the project to disseminate the project to various communities (e.g., scientists, librarians, ICT experts, teachers, publishers) through numerous conferences, meetings, and workshops. To date, already 38 events were attended by project members to present BHL-Europe mainly with talks or posters (Table 1).

Table 1. List of conferences and events attended by members of BHL-Europe in order to disseminate the project to the various audiences of the project.

Title	Start date	End date	Country
Synergy afternoon of the Finnish National Digital Library	20/05/2009	20/05/2009	Finland
Three Liaison Committees of Viikki Science Library	29/05/2009	29/05/2009	Finland
International Conference on Biodiversity Informatics	1/06/2009	3/06/2009	UK
Three Liaison Committees of Viikki Science Library	2/06/2009	2/06/2009	Finland
Three Liaison Committees of Viikki Science Library	8/06/2009	8/06/2009	Finland
ElPub2009	10/06/2009	10/06/2009	Italy
EDIT meeting "Scientific publishing in natural history institutions"	22/06/2009	22/06/2009	Slovak Republic
Willi Hennig XXVIII Annual meeting	22/06/2009	26/06/2009	Singapore
Primer Encuentro Ibérico de Biología Subterránea	10/07/2009	12/07/2009	Spain
I Iberian symposium on geometric morphometrics	23/07/2009	25/07/2009	Spain
Systematics 2009	10/08/2009	14/08/2009	The Netherlands
Workshop on Advanced Technologies for Digital Libraries 2009	8/09/2009	8/09/2008	Italy
27th Annual Meeting of the Natural History Curators of the Carpathian Basin	8/09/2009	11/09/2009	Slovakia
XXVI Jornadas de la Asociación Española de Entomología	12/09/2009	15/09/2009	Spain
International Symposium on Islands and Evolution	14/09/2009	17/09/2009	Spain
30. Österreichischer Bibliothekartag	17/09/2009	17/09/2009	Austria
BHL-Europe presentation at RBINS	24/09/2009	24/09/2009	Belgium
Athena project workshop	30/09/2009	30/09/2009	Czech Republic
DINI Jahrestagung	30/09/2009	1/10/2009	Germany
Two membership meetings at UH-Viikki	1/10/2009	31/10/2009	Finland
79. Jahrestagung der Paläontologischen Gesellschaft	5/10/2009	7/10/2009	Germany
Open Access Days	7/10/2009	8/10/2009	Germany
SUN PASIG	7/10/2009	9/10/2009	USA
Digital Libraries and Digital Preservation: ICT Call 6 Information Day	9/10/2009	9/10/2009	Luxembourg
Improving Access to European Cultural Heritage; Europeana Round Table event for aggregators	13/10/2009	16/10/2009	Sweden
2nd LIBER-EBLIDA Workshop on Digitisation of Library Material in Europe	19/10/2009	21/10/2009	The Netherlands
Biodiversity Research Conference	4/11/2009	4/11/2009	Belgium
Segundas Jornada del Departamento de Biodiversidad y Biología Evolutiva del Museo Nacional de Ciencias Naturales	5/11/2009	5/11/2009	Spain
TDWG - Taxonomic Database Working Group annual conference	9/11/2009	13/11/2009	France
14th international congress "Cultural heritage and new Technologies"	17/11/2009	17/11/2009	Austria
Biologentag 2009	20/11/2009	21/11/2009	Germany
LAPI III, annual meeting of the Latin American Plants Initiative	30/11/2009	4/12/2009	Colombia
4 Congreso Ibérico de ornitología (SPEA)	5/12/2009	8/12/2009	Portugal
MÁSTER OFICIAL BIODIVERSIDAD en ÁREAS TROPICALES y su CONSERVACIÓN 2ª edición	10/12/2009	10/12/2009	Ecuador
EDIT General Meeting 2009	15/12/2009	17/12/2009	Portugal
Launch of the International Year of Biodiversity 2010	11/1/2010	11/1/2010	Germany
SciColl Conference - International Coordination of an Interdisciplinary Global Research Infrastructure	8/2/2010	9/2/2010	Belgium
CETAF27	27/4/2010	28/4/2010	Austria

4 Target Users & their Needs

As presented above, BHL-Europe will provide three routes of access to the digitised books to respond to the different needs of the content users (see Table 2). Libraries need a tool to manage the scanning process and to identify relevant content for scanning. They will use the GRIB to fulfil their needs. Scientists need a way to nominate content for digitisation and get an overview of the entire corpus of biodiversity literature held by the various libraries of the European natural history museums and botanical gardens. This need can also be covered by the GRIB. The direct online access to the biodiversity books stored in the BHL-Europe repository will be facilitated through Europeana and the multilingual BHL Portal to be built by BHL-Europe. Users interested in community specific functionality (e.g. Taxonomic Intelligence) will use the BHL Portal. General interest readers also interested in the wider (cultural heritage) context of the literature will use Europeana as their primary route to the BHL-Europe content.

