ECP-2008-DILI-518001

BHL-Europe

Progress Report

1 May 2011 - 31 October 2011

Deliverable number/name	*D1.8*
Dissemination level	*Confidential*
Delivery date	*16 December 2011*
Status	*Final*
Authors	*Jana Hoffmann, Henning Scholz*

*e*Content*plus*

Document History

0.0 Contributors

Person	Partner
Jana Hoffmann	MfN
Henning Scholz	MfN
Melita Birthälmer	MfN
Lola Obajuluwa	NHM
Chris Sleep	NHM
Graham Higley	NHM
Jiri Kvacek	NMP
Consortium members	BHL-Europe

0.1 Revision History

Revision Date	Author	Version	Change Reference & Summary
15 Nov. 2011	Jana Hoffman	0.1	First draft
22 Nov 2011	Jana Hoffmann	0.2	Second draft
12 Dec 2011	Henning Scholz	0.3	Third draft including reviewer comments
16 Dec 2011	Henning Scholz	1.0	Final version

0.2 Reviewers

This document requires the following reviews and approvals.

Name	Position	Date	Version
Birthälmer, Higley, Scholz, Obajuluwa, Sleep, Kvacek	WP Leader	15 Nov. 2011	0.1
Martin Gordon	Project Manager, RBB	29 Nov 2011	0.2
BHL-Europe consortium		30 Nov 2011	0.2

0.3 Distribution

This document has been distributed to:

Group	Date of issue	Version
Martin Gordon and BHL-Europe consortium	22 November 2011	0.2
BHL-Europe consortium	16 December 2011	1.0

Table of Contents

2 Summary

The main objectives of the reporting period (May 2011 to October 2011) of the project were:

(1) Review approaches for the establishment and management of multilingual biodiversity digital libraries.

(2) Review state-of-the-art technologies used by Europeana and BHL for processing digital content.

(3) Evaluate and document processes used by BHL-Europe partners for the digitisation of biodiversity content.

(4) Implementation of the requirements of BHL-Europe content providers regarding metadata schema (OLEF implementation, schema mapping).

(5) Implementation of the ingest workflow of content from the BHL-Europe content providers.

(6) Continue implementation of the technical and functional architecture of the BHL-Europe system.

(7) Design and implementation of a multilingual access point for the search and retrieval of biodiversity content through at least two portals (Europeana and BHL-Europe).

(8) Refine the user requirements involving the proposed target users of BHL-Europe.

(9) Make BHL-Europe content available for harvesting by Europeana.

(10) Develop operational strategies and processes for long-term preservation and sustainability of the data produced by national biodiversity digitisation programmes.

(11) Raise awareness and ensure that the project outputs are known and used by the target users and that the proposed approach directly addresses user needs.

(12) Negotiate with Rights Holders & publishers to enable access to in-copyright content.

(13) Extend the network of content providers.

BHL-Europe provides four access points to the digital biodiversity literature of the partner libraries: the Global References Index to Biodiversity (GRIB), the multilingual BHL-Europe portal, Europeana and Global BHL. After overcoming technical problems on the part of the GRIB service provider, the GRIB was set up again in this reporting period and is ready for testing soon. We are making good progress towards the launch of the final BHL-Europe system and portal. For the first time in the project, we now have prototypes of all components available and integrated to the extent that we can run tests over the entire system. This includes ingest of content according to our specifications. However, implementation of the system, due to its technical complexity, took longer than expected and delayed ingest and integration work in the reporting period. Remedial actions are in place to bring the integration work back on track. We continue to provide access to ten collections of BHL-Europe content providers through Europeana and we update this dataset on a regular basis.

We made good progress towards a sustainable BHL-Europe service after the end of the project. The business plan is almost finished and concepts for a further involvement of BHL-Europe in new related projects are available. We are also seen by international initiatives as a valuable partner for re-use and implementation of our results elsewhere.

We are actively disseminating the project and its result. The work on the virtual exhibitions is progressing well and plans are in place to increase dissemination activities with the launch of the BHL-Europe portal. BHL-Europe is continuously involved in the activities and processes of Europeana, mainly through the Europeana Group of Projects activities and the Council of Content Providers and Aggregators.

3 Status

3.0 Status and progress of activities

3.0.1 Status and progress of activities – WP1

3.0.1.1 Resource planning, monitoring and controlling

The project coordination, project partners and BHL-Europe associates followed the project progress on the BHL Wiki. They used the information provided for planning tasks and meetings as well as for the preparation of their reports.

With the finalisation of the financial statements (see below), we are now in a position to check and revise the budget and resources of all partners individually. We have to ensure that the remaining budget is spent as efficiently as possible and avoid under- or overspending in any area of the project. Therefore, the BHL-Europe project office is now working on the budget and personnel resources spent so far, and on the budget planned to be spent by the end of the project in order to validate the indicative budget per partner until the end of the project. Partners are approached on an individual basis starting with partners having the highest risk of under- or overspending the budget. This process goes hand in hand with the redistribution of tasks, if necessary, to make sure that all work explicitly mentioned in the Description of Work is completed by the end of the project. This is a complex process also involving further budget shifts and discussions with the EC. All major decisions will be communicated to/ discussed with the consortium when all details are available.

3.0.1.2 Financial statements

The project office prepared and updated the guidelines[2] and planned the process to allow a timely delivery of all documents. This year, we were in a much better position to thoroughly check and validate the financial statements compared to last year due to the improved reporting process in BHL-Europe (activity reports). The project assistant checked all financial statements against the monthly activity reports to identify inconsistencies in personnel commitment, for example. This supports all partners in validating their internal project documentation. The financial statements were submitted to the EC on 31 August 2011. After the approval of the documents by the EC, the next tranche of the pre-funding was remitted to the MfN account. The BHL-Europe project office started to distribute the 3rd tranche to partners with completed budget revision.

3.0.1.3 Farewell to BHL-Europe colleagues

Nancy Chillingworth – The BHL-Europe Work Package 4 was officially closed end of April thus Nancy's contract ended for BHL-Europe. However, she is still located at the NHM and will be available for IPR related questions.

Adrian Smales – Adrian's contract ended end of April and Work Package 3 lead was handed over to Graham Higley, Lola Obajuluwa and Chris Sleep.

Lesa Ng – Lesa completed her nine months contract for Work Package 2 in BHL-Europe at RBGE at the end of April 2011.

---

[2] https://bhl.wikispaces.com/BHLE-FS

Zsuzsanna Forgács – Zsuzsanna's contract for Work Package 2 in BHL-Europe at HNHM ended at the end of July 2011.

The current project assistant of BHL-Europe, Jana Hoffmann, will leave the project end of November 2011. Handing-over activities were started already well in advanced to enable a smooth continuation of the project administration beginning of December. A new project assistant will be hired mid of November at MfN.

3.0.1.4 PMG conference calls and communication

PMG calls were held regularly in the reporting period. The minutes from most PMG calls are provided on the BHL Wiki[3]. As the PMG had several discussions on budget and resources, the notes of some of the PMG calls were not made publicly available.

An Open PMG Call was held on 6 September 2011 in order to communicate the project progress in the respective work packages, to present the outcomes of the Public Core Review Meeting in The Hague and to enable partners to ask questions to the PMG members. The minutes of the Open PMG Call are available on the BHL Wiki[4].

3.0.1.5 Contractual Meetings

2nd BHL-Europe Annual Review Meeting in Tervuren/ Meise, Belgium

It was agreed that the meeting will be organised by NBGB in collaboration with RMCA and held at two different places, in Meise and Tervuren, Belgium. The organisation of logistics, preparation of the tight meeting's agenda, organisation of the review meeting and multiple parallel workgroups was a time-consuming and challenging task.

The meeting started 7 June in Meise (NBGB) with a PMG meeting for preparation of the official Review Meeting. In the afternoon, two parallel sessions took place: (1) Collection Policy and (2) Evaluation and Testing. The meeting then continued 8-10 June in Tervuren, Belgium, at RMCA. The official review meeting on 8 June was attended by members of 24 BHL-Europe partners. It was a very successful review and BHL-Europe received a green flag. However, one deliverable was amended and re-submitted to provide a more detailed content ingest plan to Europeana and the BHL-Europe portal in the third project year.

In the morning of 9 June two parallel workshops were held, one dedicated to dissemination activities with a particular focus on the virtual exhibition, and a second workshop for our content provider to demonstrate the metadata mapping tool and the Pre-Ingest tool. In the afternoon there were three parallel workshops: a content provider workshop on the GRIB, a business development workshop to investigate gaps in the existing D1.5 document for further development into D1.9, and the first part of the portal development workshop for the development team (portal review). The outcomes of the workshops were presented to the entire consortium in the early evening.

The next day, 10 June, the portal development workshop continued. A parallel workshop was held to plan the next steps in finalising the Best Practice Guide for M36 of the project. The meeting was closed by a TMB session to discuss the last open questions and existing issues towards the Public Core Review Meeting in August with the developers and decision makers in the project. All details including agenda, notes and participants are available on the BHL Wiki[5].

Public Core Review Meeting in The Hague, The Netherlands

---

[3] https://bhl.wikispaces.com/BHL-E_PMG_calls

[4] https://bhl.wikispaces.com/BHL-E_PMG_calls

[5] https://bhl.wikispaces.com/BHLE_AnnualReview2

The Public Core Review Meeting took place from 30 August to 2 September 2011 at the National Library of the Netherlands, Koninklijke Bibliotheek[6] (KB) in The Hague. It was hosted by Europeana (EDLF) with organisational support of the BHL-Europe office at MfN. The main agenda items were the portal review and the round table of testers and developers to analyse the status of the development work and the progress towards the delivery of the final BHL-Europe system. Afterwards, the portal testers held a workshop to discuss the general (technical) requirements for the portal and the design concept. Simultaneously, the developers held a technical integration workshop. In the following TMB sessions the next steps and work plans were discussed and agreed. A dissemination workshop concluded the meeting.

In addition, a PMG and business development meeting were held. A side meeting with the IMPACT team was organised in parallel at the KB to plan the next steps in analysing the BHL(US) dataset to improve the OCR. More details on the outcomes of the various workshops during the meeting are provided below (see particularly the WP3 section). The full documentation of the meeting can be found on the BHL Wiki[7].

Content Provider and Technical Review Meeting in Brussels, Belgium

The preparation of the next contractual meeting 29 November – 2 December at RBINS in Brussels, Belgium was initiated. A BHL Wiki[8] site was created and relevant information is provided. A first draft of the meetings' agenda was presented and discussed during the PMG Call. Regular conference calls between the host and the BHL-Europe office will be held in order to prepare this meeting.

3.0.1.6 BHL-Europe internal reporting

Monthly activity reports were collected from most of the partners, quality checked and if necessary the partners were contacted in order to discuss necessary changes. The checked activity reports were sent out for approval by the work package leader.

The status report for the reporting period April-June and July-October were collected in July and beginning November.

The Interim Report #5 for the reporting period April-June 2011 was finalised based on the input of the partners (status reports). Unfortunately, many partners did not submit their reports in time which resulted in a delay of the finalisation of the Interim Report. The report was made available for review by the consortium in September. The finalised report is available on the BHL Wiki[9].

3.0.1.7 Risk and issue management

Risk and issues of the project were collected with the status reports from each partner and were completed by the respective project manager or coordinator. The risk and issue log of the project was discussed and evaluated during the PMG calls and appropriate counter measures were decided. A short summary of important risks, issues and dependencies is provided below.

3.0.1.8 Project management

Some specific project management work was required in the last reporting period.

---

[6] http://www.kb.nl/

[7] https://bhl.wikispaces.com/BHLE_Public_Core_Release_Meeting

[8] https://bhl.wikispaces.com/BHLE_Brussels_Meeting

[9] https://bhl.wikispaces.com/BHLE_WP1_reports_INTERIM

In month 24 of the project WP4 was officially closed as planned according to the DoW. In order to plan all related actions in more detail an additional meeting was held in London 7 April with the NHM team and the project coordinator. This was to confirm that before the closure of WP4 all relevant tasks were carried out and Work Package results were documented properly. Furthermore, it was discussed how to proceed with IPR questions arising during the final year of the project. Nancy Chillingworth, the WP4 leader, will continue to work at NHM and agreed to act as a consultant for IPR questions during the last year of the project. Nancy will also continue to work for the Legal Working Group of the CCPA, which facilitates the follow up of legal aspects related to Europeana. This will be of particularly interest when the Europeana Data Exchange Agreement needs to be finalised.

At the end of April the contract of our WP3 leader Adrian Smales at NHM ended. Handover activities of WP3 management tasks from the current Work Package leader, Adrian Smales, to the NHM project management team and future roles and responsibilities were discussed during the London meeting. Graham Higley will act as the official WP3 leader for the transition period until NHM has hired a new head of IT. Chris Sleep will be responsible for managing all technical aspects related to the development of the BHL-Europe system and Lola Obajuluwa for the overall WP3 project management. This assured a smooth continuation of all of the WP3 management tasks in this important period of the project.

3.0.1.9 Production of deliverables

All nine deliverables submitted for month 24 were accepted including the revised documents. They are available on the BHL Wiki.
The resource table in the Progress report D1.4 and D1.6 was updated according to the figures provided with the second financial statements of all partners. This section could not be finalised before all financial statements were cross-checked and approved. An amended version of the resource tables (but not the reports) was sent to the desk officer at the end of August. The annual report and the final progress reports with the updated resource tables can be found on the BHL Wiki[10].

3.0.1.10 Business development

BHL-Europe should present the business plan for long-term sustainability (D1.9) in M36 of the project. The more general plan (D1.5) was further elaborated during the Business Development Workshop in Tervuren. A survey was planned to collect information from our content providers on their future perspective in providing content to BHL-Europe after the end of the project. This survey was evaluated in the reporting period. The results were used to further develop the business plan during a workshop at the beginning of the last project meeting in The Hague (see above). It was agreed that the current draft needs to be somewhat restructured and amended. Based on current progress we should be in a position to finalise the work at the document well before M36. We plan to finalise the document, including some first ideas on follow up projects, in time for the next project meeting in the next few weeks. We are also using the Life & Literature conference in Chicago to discuss various aspects of the plan with experts attending the conference. The business plan will then be presented during the plenary session of the next project meeting in Brussels and some details (mainly on new projects) will be discussed during a workshop during that meeting. Afterwards, the business plan and the ideas for implementing the BHL-Europe business model will be presented in a workshop during the DISH conference[11]. This workshop will be held in

---

[10] https://bhl.wikispaces.com/BHLE_WP1_reports_Progressreports

[11] http://www.dish2011.nl/

collaboration with the IMPACT and CATCHplus projects. After this workshop, the draft of D1.9 will be communicated to the global BHL community for further input. It is planned to publish the plan in January 2012.

3.0.1.11 Networking activities (Europeana and Natural Europe)

The second Europeana Group of Projects meeting was held on 20 May 2011 in The Hague, Netherlands, and was attended by 13 projects. A key point on the agenda was the discussion of the 'Project share space' to be built over the summer 2011. This web space will bring together content from all of the Europeana Projects to facilitate knowledge exchange and communication. BHL-Europe is participating in the pilot project to set up this exchange service. The BHL-Europe project coordinator will coordinate these activities. An afternoon session was dedicated to a 'lessons learnt' session and small workshops were held to discuss issues and various options. The dissemination session was particularly important for BHL-Europe as it demonstrated that Europeana is the public end-user space for all content. Other project's sites or portals are either for their professional users or, if end-user directed, for a scientific audience. Thus it is not a primary task of our dissemination activities to reach out to the public. More collaboration with Europeana is certainly beneficial to cover this important user group, the public users, more efficiently.

The newly launched Europeana Search API and Search Widget API were circulated in the BHL-Europe consortium. They enable integration of European cultural and scientific heritage in content provider's online environments in a format that is convenient and meaningful to end-users. These Europeana API services can help to successfully engage online audiences by delivering richer and relevant content to them in new, intuitive, visually appealing ways. The improved user experience translates into an increase in high-quality traffic, active online communities and stronger brand association. At the moment these services are only available to Europeana network partners having an own portal. More details can be found here: API website[12]. Lizzy Komen (EDLF) presented the integration of the Europeana API by RMCA[13] as a best practice example of Europeana API integration during the BHL-Europe Annual Review Meeting.

The Natural Europe [14] project was launched in October 2010 in the framework of the ICT-PSP programme. Natural Europe aims at placing museum visitors in the centre of an exhibition that is both physical and virtual, using technology, services and learning content to provide personalised, appealing educational experiences, changing the way museum education is perceived and conducted, worldwide (Natural Europe vision taken from the 1st Natural Europe Newsletter). The project coordinator is currently investigating the level of a collaboration of BHL-Europe with Natural Europe is possible. Knowledge transfer and data exchange are two possible fields of collaboration. A face-to-face meeting is required to explore all options in more detail.

3.0.1.12 Consortium agreement

The final version of the consortium agreement including signatures of all partners except Wiley and SIL is available on the BHL Wiki[15]. As both partners are not payed for any work they are doing on behalf of BHL-Europe, it is no issue to not have their signatures. As the

---

[12] http://version1.europeana.eu/web/api

[13] http://bit.ly/AfricamuseumEuropeanaAPI

[14] http://www.natural-europe.eu/

[15] https://bhl.wikispaces.com/BHLE_WP1

consortium agreement is not a contractual obligation for BHL-Europe, it is also not causing any damage to the project.

3.0.2 Status and progress of activities – WP2

3.0.2.1 IT Development

The document 'Final and enhanced Web-database for content management and collection analysis' (D2.5) was finalised at the beginning of this reporting period with the support of Boris Jacob. The document was reviewed by the work package leader, the consortium and an external reviewer before handing it over to the project coordinator. D2.5 serves currently as a manual for the GRIB and describes the functionalities of the system.

The Digitisation Management Widget of the GRIB was finalised during this reporting period. The GRIB was planned to be up and running for the BHL-Europe Annual Review meeting, but due to technical difficulties at the VZG, the GRIB could not be put into production for the BHL-Europe partners in time. The technical issue was identified but not easily solved. Therefore intensive communication took place to ensure that the GRIB was up and running as soon as possible and that all available catalogues were imported and de-duplicated. Following-up on the GRIB finalisation a meeting with BHL-Europe WP2 leader, Boris Jacob and the VZG took place in Berlin 19-20 July. A new timeline and future tasks were discussed. Technical specifications for the BHL-Europe Public Core Review Meeting were updated. Furthermore, possibilities of integrating the GRIB into the BHL-Europe portal were discussed during the BHL-Europe Public Core Review Meeting in The Hague beginning of September. The problematic technical issues with the underlying infrastructure of the GRIB were solved at the beginning of October and the GRIB is now again displaying the re-imported catalogues. The re-import of the library catalogues into the GRIB are ongoing.
NBGB installed the Z39.50 protocol in order to facilitate harvesting by the GRIB. In April RBGE created further bibliographic records of mycological and phytopathological books for inclusion into the Global References Index for Biodiversity.
Additional work on promoting the GRIB was started by preparing posters for the following conferences and meetings:
- TDWG 2011 Annual Conference: 16-21 October 2011, Digitization Methods, Technologies, and Standards for Biodiversity Collections, New Orleans, Louisiana, USA.
- Sherborn Symposium: 28 October 2011, Anchoring Biodiversity Information: From Sherborn to the 21st century and beyond, London, UK.
- Life & Literature: 14-15 November 2011, Chicago, Illinois, USA.

The Scanlist is still maintained and is running in parallel to the GRIB to ensure a smooth change from the Scanlist to the GRIB until the GRIB is 100% stable and in use by our consortium partners.

3.0.2.2 Analysis of domain content

At the beginning of this reporting period the writing, layout, and internal and external review of 'Content analysis and management status report 2' (D2.7) took place and occupied most of the time. During the Annual Review Meeting the reviewers asked for a detailed upload and ingest plan for our BHL-Europe portal and Europeana, and requested a more detailed description of the content acquisition methodology. Over the summer the upload and ingest plan was produced in cooperation with our content providers and the WP3 leading team. Furthermore the content acquisition chapter was updated with more detailed information and

a list of potential content providers was provided. These providers will be approached as soon as the BHL-Europe portal is online.

The updated version of D2.7 was resubmitted to the European Commission end of August 2011.

3.0.2.3 Collection Policy

Discussion of a BHL-Europe collection policy was initiated at the beginning of this working period. A collection policy should help to clarify the question how we are going to organise the content contribution and prioritisation of content for BHL-Europe. The goal is to achieve a consensus with the content providers. This task was given to Michaela Hierschläger from LANDOE. First ideas from our content providers were collected on our BHL Wiki[16]. In order to reach a consensus on the BHL-Europe collection policy, a collection policy workshop was held during the Annual Review Meeting in Meise. Shortly after the workshop, a first version of the collection policy was sent out to the attendees of the collection policy workshop for review. Subsequently, the collection policy document was adapted to include feedback given by the partners. During the last three months the document was finalised and is now integrated into the Best Practice Guide. It will also be available as a separate document to support content provider acquisition.

In order to consider all issues and articles of the RBGE serial Sibbaldia for inclusion in BHL-Europe, all authors and co-authors were asked for permission. Although this process went very well, one issue was identified that is important enough to report on, as this issue also took up working time for RBGE and MfN colleagues in further investigation. One author withheld permission to include the article co-authored by him because he has concerns about the scientific quality of the serial. According to this author, a well-established peer review system and high quality editorial processes must be in place to consider a publication for BHL-Europe. Otherwise, BHL-Europe will not be perceived as a trustworthy source for scientific information due to low-quality scientific information. This is a particular problem for non-scientists and for the general public who may not be able to judge the quality of information. This issue raise a number of questions that may need discussion in larger groups. How BHL can estimate the scientific quality of a publication to be included in BHL and is it necessary? Should BHL refuse to ingest publications that are below a certain threshold for scientific quality? Should BHL establish an evaluation system for the scientific quality of publications or attach information to every digital publication to indicate the quality? We intend to further address these questions in order reach a consensus on the global quality assurance process to be established for BHL.

3.0.2.4 Management of the content acquisition process

The Annual Review Meeting in Tervuren/ Meise was used to provide detailed information related to content and to train content providers. Therefore one day was planned for workshops related to content provider's tasks. Three different Content Provider Workshops were planned to take place during this day as followed:

- Pre-Ingest
- Mapping
- GRIB

The *Pre-Ingest workshop* aimed at giving an introduction to the Pre-Ingest tool developed by AIT within WP3. Therefore, AIT was asked well before the meeting to provide a step-by-step

---

[16] https://bhl.wikispaces.com/BHLE_WP2_CoPo

tutorial in order to instruct all content providers how to use the Pre-Ingest tool for the ingestion of content after content was uploaded to the NHM server and to discuss examples regarding the Pre-Ingest file submission guidelines. Unfortunately, the Pre-Ingest tool was not available on the NHM servers at this time and thus our content providers could not be instructed how to use the tool to trigger the ingest process.

The *Mapping workshop* aimed at introducing the developed mapping tool - developed by NHMW – to the content provider in order to give them the opportunity to do the mapping of their metadata to the OLEF schema (i.e. the BHL-Europe schema) autonomously. NHMW provided a well-structured introduction to the mapping tool including live examples.

The *GRIB workshop* was a continuation of previous workshops and was aimed at giving an update on the GRIB development and introducing its functionalities and use. This workshop was held by Boris Jacob in cooperation with Andreas Krausz from the Head office of the common library network GBV (VZG). The content providers were informed about progress in the GRIB development and all functionalities available at that time were presented to the content providers.

Further, content providers were asked what their expectations were after the project life time. One major need identified is the maintenance of the communication in the community of natural history librarians. The adoption of a technical system with forum functionalities to function as a help desk, in support of an easy and fast communication between librarians, was suggested.

The management of the current content providers is an ongoing task within WP2. During this reporting period a lot of time was invested into communication with our current content providers and the WP3 team (NHM, AIT, and NHMW) in order to coordinate the upload of content and to give feedback for content providers regarding their content. Each content provider is different and has individual needs; communication with each of our content provider was and still is necessary. More up-to-date information about the content management can be found on the BHL Wiki[17] including documentation of the current status, in order to make the content management process more visible and to facilitate communications and future planning for our partners and the technical team. However, the management of the content upload during this time was very difficult due to laborious communications and changes in responsibilities for the Pre-Ingest tool. We anticipate that communications will become more efficient and effective after the Pre-Ingest tool is up and running.

3.0.2.5 *Management of the Pre-Ingest tool testing*

A first version of the Pre-Ingest tool was provided by AIT end of August 2011. Therefore preparation and management of the Pre-Ingest tool test was a major task during this reporting period. First of all, seven consortium partners were selected to take part in testing of the Pre-Ingest tool (LANDOE, CSIC, NATURALIS, RMCA, UH-VIIKKI, HNHM).

In order to receive representative feedback, content providers were selected based on their available uploaded content on the NHM server, their uploaded content structure (volume/article level) and their various backgrounds (non-technical/partly technical).

The testers were asked to test the tool by ingesting as diverse a selection of content as available in terms of the uploaded structure (volume/article level), the type of content (monographs, journals) and the file format (TIF, JPEG, PDF). The testing phase started on 29 August 2011.

---

[17] https://bhl.wikispaces.com/Content+_Management

For reporting of test results it was decided to use the BHL-Europe repository at the software GitHub - a web based service for software development projects - and label all issues, bugs and questions with 'Pre-Ingest and OLEF'. Testers were invited to GitHub in order to report existing bugs on their own or at least to be able to follow conversations and updates on the bugs/issues related to the test.

Content providers tested the tool several times by trying to ingest their own content to the BHL-Europe system. They were asked to describe any upcoming issues or inconvenient steps in as much detail as possible in order to be able to give the development team detailed information on how to improve the tool and make it more user friendly for our content providers. Finally, the testers sent their feedback directly to the WP2 leader who followed up unclear descriptions, summarised, structured and reported the recognised issues within GitHub.

In total, 23 issues were reported of which nine are critical. However, the test conducted with our content providers could only review whether Archival Information Packages (AIP) were created and whether the AIPs included the files described in the Technical Note of the Content Ingesting Process as intermediate steps of the Pre-Ingest process were visible for the user. From the users point of view it could not be evaluated whether the tool was working correctly as AIP folders were only sometimes created. In addition, the created AIP folders did not contain all described files, or else existing files did not contain any information. In some cases no AIP folders were created at all. By checking the created metadata files within the AIP folders it was noticed that the OLEF schema was not implemented in the tool and metadata has only been converted to MODS. The usability of the tool needs to be improved in order to not confuse the content providers. For this purpose suggestions were given to the developers. All reported issues and questions can be found in detail in the BHL-Europe GitHub repository[18].

After the test was finished, it was decided that NHMW will proceed with the improvement of the tool. The estimated deadline for finishing and improving the tool and reaching the stage of a productive system is the end of November 2011. After the development of the Pre-Ingest tool is finished, Michael Malicky from LANDOE will take on the role of the ingest manager and will give advice if issues occur during the Pre-Ingest process.

3.0.2.6 Assist partners in implementation and evaluation of scanning operations; discuss and distribution of data standards and specifications; attracting new content providers

This chapter gives an overview over work done in this reporting period regarding analysis and preparation of scanned material and the corresponding metadata, structured in paragraphs by individual partners.

BnF worked on the selection of documents for digitisation for BHL-Europe and checked the quality of already-digitised pages, in particular the *Annales de la Société Entomologique de France* ranging from year 1938 to 2002. In total 29,523 pages were quality checked. Furthermore, the data transfer to BHL-Europe was technically supervised. The transfer of BnF content to BHL-Europe had to be initiated twice as errors occurred during the transfer. According to the check-up of the already-transferred content by BnF, one directory is still missing after the second transfer. However, because of the immense data volume BnF cannot find out which directory is missing by using their FTP account and is currently waiting for the

---

[18] https://github.com/bhle/bhle/issues?labels=Pre-Ingest+and+OLEF&sort=created&direction=desc&state=open&page=1

completion of the ingestion of their content to the BHL-Europe portal in order to check the availability of content online. Subsequently, they will be able to provide the missing directory to BHL-Europe. In total 208,308 pages were transferred to the BHL-Europe server, including Annales de la société entomologique de France and Annales des sciences naturelles.

BnF is currently preparing their second delivery of content to BHL-Europe for autumn this year. BnF also supervised the technical transfer between Gallica and BHL-Europe.

CSIC worked on metadata for already-digitised content, uploaded several officially published documents and documents considered as "grey literature":

- "Unidad de Biodiversidad", a book on biodiversity that was produced by the FECYT (Spanish Foundation of Science and Technology) for teachers at the high school level.
- Celso Arevalo, 1929, "La vida en las aguas dulces", the first limnological monograph in Spanish.
- Three volumes of the "International Symposium and First World Congress on the Preservation and Conservation of Natural History Collections".
- "Iberus", vol. 22.2 (as it was not available in BHL-US).
- Particular mention deserves the completion of ingestion of the Spanish Malacological Journal "Iberus", which was only partially available in BHL-US. Now, taking BHL-US and BHL-Europe together, the journal is complete as following parts of Iberus were provided by CSIC: vol. 24.2, vol. 25.1, vol. 25.2, vol. 28.1, and vol.28.2.
- "Memoria presentada para optar a la plaza de conservador del Museo Nacional de Ciencias Naturales", the first natural history monograph on museums curators in Spanish (regarded as "grey literature", A.G.Valdecasas, 1984).