BHL-Europe has carried out one of two specific online user evaluations during the project (see 6.1 below). The results will be evaluated and published this summer. In addition to these large scale evaluations, we are in permanent contact with individual users representing our target users (Table 2). We receive direct feedback from these users for the further development of our prototypes into the live system. All user requirements will deliver input into the design of the BHL-Europe system and portal.

Table 2. BHL-Europe target users and their needs.

Target user description	Needs
1.1) European citizens	Direct online access to comprehensive information not currently publicly accessible to help raise the awareness and appreciation of biodiversity heritage
1.2) Scientists (e.g. Biology)	Taxonomic descriptions of species; biodiversity data of specific regions in the last centuries; full-text searching; taxonomic intelligence
1.3) Scientists (e.g. History, Cultural heritage)	Historical information on science and scientists
1.4) Citizen scientists / Hobby scientists	Search, read, download, and print articles about biodiversity in their area
1.5) Students of different levels (primary to academic)	Reliable and meaningful information and relevant images on biodiversity; minimal time to aggregate information from different sources; research resource
1.6) School teachers	Resource for teaching materials as complement to textbooks
1.7) Environmental and Conservation agencies / Government officials / Policy makers	Information on impact of climate change, environmental deterioration and human interventions
1.8) Artists	High quality images of animals and plants
2.1) Libraries	Information on the distribution of heritage material (metadata); new platform for presentation of content
2.2) Digitisation centres	Best practice guidelines for the digitisation of heritage literature; new platform for presentation of content
2.3) Digital library / Open Access networks	Best practice guidelines for the establishment of digital library networks; information about digital repositories; new platform for presentation of content
2.4) Universities / museums	Best practice guidelines for the digitisation of heritage literature; new platform for presentation of content

5 Underlying Content

Currently, BHL-Europe has 17 content providers involved as consortium members in the project: NHM, NMP, UGOE, LANDOE, HNHM, UCPH, NAT, NBGB, RMCA, RBINS, BnF, MNHN, CSIC, RBGE, SIL, MOBOT, UH-Viikki. In addition to these content providers, we expect new partners to provide further content during the project's lifetime. These new partners may be other important libraries in countries not yet engaged in the project or learned societies. Currently, all content providers have a total of more than 46,000 books to contribute to the BHL-Europe project. This number is continuously increasing as scanning continues.

The content providers within the consortium create a critical mass of high quality digital content representing the biodiversity domain. The content providers were selected on the basis of their ability to contribute key biodiversity and taxonomic literature (zoology, botany, palaeontology).

Content is not restricted by proprietary third-party rights or by any other constraints, which would limit its use in an open access environment using Creative Commons http://creativecommons.org/ licences. The digital content is either in the public domain or otherwise the content contributors will have obtained permission from IP owners to allow open access under a Creative Commons license.

The initial focus on public domain material is not a limitation of the project, as systematic biology depends more than any other natural science upon historic literature. Another reason to focus on historical literature is that many old and important monographs are themselves inherently very rare, fragile or in need of conservation. This makes "hands on" access very difficult. This project substantially reduces the need for the physical handling of these rare and valuable materials.

6 Summary of Activities

6.1 Important past activities

- The Museum für Naturkunde established effective communication mechanisms in the consortium (Email, Skype/Conf calls, BHLwiki, Google Group, project meetings). Altogether 15 meetings of BHL-Europe team members were organised and held in different locations across Europe. We also invited members of other related projects to some of these meetings to exchange experiences and identify areas for cooperation.
- Several of our partners have already started to build lists of books and serials that are relevant for our community. These lists will be input into the GRIB platform and integrated by the system once GRIB is ready in spring 2011.
- The main technology partners (AIT, Atos) worked intensively on the BHL-Europe system architecture and on an implementation plan for the German prototype and the BHL-Europe system. This also included the evaluation of available technologies/solutions to be considered by BHL-Europe. Both partners were actively involved in the discussion and in consensus-building work for both the metadata schema and for use cases in BHL-Europe. The post-processing and ingest of page images is another major step towards building the BHL-Europe system. Based on the experiences and documentation of BHL-Europe, the procedure of uploading files to

Internet Archive to generate derivatives for the BHL Portal is now established further for use by partners. Partners carrying out digitisation can use this technology to enhance and process page image files for display through the BHL Portal and Europeana. This uses existing technologies for post-processing of content in order to deliver the page images (i.e. the actual digital content) in the format agreed by BHL-Europe project partners.