Furthermore, CSIC got in touch with following copyright owners:

- Maria Soledad Vicente Rosillo, co-author of the five volumes of "Catálogo de los fondos especiales de la Biblioteca del Museo Nacional de Ciencias Naturale".
- Dr. Borja Sanchiz was contacted in order to digitise "Primer Congreso Español de Biogeografía", 1978, Universidad Complutense de Madrid, a Ph.D. thesis on Biodiversity for BHL-Europe.

To date, no final decision was received in any of these cases.

Conversations with Dr. Isabel Izquierd who is in charge of the digitisation of the Spanish entomological journal *EOS* have led to inclusion of the journal into BHL-Europe. The journal – not longer published – contains about 30,000 pages of entomological information in Spanish and other languages.

After numerous checks, reviews and feedbacks a template (Excel sheet) for content indexing was validated by the MNHM team and by their digital library provider (the vendor of the portal). Then MNHN sent the template to their digitisation and indexing service provider. First, a small batch of 50 books was ordered for tests and adjustments. The main difficulty is that each collection of books or journals has its own specifications (for example two different rules of numeration for some serials, missing TOC, etc.). Thus MNHN has to note the particularities of the collection to the service provider each time, before the service provider is able to start the work. The tests ended in April and the production was started in May. MNHN hopes to finish the indexation at the end of December 2011.

In May, MNHN updated and validated their OAI API for Gallica. This API can also be used by BHL-Europe to harvest the metadata of the ~200 books that are already available at the NHM server. 200,000 pages in TIFF format and 200,000 pages in JPG, recently scanned, are ready to be uploaded to BHL-Europe. However for the moment, they only can provide metadata in UNIMARC on journal title level. Besides, MNHN has started to receive the template for the content indexing. This is an ongoing process.

In September MNHN finished to upload the first batch of volumes which are already online on their Gallica portal. A detailed list of the titles can be found hereafter:

Title	Years available	Pages
Annales du Muséum national d'histoire naturelle	1802-1813	11,500
Archives du Muséum d'histoire naturelle	1839-1858	5,600
Bulletin de la Société botanique de France	1854-1935	64,000
Bulletin de la Société des naturalistes parisiens	1904-1932	1,200
Bulletin de la Société philomathique à ses correspondans	1791-1797	5,800
Mémoires du Muséum d'Histoire Naturelle	1815-1832	10,200
Nouvelles annales du Muséum d'Histoire Naturelle [Autre titre : Recueil de mémoires]	1832-1835	2,100
Nouvelles archives du Muséum d'histoire naturelle	1865-1914	18,000
Total		**118,400**

MNHN is waiting on the ingestion of their content to the BHL-Europe portal in order to get a feedback how their content is displayed in the BHL-Europe portal.

In April, the analysis and description of a first batch of selected new content finished (Flores, Index seminum, History of the Botanical Garden and of the MNHN). 28,500 new pages are ready to be scanned in the future, which however still depends on funding.

Part of the new digitised content was already uploaded to the BHL-Europe server. Furthermore, a copy of the Excel template for the indexation was also sent to the technical team of BHL-Europe, in order to evaluate what may be exploited under the metadata format of BHL-Europe. In September, MNHN uploaded further 87,000 pages, more than the agreed for November. MNHN will continue the uploading, until reaching the total of the 232,000 pages planned for April 2012.

A detailed list of the uploaded content can be found here:

Title	Years available	Pages
Archives du Muséum d'histoire naturelle séries 6 et 7	1926-1969	7,870
Bulletin du Muséum National d'Histoire Naturelle (toutes sections)	1895-1970	51,100
Mémoires du Muséum National d'Histoire Naturelle	1936-1950	11,560
Notulæ Systematicæ	1909-1980	5,205
Adansonia	1961-1980	11,370
Total		**87,105**

MNHN was contacted by some other institutions about its participation in the BHL-Europe project and its other digitisation projects, such as the libraries of University of Rennes 1, the École des Ponts, Paris Tech and the Interacademic library of Health in Paris. Communications between the WP2 lead and the University of Rennes 1 were successful by the end of the reporting period.

NATURALIS started a new large-scale digitisation project. At the start of this project a lot of attention was paid to selecting the right materials for digitisation, checking the availability in BHL (Europe), and setting up the work processes so that the results can be uploaded to BHL-Europe.

Selection of works was carried out in close contact with BHL-Europe partner UGOE. Their funding for digitisation (AnimalBase) has ended, but they still have a large list of biodiversity

literature which they requested to be digitised. NATURALIS has checked UGOE's list against their own collection. Eventually, 59 titles were selected from UGOE's list that now will be digitised within their new project.

NBGB tested the FTP upload to the BHL-Europe server. Further work was image quality control and checking metadata which are currently in process. They also searched for the rights owners of several books they digitised. With the help of their researchers they continue searching new content for digitisation for BHL-Europe. Their first data was uploaded to the BHL-Europe server at the end of September. NBGB also makes regular backups of all images scanned by the three Belgian institutions (RMCA, RBINS, and NBGB).

RBINS and their library personnel selected digital publications for provision to BHL-Europe and defined priorities. Furthermore, they uploaded ~ 2,000 pages as test content to BHL-Europe. The RBINS metadata is not in a standard format and thus the preparation of the mapping takes more time and is still in process. RBINS also checked metadata for PDF-files. Most of their content consists of TIFF files with suitable metadata, but they also have a lot of PDF files which do not have consistent metadata.

In April 2011 *Notes from the Royal Botanic Garden Edinburgh v.1-46 1900-1990* and *Flora of Bhutan Parts 1(1)-3(2) 1983-2001* were uploaded to the BHL-Europe server by RBGE. This content was harvested by Europeana, and quality checked by RBGE.

Following discussions and a meeting with staff at the National Library of Scotland (NLS), arrangements were set in place for staff operating the Internet Archive station at NLS, to create digital versions of selected books from the Library. The following titles were scanned and made available:

- Sutherland, James. Hortus medicus Edinburgensis… (1683)
- Alston, Charles. Index plantarum… in horto medico Edinburgensis (1740)
- Hope, John. [Dr. Hope's elementa botanici] (178?); Alston, Charles. A third dissertation on quick-lime and lime-water. (1757)
- Fletcher, H.R. & Brown, W.H. The Royal Botanic Garden Edinburgh 1670-1970. (1970)
- Alston, Charles. A dissertation on botany. (1754); Braithwaite, M.E. and Long, D.G. The botanist in Berwickshire. (1990)
- The Botanical Society of Edinburgh 1836-1936. (1936). Permission was granted by the Berwickshire Naturalists' Club and the Botanical Society of Edinburgh, to include the last two titles. These books will be provided to BHL-Europe.

Furthermore, some work on rights clearance was done, because of permission from external organisations to arrange for digitisation of three books using the Internet Archive station at National Library of Scotland. This was not charged to the project.

In early August 2011 RBGE problems with the content display in Europeana were discovered. The WP2 leader was informed about this problem, and developers were made aware of this situation. The issue was further reported in GitHub[19] and also mentioned during the Public Core Review Meeting end of August. In October a short-term solution was implemented until the new book viewer is finally ready for use.

LANDOE increased the amount of content delivered to Europeana. To date they supply slightly more than 3,500 books to the Europeana portal. The metadata for Europeana is still provided through an OAI-PMH interface.

---

[19] https://github.com/bhle/bhle/issues/200

LANDOE started the content upload to the BHL-Europe server and nearly finished this task within the reporting period. This initial upload took more than three weeks. All digital books available in Europeana have now also been provided to BHL-Europe for ingest to the portal. Missing metadata was delivered by August 2011. In the meantime they digitised and metadata-enhanced 600 additional books. This material – including more books in the current digitisation process - will be provided for the BHL-Europe Portal and Europeana.

Michaela Hierschläger spent time enhancing the metadata of further LANDOOE content, especially for the digital books for their next and final upload to the BHL-Europe server within the project period beginning of 2012.

Following bulk digitisation of HNHM journals, HNHM is working further on the compilation and key-in of the bibliographic journal article level metadata sets (xml format). The splitting of scanned journal volume to article level was continued. Metadata preparation, quality assurance and database conversation of the following journals' data were completed:

- Anthropologica Hungarica,
- Crania
- Parasitologica Hungarica.

By the end of July Zsuzsanna Forgács's contract ended. Therefore the main task was to fill in smaller gaps in former data conversion in order to complete the given journals. The metadata compilation of Acta zoological hungarica – a joint publication of the Hungarian Academy of Sciences and the Hungarian Natural History Museum- was also started.

HNHM is working on a continuous update with recent volumes of journals in order to submit more new content as requested by users. Therefore HNHM also continued the metadata checking/improving of journals on article level.

Because of an internal change in their workflow all formerly uploaded content had to be resubmitted. In order to facilitate this process - HNHM is providing a huge amount of content on article level - scripts were created to extract data from their DB (MySQL) to MARC21 format (XML). A further shell script was developed for creating the necessary library structures and facilitating the data upload to BHL-Europe.

A further task was related to the conversion and quality check of monographic metadata of their Central Library System (HUN-TÉKA). As a result their bibliographic metadata set of their digitised monographic holdings increased from ~50% to ~80% and thus will also be available for BHL-Europe (expected to be upload by the end of October). The raw data input of the new bibliographic metadata set was covered from a national fund. A new digital asset management system (JADOX) was also purchased and installed from the above fund. The new system helps HNHM to improve the workflow for the data upload to BHL-Europe.

Live demonstration and negotiation started with the regional natural history museums and the Hungarian Biological Society in order to incorporate their biodiversity literature into the BHL-Europe corpus. As soon as official agreements are available, HNHM will provide further content to BHL-Europe.

In collaboration with NATURALIS (Tom Gilissen) a list of important historic zoological books to be scanned by NATURALIS digital library was extracted from AnimalBase (UGOE) resources. These were books that had been scheduled for digitisation in Göttingen, Germany, but which could not be scanned either due to technical reasons or due to lack of funding. The list mainly derived from detailed analyses of Sherborn's Index Animalium published between 1902 and 1931, where all literature was listed that contained new taxonomic names in zoology. It was checked if the works were digitised elsewhere and, if not, they were included in the list. The first part of this list contained 140 books published before 1780 and was sent to NATURALIS in June. The further preparation of the list is a currently ongoing process.

RMCA is continuously working on selecting literature to be scanned, coordinating the scanning process and updating their local Scanlist[20], which is a common list for the three Belgian BHL-Europe partners: RMCA, NBGB, and RBINS.

From April to June more than 15,000 additional pages of biodiversity literature from the zoological department were digitised in the framework of other projects and in-house activities. All files were controlled and renamed on page level. Furthermore, metadata processing is one of the important tasks in preparation of the content for ingestion. This includes the input of new data, the enrichment of catalogue data with additional information and the extraction of metadata from the LIBIS library system.

Public domain literature is continuously uploaded to the NHM server. The first 52 items (~24,000 pages) were so far harvested by Europeana. Routine copyright clearance procedures are also one of RMCA's tasks. In the last three months RMCA was able to obtain three signed BHL-Europe agreements.

Between July and September 2011 about 12,000 additional pages of biodiversity literature from the zoological department were digitised in the framework of other projects and in-house activities. All files were quality checked and renamed on page level. Metadata processing is an important tasks in preparation of the content for the ingestion. This includes the input of new data, the enriching of catalogue with additional information, and the extraction of metadata from the LIBIS library system.

In relation to content acquisition RMCA has discussed with the publishing service of RMCA the possibility of providing 12 issues of the Journal of African Zoology (JAZ) to BHL-Europe. The agreement was achieved and it was decided to start copyright clearance procedures for all authors. Necessary steps are already undertaken.

During this reporting period UH-Viikki checked and processed the first part of digitised content for inclusion into Europeana. Currently 30 items from UH-Viikki are available in Europeana. In May the second part of publications were handed over to their digitisation provider Vaasan Micro Copy and were received in June.

UH-Viikki received the signed BHL-Europe Licence Agreements from two Finnish biological societies, Biological Society of Finland Vanamo and Societas pro Fauna et Flora Fennica. The agreements allow UH-Viikki to digitise publication series of these societies until 1940. The licence used in these agreements is Creative Commons Attribution-NonCommercial-ShareAlike 3.0.

In September the third delivery of digitised content was received from digitisation provider Vaasan Micro Copy. The total number of digitised pages is currently approximately 38,670.

UH-Viikki received the signed BHL-Europe Licence Agreement from the Botanical Museum which is part of the Finnish Museum of Natural History. The license covers most of the series Norrlinia which is published by the Botanical Museum.

BHL-Europe has finally taken off in Denmark. Two large journals, Botanisk Tidsskrift and Dansk Botanisk Arkiv with ca. 42,000 pp. were scanned, checked and uploaded to the NHM server. UCPH continues with the largest Danish (natural history) journal, Meddelelser om Grønland, comprising ca. 100,000 pages in 206 volumes. The copyright was given to a commercial editor, but in agreement with them UCPH can digitise material created before 1940; when the agreement is finalised they will also be able to scan more recent issues. As the last work in the first round of digitisation, UCPH will scan the largest Danish zoological journal' *Meddelelser fra den naturhistoriske Forening i Kjøbenhavn'* (ca. 50,000 pages).

---

[20] http://193.190.223.46/wiki_ext/index.php/BHL

With these completed they will continue to the next issue on their agenda for digitisation of the Danish biodiversity literature, i.e. works including (many) diagnoses.

UCPH has also received funds to employ a technician for three months, which will speed up things greatly and allow them more time to apply for a larger amount.

Furthermore they participated in two ½ day meetings with the Danish Europeana group, and one more is scheduled at the end of October.

UCPH is ready to inform their colleagues at the museum about BHL-Europe for the first time, and is now waiting for their own contributions to be online in BHL-Europe, so that examples and criticism can be exemplified with their own material.

After MfN purchased a scanner in spring 2011, the museum will now hire a scan operator to actually start scanning content also for BHL-Europe. The BHL-Europe team at MfN was involved in the identification of suitable candidates for the position. From November onwards, the scanning workflow will be established following the BHL-Europe Best Practice Guidelines and then digital content will be produced at MfN.

The new content provider GfBS (Gesellschaft für Biologische Systematik) has uploaded Volume 1 to 9 of its journal Organisms, Diversity and Evolutions (ODE), formerly published by Elsevier to the BHL-Europe servers. The metadata is available on article level. There is ongoing communication with the former editors of the ODES (Electronic supplement of the ODE) in order to get the PDF files and associated metadata. Further information will follow at beginning of 2012. Furthermore there is an agreement that the current publisher of the journal Springer will provide FTP access to the GfBS board in order to allow download of ODE's open access articles. All articles published in ODE by Springer will be made publicly available (open access) after two year and will thus be uploaded to BHL-Europe on an annual basis.

3.0.2.7 Assist partners in implementation and evaluation of scanning operations

In order to better understand and evaluate the software 'Goobi', UBER set up a local Goobi test system, where individual workflows are described for a digitisation process to help the content providers (see also Goobi Testsystem[21]). UBER continued exploring other new software tools, which can be used for the digitisation process, and evaluated them using the Best Practice Guidelines such as 'Fedora', 'eScidoc' and smaller Tools.

3.0.2.8 Best Practice guidelines and standards

One task within WP2 is assisting partners in implementation and evaluation of scanning operations and to deliver a first draft of the Best Practice Guidelines and Standards. This delivery aims to give a first standardised guideline for Europe. The approved best practice guidelines and standards should be understandable to every person working with it. They shall guide new and current contributors in a simple and clear way through a digitisation workflow from either the print or the digital version of an item to the digitised representation of that item in the BHL-Europe portal. The guide will describe the minimum required standard for BHL-Europe and where possible specifying the ideal standard. The first version of this guide was due at the end of the second project year, 30 April 2011. During this reporting period the first version was finalised and handed over to the project coordinator. However, the work on the Best Practice Guidelines continued. The preparation for the final version started during this reporting period in cooperation with our partner LANDOE. A third workshop was organized and held during the Annual Review Meeting in June. All contributors for the first version wereinvited to take part in this workshop and give feedback

---

[21] http://goobi.cms.hu-berlin.de/newpages/Main.jsf

on the first draft and how to move forward based on the final document. Contributors were instructed how to review the chapters for the final version in order to facilitate reading by users. Further missing information were identified and contributors were assigned to the respective chapters.

After the workshop several partners were asked to contribute to the final version of the Best Practice Guidelines and standards.

It was also agreed to provide three different versions of the Best Practice Guidelines and standards to support readers with no background of the subject matter who expect more detailed information, readers with background who expect only the most important facts and readers who only would like to get a quick overview of the workflow. Therefore in addition to the long extended version (called SPEE - **Sp**ecial **E**xtended **E**dition), we will produce on the basis of the detailed version a shorter version (called MIDI) including only the most important facts. The extended detailed version will also include extensive examples, explanations and links to further references in order to allow the reader to follow step by step. Based on the shorter version (MIDI) we will produce a very brief document (called MINI) to illustrate only the workflow steps.

Following documents are planned to be available for the Best Practice Guidelines and Standards until the end of the project:

➢ SPEE - Special Extended Edition

Purpose: very detailed version including extensive explanations, examples and links to further references.

➢ MIDI

Purpose: shorter version on the basis of the long version. No examples and detailed explanations included.

➢ MINI

Purpose: Gives only an overview of the workflow.

During this working period the detailed and final version (SPEE) and the shorter version (MIDI) were drafted. For the detailed version the Collection Policy was updated in cooperation with our partner LANDOE. Conversations wereconducted and further assistance was given on how to produce the shorter version. The already available parts of the MIDI version werereviewed and feedback was given.

The SPEE and MIDI version are in progress, however the Pre-Ingest part is still missing as it is dependent upon the development and final release of the tool. Furthermore, the GRIB chapter was delayed due to technical issues at the VZG. However the GRIB part is currently in the process of finalisation. More detailed information about the workshop and outcome can be found on the BHL Wiki[22].

The Best Practice Guide has already been used by our consortium Partner NATURALIS for setting up the work process for a new digitisation project. Although the writing of the guide is still in progress it already proves to be very useful for our partners. They have for instance used it to define the set of metadata that needs to be the result of the digitisation. NATURALIS also used it to introduce the new staff (hired for the digitisation project) to the current practice of digitisation and the best practice to make sure that the digitised content can be uploaded to BHL-Europe.

NBGB already described their internal digitisation workflow as an example for the document and contributed in collaboration with RMCA to the digitization workflow chapter for the Best Practice Guide.

---

[22] https://bhl.wikispaces.com/BHLE_WP2_BPG

RMCA staff was involved in writing parts for the Best Practice Guide, dealing with collaborative scanning workflows and collection management in Belgian natural history institutions, as well as with the use of the GRIB for scanning management.

CSIC worked on the collection retrieval chapter.

3.0.2.9 Expansion of the content provider network

Attracting new partners until the end of the project is another major task within WP2.

During this reporting period several potential content providers were contacted for acquiring new content to BHL-Europe. Communications with prospective content providers requires a lot of time and preparations in order to answer various occurring questions, for instance questions regarding IPR, metadata, OCR or the MoU. This is an ongoing task.

BHL-Europe welcomes the Université de Rennes 1 as a new content provider to this project.

Furthermore, the 'BHL-Europe in a nutshell' document was translated into German by our partner LANDOE in addition to the already existing versions in Spanish, English, French and Russian. Further information about content acquisition, e.g. the 'BHL-Europe in a Nutshell' document, can be found on the BHL Wiki[23].

The BHL-Europe project office at MfN is involved in planning the full implementation of the newly purchased scanner. As the scanner will not only be used for books but for many other applications as well, priorities need to be defined.

In May 2011, a digitisation project of the university library Frankfurt/Main was approved to be funded by the German Research Foundation (DFG). BHL-Europe was significantly involved in the project design and the initial planning process through the PCO. It is a great success for BHL-Europe that funding was received. MfN is currently discussing the final project plan and supports the digitisation project through BHL-Europe WP2.

3.0.2.10 Coordinate with Europeana

Coordination with Europeana took place with Lizzy Komen to ensure that new content will be ingested to Europeana in a timely manner. An updated ingest plan was sent to Europeana including an additional estimation per item to facilitate the planning of Europeana's ingest team.

There was an update of already-ingesteded BHL-Europe material in Europeana of the following collections (search term in Europeana: europeana_collectionName:XXXXX*) such as BHL (US), collection number 08701 plus 3890 items; Germany: collection number 08708 plus one item; Denmark, collection number 08709 plus 75 items. The total amount of BHL-Europe items currently in Europeana is 100,561 items.

There is ongoing involvement of BHL-Europe in the work of the Europeana Council of Content Providers and Aggregators (CCPA). To date, already 12 members of the BHL-Europe consortium have signed up for the CCPA. Some of them are also involved in the five working groups and actively contribute to the success of these groups. BHL-Europe members are particularly involved in the Legal working group and the Technical working group. Thereby BHL-Europe was actively involved in the work on the new Europeana Data Exchange Agreement. The BHL-Europe project coordinator, who is also secretary for the CCPA and liaison officer for the Technical WG, supported the setup of the WG and advised the chair on the goals and aims of the CCPA as discussed during the officer meetings. The work plan for the Technical WG was adopted in June during the WG meeting alongside the OpenCulture[24] conference in Birmingham. Based on this work plan, 10 white papers will be

---

[23] https://bhl.wikispaces.com/Content+_Management

[24] http://www.openculture2011.org.uk/

delivered by the WG during the CCPA plenary meeting. A conference call was held on 30 September with the core WG members to revise the current progress with the white papers. Seven out of ten white papers are drafted to date and more will follow soon.

The last CCPA officer meeting was held on 24 October 2011 in The Hague (KB). It is planned to rename the CCPA and elaborate on the strategic direction of the CCPA to better engage with the Europeana network. This will be presented in more detail on the CCPA plenary which will be held alongside the DISH[25] conference on 6 December 2011. During the plenary, the work of the first year of real work will be presented and discussed and plans for 2012 will be made. Registration for this plenary involves no costs. Partners can also benefit from a lower conference fee for the full DISH conference as a CCPA member.

3.0.3 Status and progress of activities – WP3

In this reporting period, the technical team held several workshop sessions (Tervuren and Prague) to demonstrate and review the progress of the development work and also conclude on various aspects relating to our adopted technologies and functional architecture for the final BHL-Europe system.

3.0.3.1 BHL-Europe Annual Review Meeting, Tervuren

Lee Namba (Atos), Lola Obajuluwa (NHM), and Jana Hoffmann (MfN) organised two workshop sessions on the Portal Development:

The first session consisted of the Evaluation and Review of Existing Portal Features that had been implemented so far. During this session AIT (Bernd Sproger) provided us with a step by step walk-through of the existing portal features and their functionalities in the AIT work environment. Bernd demonstrated the BHL-Europe portal core features which were further implemented and refined e.g. search queries can be saved, search results can be viewed in different layouts (list, grid), different XSL transformations can be applied to the results view, highlighting of search terms in the results views. This was the first opportunity to gather initial feedback from both users and developers that attended the workshop. At the end of the first session the demonstrated portal functions, especially Search and Retrieval, were evaluated positively and the implemented features are generally consistent with the description in the CoR (Catalogue of Requirements). Three versions of the Advanced Search were implemented:

- A Google-like search, where the user can search in the categories "with all words", "with exact phrase", "with any word" and "without words"
- A Europeana-like search, where the user can dynamically add search fields and concatenate them with logical operators.
- A code word search, where the user can enter code words in order to make the search query.

In the first 2 searches there are also elements for selecting the metadata fields (ESE) to search in. The searches are configurable via a user configuration menu.

The 2nd session was used to review the Priority 1 features yet to be implemented for M28 and review/prioritise the remaining work for the portal implementation.

On the last day of the Annual Review Meeting, the TechGroup held a Closed TMB session. During this session various topics were covered and decisions made on ongoing work within the Development process. Areas of focus include an update from Atos (Lee, Hengdi) on the Shared Development Environment (GitHub) which is now up and running. This software

---

[25] http://www.dish2011.nl/

environment strategy and planning tool will allow an organised and harmonised shared development environment between our multiple development partners. Atos presented and trained the consortium members on repository administration, issue tracking organisation, milestones, and development and testing workflows.

This included a demo on the use of GitHub as a source code repository which will also allow for a distributed development and collaboration between the various team. Furthermore, now that we are at the development phase of the project in terms of the portal, all Issues/Bugs were migrated to the Issue Tracker. The Issue tracker will be used as a tool for tracking Issues, such as features, enhancements, bugs, etc. Other topics covered include Kick-off discussions to determine the scope of work required for our next deliverable D3.8 (M30) as well as planning for the Public Core Release meeting which was scheduled to take place at the end of August in The Hague.

During the Evaluation and Testing Workshop, Atos gave an overview of the different types of software testing processes and team organisation to the consortium. Best practice recommendations were also given of testing processes and organisation which drove the creation of a test manager (Jiri Frank) and test lead roles (QA team) as well as training consortium members on testing methodologies.

3.0.3.2 Public Core Review Meeting, The Hague

The Public Core Release which was previously planned as the next major milestone for WP3 (M28) has now been rescheduled for November 2011.

This was largely as a result of the many technical and management challenges involved in achieving full automated integration and deployment of the various system components through the shared development environment. We aim to deploy a fully working development version of the BHL-Europe system in time for the November release.

The Public Core Review meeting was held in August 2011 in The Hague which consisted of various workshop sessions including:

The Portal Review Session – the QA Manager (Jiri Frank) presented the current status of the developed portal features. This was an opportunity for the testers to give their feedback on the quality of the implemented features as well as to evaluate the overall status of the portal. Excellent progress was also made in planning for the production server environment, revising the BHL Wiki directly during the session to deliver the planned layout. We determined which components could be treated as the operating system, and updated automatically, and which would be maintained via GitHub/Jenkins.

The developer's workshop enabled focus on the upload of code, the BHL-Europe theme and test data to GitHub to work toward the release scheduled for November, and enabled the developers to resolve a significant number of integration issues constraining our ability to deliver further features.

Overall, good progress was made in preparation of the PCR scheduled for November. In order to move forward with the development work it was agreed to separate the tasks into logical groups. Tasks within these groups have since been allocated to WP3 partners with agreed delivery dates and funding. We are currently in the final stage of formally agreeing the work and funding that needs to be distributed amongst partners. This process will ensure that ongoing development work from now until the end of the project will be attainable and fundable.

3.0.3.3 Development of Pre-Ingest Tool

The Pre-Ingest system was prepared for installation on the NHM servers. The Pre-Ingest system implements the processing chain for each digital object. The digital objects are uploaded to the server by the content providers according to the submission guidelines delivered with Pre-Ingest.

The system includes a browser front-end to select the provided folders and submit them to the processing queue. Objects in the queue can be monitored while they are being processed.

During processing of submitted content, Tesseract is used to extract text from the scanned books. JHOVE is used to extract digital metadata. External services are called to allocate a unique ID for the object and to extract taxonomical terms. Pre-Ingest generates an information package containing the OCR texts, the extracted metadata and a METS file suitable for insertion into a Fedora archival system.

3.0.3.4 Collaboration on OCR improvements (IMPACT)

In April 2011 we started to produce a test set for ground truthing of page images through the EU-funded IMPACT project. The test set selection was based on the IMPACT guidelines for building a test set. BHL-US content was used for this test set as their content has the variability of material we need for the process (e.g. languages, font types). The BHL also has experience with OCR based on previous projects and initiatives to work on the challenges with high quality OCR. Henning Scholz and Chris Freeland were the main actors on our side to prepare and finalise the test set in collaboration with IMPACT. Francisco Welter-Schultes is now the liaison for the QA of the ground truth process.

The process has already started and involves the following steps:

- Conversion of the supplied JP2 images to TIFF required for OCR and long-term storage within the IMPACT infrastructure. This brought one particular problems, as the JP2 images had no resolution information included, which is necessary for the process. Some manual work was necessary to finalise the conversion.
- OCR of the derived TIFF files to produce PAGE.XML, the format used for ground truth and evaluation in IMPACT.
- Assigning unique IDs to both images and initial OCR result.
- Ingest of images and metadata into our central repository.
- Selection of a small set of representative images to approach potential service providers for the rekeying process (tendering process).
- Sending out a tender to a number of qualified service providers. This call for tenders is expected to be finished in September 2011.
- Production of ground truth files by an external service provider in batches of ~500 pages per week. As we have provided 5,000 pages, the process is expected to take about 10 weeks.
- QA on ~10% of every batch to finally achieve an accuracy of 99.95%, which means 5 characters wrong on 10k pages. For this process we received software developed within the IMPACT project for testing purposes.
- After one batch is finished, IMPACT will start running evaluations parallel to the ground truth production. IMPACT will develop an evaluation workflow with us and for us. We can also run our data against the workflows of pure library projects to compare the outcomes (maybe we need a domain specific workflow).
- As a result we get lots of statistics and we know what combination of IMPACT tools give the best results for us. We can also compare this with previous experiments run by BHL-US.

- We will prepare an OCR session during the fall meeting in Brussels with IMPACT to show first results and discuss options to improve or adapt the workflow.

As the IMPACT tools are not open source and not freely available, we cannot simply use them. However, we may be able to charge costs to a follow up project of BHL-Europe and intensify the collaboration with IMPACT. Although the IMPACT project will end this year, it will continue as a Centre of Competence in the future.

3.0.3.5 Integration of Catalogue of Life

In June Frank Bisby (Sp2000) visited Heimo Rainer at the Vienna Natural History Museum to discuss the request for Species 2000 to provide the Catalogue of Life to both the BHL-Europe and OpenUp! projects.