- A first set of IPR working documents was delivered by the BHL-Europe WP4 IPR, led by the Natural History Museum London (NHM). It contains a best practice guide, a due diligence guide, and pro-forma agreements. These documents are an important pre-requisite to help our partners with IPR questions in clearing copyrights for content. The documents are already used and appreciated by project partners and can be downloaded under www.bhl-europe.eu/presentations.php.

- While the first extensive user requirements survey of BHL-Europe was planned for the end of the first year of the project, a test survey was launched in October 2009 within the BHL-Europe consortium. The goal of this test survey was (1) to test the survey methodology, (2) to identify appropriate questions including the wording of the questions, (3) to collect first user requirements to be used for the initial design of the BHL-Europe prototype. 52 users completed the questionnaire. The experiences gained from this survey were extremely valuable in order to design the extensive survey, which was launched on 15 March 2010. The survey is now closed and the results are expected to be published on www.bhl-europe.eu/presentations.php by 31 July 2010 at the latest.

  The test survey also gave us some first important input for the further development of the BHL Portal and the implementation of the BHL-Europe system. Users need an improved search, filter, and sorting functionality for the portal. This is very important in order to find the relevant literature and quickly access the books and articles. Both downloading books as PDFs and reading the book online is required. Time is a very critical aspect as web sites, books and book pages must load quickly into the browser window. With the second BHL-Europe user requirement survey launched on 15 March 2010, these results are further refined and we expect more specific input for the development of the BHL-Europe system and prototypes.

- During the first months of the project the BHL-Europe WP5 Dissemination led by the Narodni Muzeum Praha (NMP) designed and established the project Web site under www.bhl-europe.eu. Furthermore NMP produced an initial set of promotional material (flyer, business cards, posters, roll-ups) to present the project adequately at conferences and similar events.

6.2 Important future activities

After the consensus building and conceptual activities of the first year, we now focus upon on the first important tangible results for delivery to the user during the second year of the project. The most important and crucial ongoing activities are focused on the German prototype of BHL-Europe and a fully functional GRIB system. The German prototype of BHL-Europe is expected to be delivered in October 2010 and should allow search and retrieval of biodiversity literature in English and German. The current BHL Portal under www.biodiversitylibrary.org is only working in English, thus the German prototype is a first step towards a multilingual BHL-Europe portal.

After establishing the GRIB infrastructure and functionality in the first year of the project, the second year will be devoted to increasing the quantity of catalogue records by connecting the BHL-Europe partner library catalogues to GRIB. This will give us the opportunity to analyse all relevant content and to manage the content acquisition process. All users can search for biodiversity related literature and find it even if it is not available digitally. The user will then have the chance to search for physical copies of the books in a library and also identify digital resources. If books are not available digitally, the user has the option to nominate books for scanning. This does not mean that the item will be scanned, but it significantly increases the probability of scanning. BHL-Europe will manage a priority list of items to be scanned by those partners that have the ability and capacity for scanning. For librarians, GRIB presents the possibility of managing the scanning process in its entirety, from the initial process of bidding for books or collections until the final ingestion of the items into the BHL-Europe repository for search and retrieval.

The end of Year Two of BHL-Europe coincides with the Danube release of Europeana. It is important that at least 50% of the BHL-Europe content is available by April 2011 via Europeana. We will align with Europeana to adapt and improve data mapping and ingest procedures to ensure an effective data harvesting procedure.

7 Impact & Sustainability

BHL-Europe contributes to the i2010 European Digital Library initiative of the European Commission to improve the online accessibility of digital content. The results of the project will remain accessible online also after the end of the duration of the project through various platforms. One of these platforms is Europeana, plans for which extend beyond the end of the BHL-Europe project.

BHL-Europe is also embedded in a global framework through BHL. The BHL community is currently growing and new regional partner projects are joining the community. BHL-China and BHL-Australia have already signed a Memorandum of Understanding with BHL and other partners are expected to join soon. This quick growth underpins the importance of BHL on a global scale. This helps BHL-Europe to maintain a sustainable service in Europe for the future. This global community also helps BHL-Europe to bring together and repatriate that part of the European cultural heritage that is currently not available in European institution and to provide a digital copy to BHL-Europe and Europeana.

8 Conclusion

For the first year of the project, BHL-Europe is healthy and well on-track. We successfully managed staffing issues, communications among partners are effective, the work is being carried out in a timely fashion and there are currently no barriers foreseen which could significantly delay the major outcomes of the project. The project is also well on budget, but the most time-consuming work is still ahead of us. We now look forward to the next year which will deliver the German prototype of BHL-Europe and a fully functional GRIB system. The implementation of these prototypes will also lead to the increased harvesting of content and data from our partners and the necessary adaptation of the data quality. We also will

continue to make more BHL-Europe content available for the Danube release of Europeana in April 2011.