The principal discussion was on the structure and usage that can be made of the Catalogue of Life synonymy, to provide synonymic indexing and a taxon user-view for both the BHL-Europe and OpenUp! portals. Furthermore different ways of delivering the Catalogue of Life – from imported editions to Webservices were discussed as well as the new Webservices being implemented for the CoL in the 2nd batch of services in the 4D4Life project, due for testing at the end of June 2011, and for public release in October 2011. These will be made available for test during the summer. The nature of the license or permission to use the CoL that will be provided by Species 2000 to BHL-Europe and OpenUp! was clarified another topic. The permission will be to use the CoL as an indexing and annotation tool, but will be restricted to this use and will not extend to publishing the CoL or derivative works.

3.0.3.6 Digital Preservation Summit 2011[26]

Melita Birthälmer and Francisco Welter-Schultes attended the first summit on digital preservation held 19-20 October 2011 in Hamburg, Germany. 120 IT experts from all over the world discussed the current challenges of digital preservation. Of many issues raised at this meeting, a few points stood out and were often repeated. Public awareness of the need of long-term preservation of digital content is still very low, resulting in permanent budget pressure for all involved institutions. A too-narrow frame provided by the legal rights backgrounds presents a serious obstacle to effective preservation of digital content in the public and the scientific domains. Except for image content, for which TIFF seems to be widely accepted, there is no common acceptance on reliable file formats that can or should be used for storage of digital content. Technical solutions for digital preservation are currently in a transitional stage between initial pioneer work and final production.

3.0.3.7 Development of the Portal System

A preliminary Drupal 7 theme was created based upon the official BHL-Europe Drupal 6 theme. Drupal 7 was setup to use different localisation and internationalisation modules thus facilitating the implementation of different languages. Therefore all seven necessary languages prepared for the Annual Review were incorporated.

The current available version of Islandora was not compatible with Drupal 7 and therefore Atos (Lee Namba and Hengdi Zhang) have rebuilt some of the essential modules (the Islandora digital repository module and book module) to make the basic functions available on our Drupal 7 portal and have replaced Fedora item bootstrap class with Fedora API module for Drupal 7.

---

[26] http://www.digitalpreservationsummit.de/

Atos have also developed Webservices for downloading books in various formats e.g. PDF pages and books with a range, OCR packages of books, images and books. These Webservices enable the download of different formats of ingested pages and books from the Fedora repository.

Furthermore, Pentaho Drupal Integration was tested and implemented using HTTP Post with METS to create new Drupal nodes in the portal. Data harmonisation for RBGE, RMCA, CSIC, and UH-Viikki was developed, specifically pre-processing steps to harmonise the structure of submitted files. This pre-processing is done on a per-collection basis using various bash-scripts and XSL transformations. Also, the IA BookReader for preview purposes of created archival information packages was created. An image conversion backend was created using Perl and PHP, the PHP backend is currently in use. New tests needed to be conducted using Tesseract 3 since old tests and development plans were based on Tesseract 2.

Pentaho Drupal was used to create ESE mappings for metadata. It has also been used to create transformations specifically needed for OAI-PMH. During development it proved itself to be a tool which should be used for transforming file/folder structures so that submissions adhere to the pre-ingest file submission guidelines.

3.0.3.8 Maintenance of the portal

Andreas Kohlbecker (BGBM) handled the planning and setting up of the server infrastructure together with Lee Namba and Chris Sleep. He was also responsible for the installation and configuration of the Apache Solr server. This server is responsible for indexing the whole content stored in the BHL-Europe system and thus is the technological base for making the content accessible from the portal via search and browse functionalities.

He started working on binding the portal to a new Solr search index which is now updated by a special component of the Fedora system, called Gsearch. The index is thus updated every time new documents are ingested or when existing documents are updated.

3.0.3.9 Implementation of BHL-Europe features and preparation of the PCR Meeting

The development of the Book Viewer: Atos took over the development of the Book Viewer and developed functionalities such as table of contents, search in document, add pages to download, download in various formats, and download in various resolutions. Various user interfaces for book viewers were evaluated such as BHL US and BHL Australia. Some code from BHL Australia was reused and adapted to the requirements of BHL Europe. The book viewer was also integrated into the BHL Europe portal.

Further development of the search functionality of the portal: Andreas Kohlbecker (BGBM) was continually working on preparing the improvement of Unicode support of the portal (correct handling of diacritic characters). He also investigated options to allow intelligent fuzzy search and how to improve multilingual features in the search facility.

Other modules that were developed for the portal by NHMW:
(1) The VIAF[27] (Virtual International Authority File) Webservices.

It provides Drupal functions which will be used to amplify search terms. The advantage of this module will be the ability of the user to search for a single author name and at the same time expand the search to all alternative spellings of the same author. This will make it possible for the user to retrieve all results authored by the same person even if alternative spellings were used.

---

[27] http://www.oclc.org/developer/documentation/virtual-international-authority-file-viaf/using-api

(2) In addition a module which invokes both the PESI[28] (Pan-European Species directories Infrastructure) and the CoL[29] (Catalogue of Life) Webservices was developed. It provides Drupal functions which can be used to translate common names into their scientific names. The module was further extended in order to provide synonyms searching for scientific names. This will allow the user to search both using common names and scientific names and automatically have the search expanded to all synonyms provided by the services. Again this will yield a larger result set for the user and provide hits including synonyms, scientific names and common names.

For the November Portal Release, these developed Drupal modules will be integrated into the portal. The search functionality will be leveraged to invoke those modules.

3.0.3.10   Portal Testing

Andreas Kohlbecker and NMP organised the Evaluation and Testing Workshop during the BHL-Europe Annual Review in Tervuren and Meise, Belgium. This workshop was to decide on the various testing approaches (including the technical testing) and identify responsibilities. The minutes can be found on the BHL Wiki[30]. Andreas also requested quotations from usability testing agencies and evaluated them.

Jiri Frank (QA Manager) and Jana Hoffmann (MfN) established the testing group and created a testing workflow to do the initial testing and provide feedback for the developers. The main focus for the developers and the QA team in the first testing phase was to ensure that the Priority 1 Core Feature functionalities are tested and fully functional ahead of the Public Core Release meeting planned for the end of August.

We originally planned to kick off the testing process by the end of July, however as this was during the height of the summer holiday season, many key team members were on annual leave which caused a slight delay in the testing process. The testing process was managed by our QA manager Jiri Frank and Jana Hoffmann. Further supporting documentation was prepared which consisted of a revised version of the Catalogue of Requirements (the CoR document captures the specifications of all relevant features for the portal development and also serves as a reference for testing). More importantly, this document also captures all the relevant information from a user's perspective of all the key functionalities of the BHL-Europe portal as well as prioritising the features according to user's feedback and contractual obligations. The document has since been revised by Jana Hoffmann based on further feedback from the Tervuren Workshop held in June and was made public to all consortium members.

In August the testing process finally kicked off. The testing team, which consisted of an internal testing group, began by testing 20 features. This process was handled by the QA Manager and Jana as a separate process i.e. once tested the team sent their comments and feedback on each of the features (e.g. observations, detailed descriptions of any bugs discovered and recommendations to resolve any issues). The QA team compiled all the information from the testers in GitHub which was then assigned to respective developers to fix such bugs that had been reported.

The QA team continued to do the post-testing cross-checks of bugs fixed during the feature freeze of the developing phase up until the end of August.

---

[28] http://www.eu-nomen.eu/portal/webservices.php

[29] http://webservice.catalogueoflife.org/

[30] https://bhl.wikispaces.com/file/view/Evaluation+and+Testing+Workshop.pdf

3.0.3.11 Shared Development Environment

Between July and the end of August, work was carried out to link up the various development components – notably the GitHub codebase was connected to the deployed development system using Jenkins (the selected Continuous Integration tool). This meant that the coding carried out by the developers on their local systems can now hook in so that once code is committed, that code could be deployed to the virtual servers at NHM. During this period the impact of vacations meant that there were delays to getting commitment of all development code into GitHub, limiting the value this integration could deliver.

During this period, only the development server was integrated for deployment in this way, which meant that the developed code does not yet deploy fully to an environment suitable for the testers. A self-contained virtual instance holding the portal and early ingested test data was supplied by AIT in Tervuren, and this virtual machine was brought online to act as the test server.

Because Pre-Ingest at this point was not yet producing OLEF output, this highlighted some limitations on the metadata held in the test data set, limiting our ability to fully test. For example, date metadata for records is not OLEF aligned and thus inconsistent search results were seen during testing, the testing team identified a number of issues which were identified as metadata related.

The other major challenge in this timeframe was to achieve GitHub commitment of all development components, including a base of test data, to enable the deployed development code to be validated. This test data issue had a constraining effect in this period, meaning that it was not possible to validate the functionality of the developed code deployed on the NHM virtual servers, as searches/browsing had no underlying data to work against.

To achieve full commitment of code, resolve the test data issues and determine the best route forward, a Developer's Workshop during the Hague meeting was held.

3.0.3.12 Progress on Development/Production Environment (NHM)

Following the The Hague meeting, NHM reconfigured the development servers, and deployed additional clean virtual machines to make up the current service infrastructure.
We now have split virtual servers between four layers:

- Integration, where developed components are brought together for initial validation.
- Testing, where stable components will be deployed for the testing phases.
- Production, where the final components will be deployed.
- Common – where common components across Integration/Testing/Production reside, including the database server and a web proxy in the NHM DMZ.

The integration server and proxy have both been brought online, and deployment to them via continuous integration is live.

Configuring the automated deployment jobs in "Jenkins" (continuous integration server): all changes to the whole system made by developers need to be built and deployed to the integration server. This server holds the whole BHL-Europe system and always reflects the most recent state of development. Once the system has proven to be mature enough to be tested the BHL-Europe system components are deployed to the next stage, to the testing servers. All in-depth testing can than take place in a stable environment. Once testing has proven the system to be ready for release it is deployed to the production servers and the new features and bug fixes will become available. All these builds and deployment tasks are performed by a special server, the continuous integration server, in which specific jobs need to be configured of each component of the system and for each stage mentioned above.

3.0.3.13 System Architecture

Atos helped to establish and document the system architecture including the distribution of software modules on different servers and environments (development, integration, test, production) within the context of scalability and clustering. Also included was the maintenance and upgrade strategy for different components such as operating system, languages, SDKs, libraries, packages, application servers, and applications.

3.0.3.14 D3.7 Key Components documented for output of D3.5

D3.7 was submitted to the EC in May 2011. The main objective of the document was to cover the technical architecture status in terms of where our technology stands to date with regards to the Implementation Phase. Overall the document described the technical architecture, workflows and interdependencies for the Key Components previously identified in deliverable D3.5 including: - BHL-Europe Portal, Pre-Ingest tool and Archival Storage.

3.0.3.15 D3.8 Sustainability policy for continuation of services

Consultations have taken place with the management/executives of several partner organisations to determine the future level of technical support. At this stage only NHM and NHMW were able to make firm commitments for post-project support. However, MfN and RMCA are in active discussion with their administrations to determine what they can do.

Thus, D3.8 is a mix of commitments and ambitions. It will take some months for the final disposition to become clear, so D3.8 is a statement of the current position. A final version of D3.8 will be issued in April 2012 (or sooner if commitments are agreed sooner).

3.0.4 Status and progress of activities – WP5

3.0.4.1 CWG meetings

During the Annual Review Meeting, Tervuren and Meise (Belgium) a CWG meeting was held. Several employees of the RMCA interested in the project joined this meeting of the communication working group. Jiri Kvacek and Jiri Frank presented their work on the virtual exhibition BLE[31] on spices. All the different features were discussed and suggestions were made. It was agreed that BHL-Europe will continue working on a virtual exhibition on expeditions until the November release. Patricia Mergen presented RMCA's virtual exhibition on expeditions[32], where they combined information from different projects (BHL-Europe, OpenUp!, Europeana). Furthermore the timeline for dissemination activities related to the portal launch were discussed and agreed. For further information please refer to the minutes on the BHL Wiki.[33]

During the Public Core Review Meeting, The Hague (The Netherlands) a CWG meeting was held. Due to delays in portal development we had to postpone prepared dissemination activities. Major dissemination events associated with public release of the BHL-Europe portal were discussed, including soft launch (implementation of virtual exhibition BLE) and hard launch (press releases and social media dissemination). NMP suggested having a pilot project based on dissemination of spices and spices plants from BLE in farmer markets and

---

[31] http://www.biodiversityexhibition.com/

[32] http://cybertaxonomy.africamuseum.be//drupal7/

[33] https://bhl.wikispaces.com/BHLE_AnnualReview2_minutes

Christmas markets in the Czech Republic. Jiří Kvaček and Jiří Frank presented their ongoing work on the virtual exhibition BLE[34] on spices.

3.0.4.2 BHL-Europe Website - *http://www.bhl-europe.eu*

During July and September we repeatedly edited the project web site and also the BHLWiki, which is a platform for internal communication of the project. The number of languages of the BHL-Europe web site was increased to ten; two other languages – Hungarian and Dutch are in preparation. RBINS is currently working on a Dutch translation.

3.0.4.3 Europeana Project Share

During the last meeting of the Europeana Group of Projects (May 2011) it was agreed to build a project share space as well as considering BHL-Europe for the pilot phase of the shared space. Henning Scholz collaborated with Europeana to identify suitable project documents and uploaded them to the web space. Due to some technical reorganisation of the Europeana sites, the project share space is not yet live. When this system is up and running it should facilitate the communication and knowledge exchange between all projects of the Europeana group. This should help, for example, to avoid duplication of work.

3.0.4.4 Life and Literature, Chicago, USA

The Life and Literature[35] conference in Chicago from 14 to 15 November 2011 will convene librarians, biologists, computer scientists, publishers, students, and other stakeholders to set the agenda for biodiversity literature digitising and its networked environment for the next four to five years. BHL-Europe will be represented by a number of representatives and is presenting four posters. The abstracts can be found on the BHL Wiki[36] and are online available on the Life and Literature homepage.

3.0.4.5 Usability and testing

Jiří Frank (NMP) as a testing manager and Jana Hoffmann (MfN) as product owner have coordinated testers and developers in order to achieve final parts of the work on the project portal. Therefore they closely cooperated with the Tech group and the WP3 management. The testing infrastructure was set-up and maintained in GitHub.

---

[34] http://www.biodiversityexhibition.com/

[35] http://www.lifeandliterature.org/

[36] https://bhl.wikispaces.com/Life+and+Literature+Poster+Session
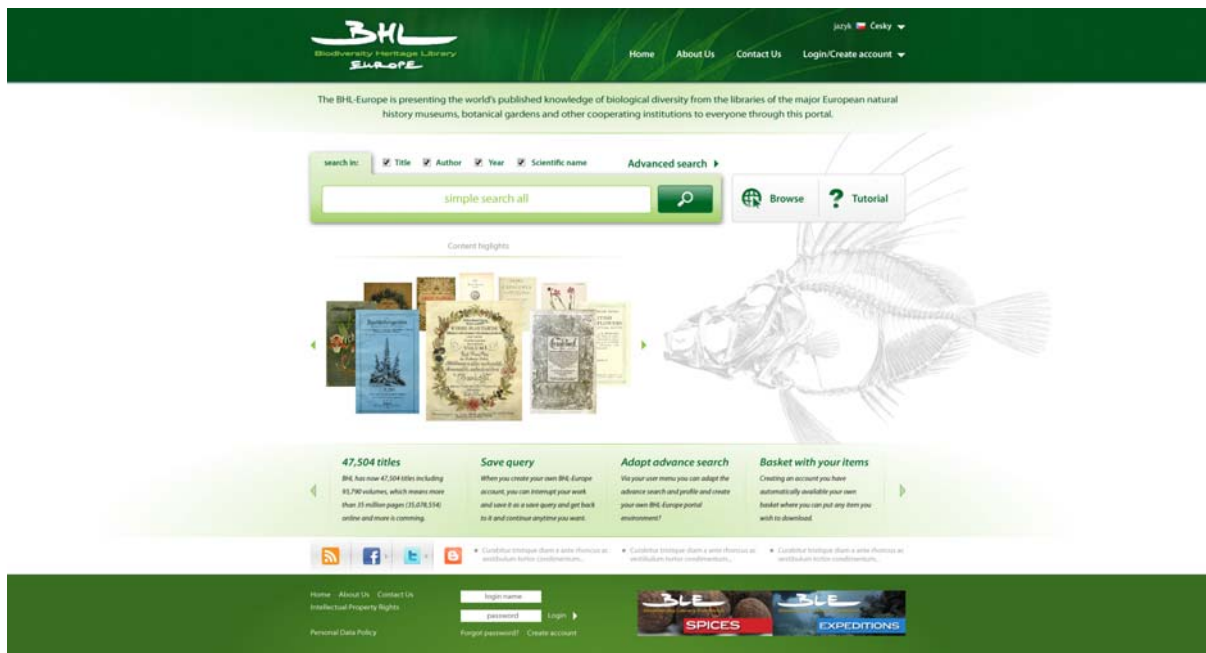
Figure 1: User Interface (UI) for the BHL-Europe portal.

First initial tests on closed issues/ features of the BHL-Europe portal were carried out by the testers and reported in the issue tracker in August. Relevant bugs and general observations on portal behaviour were reported in the issue tracker enabling the test manager (Jiří Frank, NMP) to produce a thorough test report for the Public Core Review Meeting in The Hague.

During the Public Core Review Meeting in The Hague a tester's group workshop was held in order to further plan the portal development, specify general user requirements and add relevant features.

During September-October Jiří Frank (NMP) provided and finalised the drafts for the new User Interface and portal functionalities which were then discussed in the Techgroup and during the Public Core Review Meeting in The Hague. The BHL-Europe portal homepage is given as an example (Figure 1).

3.0.4.6 BLE – Virtual Exhibition

In May and June most NMP staff efforts was focussed on finishing a draft of the virtual exhibition BLE[37]. The purpose of BLE is to promote the content of BHL-Europe and the BHL-Europe portal to a wider public. In cooperation with IT4Care we finished the topic "Spices" which includes 52 books from 19 countries, in 17 languages ranging from 15th to 20th century (Figure 2). They are divided into 10 subjects and can be sorted by 50 spices plants (groups of plants in some cases). There are 52 spices cards underlined with more than 150 pages of interesting information. In total there are 200 illustrations of spices on display. The BLE was successfully presented during the Annual Review Meeting, Tervuren and Meise, Belgium.

During September NMP worked on improving the virtual exhibition BLE. A new exhibition on "Expeditions" will be finalised in November.

BLE and other social media (Twitter, BHL-Europe blog, Facebook) will be used to massively disseminate the BHL-Europe portal after the Core Release in November.

---

[37] http://www.biodiversityexhibition.com/

Figure 2: BLE entry page.

3.0.4.7 Campaign together with Farmářské trhy – Farmer's Markets company

Farmer's markets in Prague are very popular events for Prague citizens. They take place every weekend around Christmas in various places in Prague. The most successful Farmer's Market "Farmářské trhy" in Praha/ Dejvice is becoming more international. It offers products including spices from Asian countries.

The "Spices event" will be based on Christmas spices, which will be offered by farmers and merchants. BLE "Spices" will be presented during this event and will thus disseminate BHL-Europe to the wider public.

In the period of 8-24 December 2011 there will be three spices events. The program includes a presentation of BLE and especially for children "Testing of spices", "Games with spices", "Recognition of spices", and "Spice art design". Various posters presenting BLE with BHL-Europe and BLE logos were prepared.

3.0.4.8 Develop, prepare and update promotion kits and related printed matters

We prepared a new design for the year three of the BHL-Europe project with a fossil ammonite. In September and October we worked on the new BHL-Europe project flyer and calendar for 2012. In October a new poster master template was created including the design theme for the third year of the project (Figure 3).

3.0.4.9 BHL-Europe Newsletter

The annual report was published in issue #8[38] of the BHL-Europe Newsletter (Fig. 4). It was finalised in July 2011. 100 copies were printed and distributed among partners during the Public Review Meeting in The Hague, the Europeana Tech meeting in Vienna and sent to the EC. The work on the graphical design of year three was completed. Newsletter #9 was prepared in October. However, because of internal problems it is not yet finished. Henning Scholz revised and approved this newsletter issue and RBGE did the English editing.
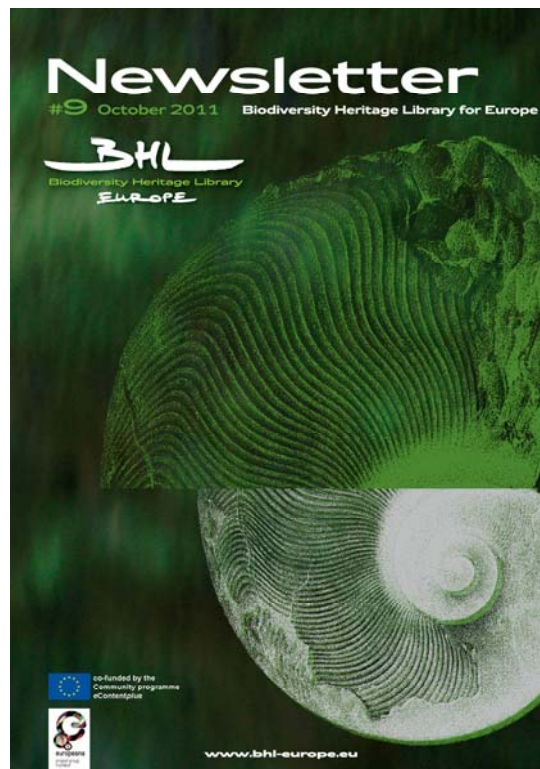
---

[38] http://www.bhl-europe.eu/de/newsletter/neue/bhl-europe-newsletters-8

Figure 3: Design for the year three of BHL-Europe
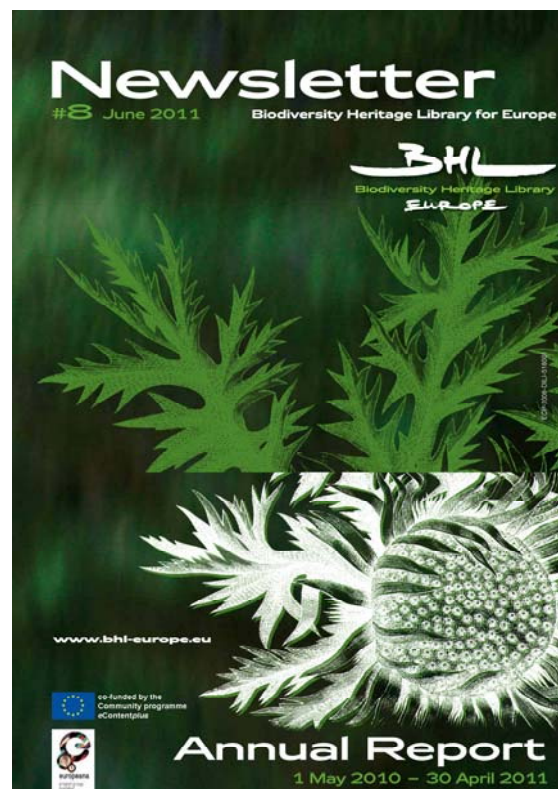

Figure 4: Cover of BHL-Europe Newsletter

3.0.4.10   Facebook, BHL-Europe blog, BHL-Europe Twitter and Netvibes

At the beginning of October 2011 BHL-Europe had 129 'members' on Facebook[39] and the BHL-Europe Facebook page "likes" all available Facebook pages of our partners (14).

Beginning of October BHL-Europe had about 100 followers on Twitter [40] and was following 122 other accounts. Monthly analysis of the BHL-Europe Twitter account was undertaken by Boris Jacob (RMCA) and discussions with staff from BHL-US took place on how to enhance the social media activities.

Various short texts about the BLE "Spices" Exhibition were prepared which will help to disseminate BLE and BHL-Europe in the social media.

3.0.4.11   Project publications in the current reporting period

Partner's publications[41]:

- Anonymous 2011: La numérisation de la biodiversité - Een digitale biodiversiteitsbibliotheek. Le Journal électronique du Jardin botanique national de Belgique - De digitale nieuwsbrief van de Nationale Plantentuin van België, N° 26, 2011; in French[42] ; in Dutch[43]
- Duteille, C. and Détienne, V. 2011: Les collections historiques du Muséum national d'histoire naturelle dans Gallica. Blog post on the Gallica Blog 13 May 2011 http://blog.bnf.fr/gallica/wp-trackback.php?p=2820
- Fabri, R. 2011: BHL-Europe: bibliothèque européenne virtuelle de la biodiversité Machoechel (Journal du personnel du Jardin botanique national de Belgique) Vol. VIII, 2, 2011, p. 4.
- Kvaček, J. and Scholz, H. 2011: Biodiversity Heritage Library for Europe. International Botanical Congress, Melbourne 2011, Abstract book, p. 472
- Scholz, H. 2011: Biodiversity Heritage Library for Europe - Vom Suchen und Finden wissenschaftlicher Literatur im Internet Beiträge zur Paläontologie 32: XV. - conference abstract.
- Scholz, H. 2011: Virtualisation of literature research through the Biodiversity Heritage Library for Europe Beiträge zur Paläontologie 32: 34. - conference abstract
- Scholz, H. 2011: Europe's local biodiversity library. International Innovation, Environment, Issue 1: 71-73; Research Media Ltd, Bristol, UK.
- Valdecasas, A.G., Cayuela, L. and Becerra J., M. 2011: Biodiversidad and Cibertaxonomía that includes information on BHLE In: Biodiversidad. El Mosaico de la Vida. Madrid, 2011, www.fecyt.es

3.0.5 BHL – Extract from the Program Director's Report

On behalf of Tom Garnett, Director of the BHL, Mr. Graham Higley, head of library and information services at the Natural History Museum in London, accepted the 2010 John Thackray medal[44], which was awarded to the BHL in recognition of the past three years of work building its digital library.

---

[39] http://www.facebook.com/pages/BHL-Europe/151086001600041

[40] http://twitter.com/BHLEurope

[41] https://bhl.wikispaces.com/BHL-Europe-publications

[42] http://www.br.fgov.be/PUBLIC/GENERAL/VISITORINFO/VISITORINFOFR/MUSA/musa_2011_2.html

[43] http://www.br.fgov.be/PUBLIC/GENERAL/VISITORINFO/VISITORINFONL/MUSA/musa_2011_2.html

[44] http://www.shnh.org.uk/awards/the-john-thackray-medal/2010.html

Chris Freeland and William Ulate attended the Society for the Preservation of Natural History Collections annual conference[45] in San Francisco, where BHL was offered a table in the vendor hall. Chris Freeland also presented a poster on the Engelmann's project and attended the stand with William.

Bianca Crowley and William Ulate attended the BHL-Europe Annual Review and shared with BHL-Europe colleagues on their common experiences, requirements and lessons learned. William Ulate gave a talk during the BHL-Europe Annual Review to update participants on the advances of the Global BHL Project.

Chris Freeland attended the Technical Workshop of the Digital Public Library of America[46] at the Associate Librarian for Library Services Office at the Library of Congress, which convened a small working group meeting to begin to make recommendations for the overall technical architecture of a Digital Public Library of America (DPLA) and to converge on a set of key technical principles upon which the DPLA should be built. Particularly, the objective was "*to surface and identify concrete insights, including areas for future research; development and policy proposals; and other tangible outcomes*". Chris Freeland led an initial discussion on the overall goals and architecture of the DPLA. His presentation "DPLA Technologies: Foundations for Growth & Sustainability"[47] was used as the base for discussion during the first day of the meeting. Chris Freeland worked together with participants at the end of the DPLA meeting to summarise the key takeaways from the workshop: http://blog.chrisfreeland.com/2011/06/dpla-summary-from-jun-mtg.html.

Charles Darwin's Library[48], a BHL special collection that draws on original copies and surrogates from other libraries and provides full transcriptions of his annotations and marks, was released online. This collection is composed by a digital edition and virtual reconstruction of 330 of the most heavily-annotated books owned by Charles Darwin's at the time of his death.

Chris Freeland attended the International Botanical Congress 2011[49] in Melbourne, Australia. During the event, which concentrated more than 2,100 participants, several botanists praised BHL during their talks. During the International Botanical Congress 2011 in Melbourne, Australia, Chris Freeland helped to organise the Symposium "Informatics tools for the semantic enhancement of taxonomic literature"[50] with BHL-Australia team lead for Atlas of Living Australia, Dr. Elycia Wallis, from the Victoria Museum, who chaired the Symposium. The 150 to 200 attendees listened to renowned speakers on the area of tool development:

- Biodiversity Heritage Library: a global resource for open access scientific literature (Chris Freeland).
- Link yourself or perish? Presentation of *PhytoKeys*, the next generation journal in systematic botany (Lyubomir Penev).
- *Taxonomic Literature (TL-2)* as a major index to help guide selection and prioritization of botanical literature digitisation (Laurence Dorr).
- A schema for description and exchange of taxonomic publications' content (Donat Agosti).

---

[45] http://www.spnhc.org/news/view/26

[46] http://bit.ly/oIvb7J

[47] http://www.slideshare.net/chrisfreeland/dpla-technologies-foundations-for-growth-sustainability

[48] http://www.biodiversitylibrary.org/collection/darwinlibrary

[49] http://www.ibc2011.com/

[50] http://www.ibc2011.com/downloads/IBC2011_Program_Book.pdf

- Names-based retrieval of library and internet resources in the Virtual Library of Biology (Vifabio): issues in using taxonomic names in library catalogues and metasearch tools (Gerwin Kasperek).
- Fine-grained semantic markup of descriptive data (Hong Cui).