9 Further Information

Please consult the BHL-Europe Web site www.bhl-europe.eu for any further information on the project. You can register for the quarterly newsletter and can find all publicly available project results. You can also contact the project office in Berlin below:

Dr. Henning Scholz
BHL-Europe Project Coordinator
Museum für Naturkunde - Leibniz-Institut für Evolutions- und Biodiversitätsforschung an der Humboldt-Universität zu Berlin
Invalidenstrasse 43
D-10115 Berlin
Germany

Email: bhl-europe@mfn-berlin.de

Appendix 1 Consortium

Museum fur Naturkunde - Leibniz-Institut fur Evolutions- und Biodiversitatsforschung an der Humboldt-Universitat zu Berlin (MfN)

With more than 30 million 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.

Role: Project Coordinator, Work Package leader, Technology Provider, Library expert, Taxonomist/User, Disseminator

Natural History Museum (NHM)

The Natural History Museum in London is one of the world's greatest museums, with over 3.9 million visitors and 13 million online visitors per year. The NHM is also an international leader in the scientific study of the natural world. The NHM Library has the largest collection of natural history material in the world, with over 1 million books (from 1469 onwards), 25,000 journal titles and 600,000 works of art. The NHM is a leading participant in the BHL project and an active player in the Encyclopedia of Life project.

Role: Work Package leader, Content Provider, Technology Provider, Library expert, IPR expert, Disseminator

Narodni 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 Naprstek Museum of Asian, African and American Cultures, Czech Museum of Music and 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.

Role: Work Package leader, Content Provider, Library expert, Taxonomist/User, Disseminator

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.

Role: Technology Provider, Library expert, Disseminator

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).

Role: Technology Provider, Disseminator

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.

Role: Technology Provider, Disseminator

Freie Universitat 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.

Role: Technology Provider, Library expert, Taxonomist/User, Disseminator

Georg-August-Universitat Gottingen Stiftung Offentlichen Rechts (UGOE)

The EZOOLO/AnimalBase project is located at the Georg-August-Universitat of Gottingen. It was initiated as a joint venture of Gottingen 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.

Role: Content Provider, Library expert, Taxonomist/User, Disseminator

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. Additional a library with many historically important volumes is available. The library with c. 6,000 scientific journals and tens of thousands of books complement the National Library of Austria and the University's libraries nearby.

Role: Technology Provider, Library expert, Taxonomist/User, Disseminator

Land Oberosterreich (Oberoesterreichische 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.

Role: Technology Provider, Content Provider, Library expert, Taxonomist/User

Hungarian Natural History Museum (HNHM)

HNHM holds more than 10 million natural history items. HNHM Library contains more than 300,000 volumes. 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.

Role: Content Provider, Library expert, Taxonomist/User, Disseminator

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.

Role: Library expert, Disseminator

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).

Role: Technology Provider, Content Provider, Taxonomist/User, Disseminator

Stichting Nationaal Natuurhistorisch Museum Naturalis (NAT)

Naturalis was founded in 1820 and much of its collection dates back to the 19th and 20th century. The collections of zoological, palaeontological and geological objects are estimated to total about 12 million objects. 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.

Role: Technology Provider, Content Provider, Library expert, Disseminator

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.

Role: Technology Provider, Content Provider, Library expert, Taxonomist/User, Disseminator

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 million 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. Information on biodiversity is repatriated to African partners, suitable training in taxonomy is also provided.

Role: Technology Provider, Content Provider, Taxonomist/User, Disseminator

Royal Belgian Institute of Natural Sciences (RBINS)

The RBINS houses a diverse and exceptionally rich collection, comprising about 37 million 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. RBINS has begun digitisation of the library catalogue and more than 185,000 titles are online.

Role: Content Provider, Taxonomist/User, Disseminator

Bibliotheque 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.

Role: Technology Provider, Content Provider, Library expert, Disseminator

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. 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.

Role: Content Provider, Library expert, Disseminator

Consejo Superior de Investigaciones Cientificas (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 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.

Role: Content Provider, Library expert, Taxonomist/User, Disseminator

Universita 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).

Role: Library expert, Disseminator

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.

Role: Content Provider, Library expert, Disseminator

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.

Role: Taxonomist/User, Disseminator

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.

Role: Content Provider, Disseminator

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. The BHL project is led from SIL.

Role: Content Provider, Library expert, Disseminator

Missouri Botanical Garden (MOBOT)

The mission of MOBOT is to discover and share knowledge about plants and their environment 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 project and is supporting the development of the system's infrastructure, application layers and interfaces.

Role: Technology Provider, Content Provider, Disseminator

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.

Role: Technology Provider, Content Provider, Library expert, Disseminator

Humboldt-Universitat 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.

Role: Technology Provider, Disseminator