Furthermore, Chris Freeland met with staff from the Museum Victoria, Australia, to coordinate on issues related to the BHL Australia project and its next steps and interactions with others BHL programs at a global level.

William Ulate has tweeted on major Global BHL related events and news from most BHL Programs translated into several languages thanks to complimentary collaboration from other colleagues.

Chris Freeland and Jay Page presented at the first Planning Meeting in Hawaii of the recently awarded NSF-funded proposal "*Advances in Bioinformatics: Innovation: Collaborative Research: The 'Global Names Architecture,' an infrastructure for unifying taxonomic databases and services for managers of biological information.*" The meeting was also an opportunity to gather requirements to allow the Biodiversity Heritage Library's CiteBank, coupled with a scientific name 'usage bank', to create a document submission module to enable users to share key reference material in compliance with the "safe harbor" principles of the Digital Millenium Copyright Act (DMCA).

Chris Freeland attended the BHL-Europe Public Core Review Meeting at the National Library of the Netherlands (KB), The Hague, The Netherlands. The progress of the BHL-Europe program was reviewed and discuss along with the latest BHL-Europe Web Portal design and new developments.

Trish Rose-Sandler attended 2011 LITA National Forum[51]: Rivers of Data, Currents of Change in St. Louis, Missouri, USA, where she presented "Building the New Open Linked Library" along with colleagues from the Smithsonian Institution.

3.0.6 Risk and issues

We have identified a number of issues and delays in the project that have implications for the success of the project. Important issues are explained in some detail in the following.

The timely production of high quality status report by all partners is still an issue for the project. The same is true for the monthly activity reports and the financial statements. Experience has shown that it is very difficult to get the full set of documents in time. Status reports need to be narrative on task level to facilitate the compilation of the full report. Activity reports need to be on task level so we are able to match the hours with the narrative report. Eventually, the financial statements need to reflect all work so the narrative progress report is consistent with the financial report to justify all costs. If we are not able to produce all reports in high quality more quickly next year, we risk not being finished by the end of the project and we will then lack the resources to finalise the paperwork. This will at least delay the payment of the final rate to the project.

Currently there is an issue with the GRIB. In spring the technical infrastructure was changed for the GRIB which caused serious technical issues at the hosting site affecting the availability and finalisation of the GRIB prototype. The technical issues were finally solved beginning of October, however this caused delays in the work with the GRIB and all library catalogues need to be re-imported to the database and deduplicated. This circumstance gives

---

[51] http://www.ala.org/ala/mgrps/divs/lita/conferences/forum/2011/concurrent.cfm

us a shorter time of period for smoothly switching from the Scanlist to the GRIB and therefore there is still a risk of not being able to replace the Scanlist with the GRIB in time.

By the time of writing, the Pre-Ingest tool is not useable by content providers. The deadline for finalising the tool had been repeatedly postponed since end of last year. The responsible developers were asked in good time to provide a Pre-Ingest tool tutorial during the annual review meeting in June in order to give our content providers the possibility to learn how to use the tool and ask questions. At that time the Pre-Ingest tool was not accessible and thus no workshop on the tool was given. Finally the tool was made available for testing end of August 2011. However the test conducted with some of our content providers showed that the tool is not usable as is (see section 'Management of the Pre-Ingest tool testing' in WP2). However, new resources were found after the test to finalise the tool. Furthermore the newly-created position of Pre-ingest Manager is a good way of assisting our content providers in the future with ingesting their content in case of any problems.

The issue with the Pre-Ingest tool has a series of implications for several areas in the project. First, we cannot run all available data through the process and produce the various files we need for the various OAIS components (SIP, AIP, and DIP). This delays the testing of the entire ingest process and the test of the technical infrastructure. It also delays the test of the data performance in the portal, for example. This again influences the entire data upload to NHM servers. Content providers need to understand how the data is displayed before uploading the full corpus. Quality checks need to be carried out at an early stage of the process to avoid duplication of work. Eventually we risk not being able to ingest all content we want and need into our system by the end of the project. We will run parallel instances of the ingest process once finalised to mitigate that risk.

Second, our content providers need time, training and clear instructions to fully understand the process to work with the tool independently. We need to understand the work required by all content providers to make sure sufficient resources (and expertise) are available on the content provider side. This is particularly important for the non-consortium content providers, as they don't have any budget provided by us to do the work. Currently we cannot expect our content providers to do the work with the Pre-Ingest tool. We have identified an ingest manager, who is working in close collaboration with the development team to ensure all data will be processed efficiently and quickly.

Third, a delayed Pre-Ingest tool delays the content acquisition process. New and prospective content providers need detailed information on what we expect them to do and how to prepare the content best for us. This requires a successful test of the Pre-Ingest process to identify problems and errors that may require a revision of the current file submission guidelines. Additionally, a public portal with content properly ingested into the system is also required to actively promote our work to prospective content providers and thus facilitate the content acquisition process. Otherwise, we are losing time to negotiate with new content providers.

Fourth, the Pre-Ingest is an important part of the Best Practice Guide (D2.9). Without being successfully tested and used by our content providers, this section in the Best Practice Guide cannot be finalised. A delayed D2.9 will also delay the work on deliverable D5.11, which is based on D2.9. The three WP leaders involved here are in regular communication to mitigate the risk.

It is also worth mentioning that more and more content providers ask for visibility of their already-uploaded content and would like first to see their content online before uploading further content. This is an understandable requirement from our stakeholders after two years

project lifetime. There is a high risk of losing the interest of our content providers in the project if we cannot show them the online availability of their content.

3.1 Resources employed

Resources employed for the reporting period (person-months)[52]						
Beneficiary short name	**WP 01**	**WP 02**	**WP 03**	**WP 04**	**WP 05**	**TOTAL**
MfN	7.41	6.69	1.64	0	0.9	16.64
NHM	0	0	3.93	0	0	3.93
NMP	0	0.14	0	0	9	9.14
EDLF	0.21	0.23	3.76	0	0.14	4.34
AIT	0.02	0	11.85	0	0	11.87
ATOS	0	0	7.79	0	0	7.79
FUB-BGBM	0	0	4.19	0	1.01	5.20
UGOE	0.21	0.38	0	0	1.95	2.54
NHMW	0.2	0	3.31	0	0	3.51
LANDOE	0	3.3	2.5	0	0	5.8
HNHM	3.63	0	0	0	0	3.63
MIZPAS	0.01	0.2	0	0	0.12	0.33
UCPH	0	0	0	0	0	0
NAT	0.08	2.06	0	0	0	2.14
NBGB	0	2.09	0	0	1	3.09
RMCA	1.25	8.5	1	0	0.5	11.25
RBINS	0	1.03	0	0	1.54	2.58
BnF	0.05	1.07	0.11	0	0	1.23
MNHN	0	2.84	0.43	0	0.12	3.39
CSIC	0	1.06	0.05	0	0.18	1.29
MSN	0	0.22	0	0	0	0.22
RBGE	0.03	0.25	0.3	0	0.06	0.63
Sp2000	0	0	0.05	0	0	0.05
Wiley	0	0	0	0	0	0
SIL	0	0	0	0	0	0
MOBOT	0	0	0	0	0	0
UH-Viikki	0	0.26	0	0	0.12	0.39
UBER	0	2.08	3.55	0	0.02	5.64
TOTAL	**13.1**	**32.40**	**44.46**	**0**	**16.66**	**106.62**

---

[52] The calculation of PM is based on productive hours mentioned in CPD form of partners.

3.2 Work Package Overview

Work package description

Work package number :	1	Start date:	M1	End date:	M36
Work package title:	Project Coordination and management				

Objectives for the period

Administrative objective: Ensure adherence of the consortium to the rules, regulations, and financial guidelines of the *eContentplus* programme; establish the project in the European biodiversity community.

Technical objective: Implement the project as set out in the work plan; ensure exchange of information and communication between partners; ensure progress of the project; guarantee timely deliverables.

QA objective: Ensure verifiable progress of the project and be committed to high quality output that has tangible impact on *eContentplus* programme objectives.

Business development: Develop a business plan for long term sustainability of BHL-Europe.

Description of work carried out and achievements

Task 1.1 – Administrative coordination

1.1.1 – Resource planning (financial, personnel, material), monitoring and controlling.

Achievements: The BHL-Europe project office is working on evaluating the budget of all partners in order to spend the budget efficiently until the end of the project. This goes in hand with redistribution of tasks if necessary in order to assure that all tasks in the DoW will be completed by the end of the project. For the technical development of the project (WP3, Task 2.1), the issue tracker software GitHub[53] is used.

1.1.2.1 – Liaison between the Commission, consortium members, EUROPEANA, BHL, and external experts; effective communication with the consortium members, Work Package leaders, the Commission, and interested external parties; coordination of meetings and progress reviews.

Achievements: Email, phone, Skype, BHLWiki, and regular personal meetings ensure effective communication between all parties involved in BHL-Europe and related to the work carried out in the project. Every Tuesday, the Project Management Group of BHL-Europe has a conference call to update on progress and to discuss recent developments and issues. An Open PMG call between the project coordination/ work package leaders and the consortium was held in September. The IT team has a conference call every other Monday to review work progress and update the GitHub issues. Furthermore, mailing lists are used to facilitate the communication between content providers.

Partners report their work on a monthly basis and provide a narrative report every three months. Internal reports are produced quarterly to summarise the project progress for all consortium members.

---

[53] https://github.com/bhle/bhle

PCO of BHL-Europe is communicating with Europeana on a very regular basis on various levels (e.g. project management, content ingestion, CCPA). In May 2011 the PCO attended the Second Europeana Group of Projects meeting in The Hague where the set up of a project share space for BHL-Europe was initiated. Furthermore, the Europeana Search API and Search Widget API were circulated in the BHL-Europe consortium.

1.1.2.2 – Coordination of meetings and progress reviews.

Achievements: Two contractual meetings were organised in the reporting period, the 2nd BHL-Europe Annual Review Meeting in Tervuren/Meise, Belgium in June 2011 and the Public Core Review Meeting in The Hague, The Netherlands in August/ September 2011.

1.1.3 – Production and consolidation of periodic external reports, including cost-statements; internal quality assurance; set-up of the Web-based project management portal.

Achievements: The resource tables in the progress report D1.4 and D1.6 were updated according to the numbers provided with the 2nd financial statement. The financial statements of partners were cross-checked, finalised and sent to the EC in August 2011.

1.1.4 – Representing the project.

Achievements: The project was presented at various conferences and meetings (see below). The Project Coordinator represents BHL-Europe and the European museums in the Council of Content Providers and Aggregators of Europeana.

Task 1.2 – Technical coordination

1.2.1 – Work-package and task coordination.

Achievements: This is an ongoing task using the tools and services implemented before (Project Server and BHLWiki).

In month 24 WP4 was officially closed. An additional meeting was held between the PCO and the NHM team to assure that WP4 tasks were fulfilled according to the DoW and results documented properly. The former WP4 leader will still be available for upcoming questions until the end of the project. Furthermore the handover of WP3 lead to the new WP3 leader team and redistribution of responsibilities was supervised.

In the last weeks and months, some reorganisation of important areas in WP3 was necessary. This required the analysis of budget and the revision of work plans for partners. A revision of tasks per partner based on the catalogue of requirements was also necessary. Some PMG calls and PMG meetings in the reporting period were dedicated to these aspects. This topic was also discussed during the Tech Calls and Tech Meetings. As the project office at MfN has the full and detailed documentation of budget and workload per partner the PCO was involved in the entire process to support the management and make the necessary arrangements with the EC.

1.2.2 – Project plan maintenance; monitoring of project progress and milestones; identification and trouble shooting of technical and organisational problems.

Achievements: This is an ongoing task using the tools and services implemented before (Project Server, MS Project, GitHub).

1.2.3 – Timely production of deliverables.

Achievements: The Project Office of BHL-Europe ensures timely production of deliverables. In cooperation with the WP leaders, the members of the consortium were asked to

fulfil their commitments according to the Description of Work.

1.2.4 – Quality control against the technical and contractual aspects; coordination with EDL Foundation Office.

Achievements: We have a strict internal reporting procedure in place. Partners have to report all their activities on a monthly basis and provide a narrative report of their activities on a quarterly basis. This way we ensure partners work as planned and we are able to identify potential problems as early as possible.

Task 1.3 – Quality assurance: Definition and communication of quality assurance procedures via project management portal, coordination of quality assurance process (internal as well as external review procedures for the various project results and deliverables).

Achievements: We actively maintain a risk and issue register in Google Docs to manage the risks and issues identified by the consortium and the PMG. The risk and issue log for the project is maintained by the project office. Appropriate counter measurements are discussed and decided during the PMG calls.

External reviewers are identified for most deliverables and will be consulted to review the deliverables of the project.

Task 1.4 – Business plan

1.4.1 – Investigate services that help the sustainability of project results

Achievements: The VZG helps BHL-Europe in sustaining the GRIB. The use of open source technologies that are supported by a large community (e.g. Drupal) helps to maintain and further develop software components. The close collaboration with Europeana and Global BHL is particularly important to ensure the long term access to the aggregated content. A sustainable Europeana on the other hand is important, to get more funding for scanning initiatives, which helps BHL-Europe to grow organically.

1.4.2 – Identify related networks or organisations interested in the implementation of project results

Achievements: BHL-Europe is in discussion with two new EU projects of the biodiversity domain to investigate areas of collaboration or options for implementing project results, i.e. Natural Europe and OpenUp! One new proposal of several European natural history institutions will be submitted soon in the FP7 programme where BHL-Europe is partner through MfN. The closer alignment of the Europeana Group of Projects since autumn 2010 will potentially result in new opportunities for the implementation of project results. In this context we investigate the upcoming FP7 and ICT-PSP calls to realise the current ideas for a future development of BHL-Europe in collaboration with Europeana.

1.4.3 – Develop a business plan for long term sustainability with WP2 and WP3

Achievements: BHL-Europe has drafted the business plan for long-term sustainability (D1.9).

Task 1.5 – Networking and clustering activities

1.5.1 – Coordinate with relevant partner networks; Identification of new important networks as potential disseminators of project results; Identification and analysis of stakeholders

Achievements: The most important partner networks for BHL-Europe are Europeana (data ingest), BHL (alignment of standards and technologies, content ingestion procedures), Global BHL, and OpenUp!. In addition, first contact was made and possible

collaborations discussed with the new project Natural Europe.

1.5.2 – Contribute to the clustering activities of the EC.

<u>Achievements:</u> No cluster meetings of the EC were held.

Deviation from work plan & remedial action
No deviations from work plan.

Work package description

Work package number :	2	Start date:	M1	End date:	M36
Work package title:		Analysis of domain content and management of the content acquisition process

Objectives for the period

IT objectives: Establish bibliographic database system, metadata repositories and Web-based content management system.

Management objectives: Ensure that all relevant biodiversity literature is listed to be scanned following a priority list; ensure that all content providers agree on the technical architecture of the project; ensure effective scanning in all content providing institutions; ensure linkage to Europeana; ensure extension of the content providing network.

Description of work carried out and achievements

Task 2.1- IT Development

2.1.3 – Develop a Web database to support analysis of domain content and management of the scanning process based on 2.1.1 and 2.1.2

Achievements: It is also worth mentioning that more and more content providers ask for visibility of their already-uploaded content and would like first to see their content online before uploading further content. This is an understandable requirement from our stakeholders after two years project lifetime. There is a high risk of losing the interest of our content providers in the project if we cannot show them the online availability of their content.

Tasks 2.2 – Analysis of domain content

2.2.2 – Use of the Web-database to identify the distribution of this relevant literature in the libraries of the content providers.

Achievements: Several partners are running individual processes to establish a list of relevant literature. The full details are given in the status section above. Currently, the BHL ScanList is still actively used to manage scanning activities and support the content analysis. After the GRIB is fully up and running it is planned to merge the BHL ScanList into the GRIB.

2.2.3 – Identify responsibilities for content contribution (which institution should provide the identified content after consideration of technical qualification, data standards, IPR).

Achievements: BHL-Europe has prepared a collection development policy to find a consensus on the strategy for the extension of the corpus of digital biodiversity literature to be considered by BHL-Europe. This collection policy was finished and is now integrated into the Best Practice, but may also be used as a separated document for content provider acquisition.

Task 2.3 – Management of the content acquisition process

2.3.1 – Identify content holder requirements

Achievements: Although BHL-Europe finished the work on the MoU in the 2nd reporting period, we now have revised the MoU to further prepare for the extension of the content provider network. The finalisation of this revised MoU was dependent on the

finalisation of the Europeana Data Exchange Agreement. An Appendix will be added to the MoU for taking the Europeana Data Exchange Agreement into account.

2.3.2 – Assist partners in implementation and evaluation of scanning operations; control duplicate scanning of literature (using results of task 2.2); discussion and distribution of data standards and specifications; work with individual donors and governments to facilitate the funding of the scanning; attracting new content providers.

Achievements: The preparation for the final version D2.9 'Delivery of the final and revised best practice guidelines and standards' started during this reporting period. Missing information was identified and contributors were assigned to the respective chapters.
The final version is in progress - the Pre-Ingest section is still missing as it is dependent upoon the development and final release of the tool, the GRIB chapter was delayed due to the technical issues at the VZG, however the GRIB section is currently in the process of finalisation.

Management of the Pre-Ingest tool testing: A first version of the Pre-ingest tool was provided end of August 2011. The tool was tested in collaboration with our Content Providers within WP2 and feedback was reported to WP3 using the BHL-Europe repository software GitHub.

2.3.3 – Coordinate with Europeana, BHL and national scanning projects to ensure that material scanned by BHL-Europe is available through these portals.
Achievements: An updated ingest plan was sent to Europeana including an additional estimation per item to facilitate the planning of Europeana's ingest team.

2.3.4 – Addressing IPR issues in cooperation with WP4.

Achievements: Various content providers are in negotiation with Rights Holders to make in copyright content available for BHL-Europe, see status update above.

2.3.5 – Take into account multicultural and multilingual aspects.

Achievements: The content in progress to be ingested into the BHL-Europe reflects the multicultural and multilingual nature of European natural history libraries (see the recent additions to Europeana for example).

Deviation from work plan & remedial action

(1) There is a delay of the GRIB caused by technical difficulties at the VZG. Communications took place to ensure that the GRIB is up and running as soon as possible and that all available catalogues are imported and deduplicated.

(2) For dependencies with WP3 affecting WP2 work see the description below and the risk and issues section (3.0.6).

Work package description

Work package number :	3	Start date:	M1	End date:	M36
Work package title:	Technological implementation				

Objectives for the period

Management and coordination of technological development and associated standards to allow for a pan-European, distributed and multilingual BHL-Europe.

Technical Objective: The technological implementation will concentrate on the innovative application of proven technologies to deliver stable and sustainable solutions. For this period this consisted of the following:

Continual work on use case requirements for the final BHL-Europe system; Adaptations and Service Continuations of all system components (Feature Improvements, Performance Testing and Integration); processing of content and Pre-Ingest Testing and continuous preparation for the Portal Release.

Description of work carried out and achievements

Task 3.1 Technological Implementation (Overall coordination)

3.1.1 – Management of the technical development team.

Achievements: In this reporting period the technical development team has mainly focused on further development and the implementation phase of the technical and functional architecture of the final BHL-Europe System. The tech group has continued to hold intensive technical meetings through various forums (conference calls, Skype and TMB meetings) to discuss ongoing development work within WP3/WP2. The development team has continued to work intensively on the BHL-Europe System architecture and on the implementation plan for BHL-Europe system. In preparation for the 2nd Annual Review Meeting in Tervuren and the Portal Review Workshop in The Hague the development team continued with the Portal development work and implementation of the BHL-Europe Core Features.

3.1.4 Liaise with scanning centres of the national initiatives for post-processing of content.

Achievements: Ongoing task which includes: content providers uploading metadata and content to the NHM infrastructure for ingest into the BHL-Europe system; the creation of FTP accounts for content providers; QA processing to verify the result of the mappings in collaboration with Content Providers as well as providing technical support.

3.1.5 – Development and adaption of specific tools; Implementation and adaption of TaxonFinder and name recognition tools; improvement and implementation of OCR techniques.

Achievements: Development of the Pre-Ingest tool - The deployment of the development version of the Pre-Ingest tool on NHM's servers was successfully completed. The currently-deployed version is not yet configured and ready for Content Providers to access directly, however the deployment of the shared codebase via Jenkins will allow

future developments to be automatically deployed and visible to the users, which will enable access for review and testing as soon as that can be supported.

WP3 recognises the impact that the availability of the Pre-Ingest tool for WP2 and Content Providers has upon the dependent WP2 processes, and is working to resolve this issue as a priority. Content Provider logon access to the tool was requested and this is due to be implemented shortly.

Schema Mapping Tool - Wolfgang Koller (NHMW) ran a workshop session during the Annual Review Meeting and presented a live demonstration of the schema mapping tool. Further improvements and some adaptations of the schema mapping tool have also been implemented in order to improve the handling for the content providers. Metadata provided by content providers were mapped to the OLEF schema in order to be able to ingest them into the BHL-Europe-Portal. Over the next few months NHMW will map more content providers to OLEF as well as assist with the portal development.

Tesseract 3.00 is being used to implement the OCR micro service. So each page will be OCR'ed during Pre-Ingest. A micro service was created to include TaxonFinder (Perl Server) and TaxonFinder Webservices (Ruby). Each page's taxa are now recognised once the page's OCR'ed text is sent through the TaxonFinder micro service.

OCR - A testset of page images was prepared to analyse the potential OCR improvement using IMPACT tools. In a next step, the design of the ground truth process in collaboration with IMPACT was finalised.

Task 3.2 – Technical integration with EUROPEANA, BHL and national platforms

3.2.1 – Ensure that the standards for data management and image formats are consistent with international requirements and specifically the EUROPEANA and BHL, adoption of EUROPEANA Semantic Elements (ESE).

Achievements: The ingestion process for metadata into Europeana was described taking into account the newly released version 3.4 of the Europeana Semantic Elements (May 2011). Once the content of the partners is in the OAI repository the partners receive a link to their content in the Europeana content checker tool where they can preview their records. Partners gave feedback on their content presentation in the content checker tool. Their feedback was incorporated before the data was finally provided to Europeana. Europeana office also received the link to the new content in the content checker and provided feedback. The May 2011 Europeana publication included the following BHL-Europe sets: BHLUS including NHM (87,572), CSIC (9), LANDOE (3,511), NBN (3,516), RBGE (183), RMCA (52), UBBI (1,634), UBER (60), UCPH (28), and UHVIIKKI (30). This set was updated further in the reporting period. This update produced a new total of 100,561 BHL records in Europeana.

The BHL-Europe Europeana ingest planning was revised and a new planning till project end was forwarded to Europeana office.

The portal code was submitted to the BHL GitHub repository. The modules provided by AIT implement a submit handler for metadata records, the connection to the ApacheSolr indexing system and corresponding portal functionality.

3.2.2 – Develop a distributed data model that will allow countries to retain control of their data, while enabling the material to be available through Europeana, the BHL Portal (with Taxonomic Intelligence tools), and any national portals.

Achievements: Translation modules for the portal were added and a workflow for the translation process was designed. The style and the pagination of the result list were adapted. Additionally, there are now configuration options for the view of the result list and of a record. Search queries can now be saved and managed in a user interface. Nodes of the BHL-Europe portal were installed for test purposes using a virtual machine. This virtual machine was handed over to NHM and was mounted at the NHM facilities in London as a test engine.

Task 3.3 – Addressing distributed access and storage – long-term sustainability

Develop a distributed access and storage system to enable national and international storage of the scanned materials; to develop the storage system in such a way that long-term sustainability of the data is secured.

Achievements: US Data Transfer to BHL-Europe Servers - Over the past couple of months Chris Sleep (NHM) liaised with our US partners with regard to the upload of content. Once the decision was made to ship the data from the US on a set of discs holding the remainder of the initial content, this significantly sped up the process of data transfer and allowed the progress to the next phase of data synchronisation over the WAN (Wide Area Network). The process of copying data from the last batch is now complete and spot checks were run before the drives were shipped onward to Bibliotheca Alexandrina in Egypt.

GUID Infrastructure Setup - Andreas Kohlbecker (BGBM) and Dennis Zielke (UBER) coordinated the development and setup of the GUID-Minting and Handle.net resolver infrastructure. This involved the installation and set up of the Handle.net server on the BHL-Europe server i.e. adapting the Apache Webserver configuration; adapting the Webservice for Noidservice into BHL-mandible server and setting up the user interface for the Noidservice as well as the registration of the handle prefix. This Noid service is used to create generic "nice opaque identifier".

Task 3.4 – Enabling BHL Portal access in European languages – interfaces, usability and mobility

3.4.1 – Work with the Europeana to create multiple language access to the BHL Portal.

Achievements: Port functionality from Prototype to Access - Data harmonisation of RBGE, RMCA, CSIC, and UH-Viikki was continuously improved and been transformed into ESE as part of the negotiation of the Europeana harvesting procedure. The BHL-Europe portal prototype basic search functionalities were translated into seven European languages.

3.4.3 – Apply the model to the key European languages (English, French, German, Italian, Portuguese, Spanish, Danish, Dutch, Czech, Hungarian, Polish, and Swedish), enabling access to the BHL Portal throughout Europe.

Achievements: One of the outcomes of the Portal Development workshop held in Tervuren was for the Catalogue of Requirements to be further revised into a single specifications document capturing all the relevant features for the portal development and also serves as a reference for testing. This document (Catalogue of User Requirements BHL-Europe Portal Feature List & Priorities) also captures all the relevant information from a user's perspective of the key functionalities of the BHL-Europe portal as well as prioritising the features according to user's feedback and contractual obligations. The document has since been revised by Jana Hoffmann and was made public to all the consortium members.

Furthermore, the list of features were migrated to the issue tracker and prioritised. The developers worked continuously through all portal issues (features, enhancement, bugs etc.) to ensure that core features are tested and fully functional ahead of the Public Core Release scheduled for November 2011.

Deviation from work plan & remedial action

1) Decision to develop the Pre-Ingest tool further in the light of user feedback.

Now that Content Providers have had the chance to use the Pre-Ingest tool, it is clear that some modifications are needed to provide a 'best practice' tool for users to upload their data. We have decided to make these changes before importing significant quantities of data to ensure consistency of metadata throughout the corpus.

2) The Public Core Release (M28), which was previously planned for August 2011, has now been rescheduled for the end of November 2011. Late delivery of the Portal system delayed the integration of the whole IT system onto NHM servers, and we decided to ensure that all components had been tested in the integrated environment before going live. The main focus for the development team for the November Release is to deploy a fully working development version of the BHL-Europe system in a single technical environment.

Work package description

Work package number :	5	Start date:	M1	End date:	M36
Work package title:	Dissemination, Exploitation and Evaluation				

Objectives for the period

Dissemination objective: Raise awareness, understanding and action of the project among the community and stakeholders in EU member states; ensure effective dissemination of project goals and results to the target users; ensure good communication within the European scientific community.

Presentation and demonstration objective: Ensure dissemination of project results at conferences, public events and among the networks of the consortium members.

Evaluation objective: Monitor the level of use of BHL; survey the users of the BHL Web Portal.

Description of work carried out and achievements

Task 5.2 – Develop and establish Web environment

5.2.1 – Plan, design, publish, maintain, and update BHL-Europe Web site for internal and external communication with publication of all relevant results and links to partner networks.

Achievements: The BHL-Europe Web site is updated regularly. Texts are now translated into Czech, Finnish, French, German, Italian, Polish, Russian, Slovak, and Spanish. Translation into Dutch and Hungarian is now in progress.

Task 5.3 – Develop and prepare promotional materials and newsletters, press relations

5.3.1 – Develop, prepare, and update target group specific promotion kits (fact sheets, flyers, posters, presentations) and related printed matters (e.g. illustrated book on best practice guidelines and standards for the public).

Achievements: For conferences a dissemination kit (folder, annual report (Newsletter) and flyer) and number of posters were prepared during the reporting period. In cooperation with Europeana, we prepared our virtual exhibition BLE (Spices and Expeditions).

5.3.2 – Prepare a quarterly newsletter to be published online.

Achievements: Newsletter No. 9 was finalised in October, but publication was delayed due to final editing issues. Newsletter No. 10 is in preparation and its release will be synchronised with the portal release in November.

5.3.3 – Ensure proper press relations; maintain the press review; distribution of articles for various types of print and online media (journals, newspapers, blogs).

Achievements: We have prepared text for BLE and for social media. As well as Facebook we also have Twitter and Weblog accounts. All of these channels are prepared for the portal release.

Task 5.4 – Demonstration and awareness raising

5.4.2 – Maintain and update the database of conferences and public events relevant for BHL-

> Europe; identify consortium members to present the project at selected events.
>
> <u>Achievements:</u> We maintain our database of upcoming conferences and have identified project partners to attend these.
>
> 5.4.3 – Alignment of public awareness activities of related projects (e.g. EDIT).
>
> <u>Achievements:</u> We aligned our public awareness activities with related projects. The Europeana Group organised the Europeana Tech conference where we presented BHL-Europe in October. We cooperate closely with OpenUp! on dissemination – particularly on the virtual exhibition BLE.
>
> 5.4.4 – Organise public events and talks.
>
> <u>Achievements:</u> We prepared a pilot project addressing public awareness during Christmas farmer markets in Prague. We will use the topic "Spices", which will be based on our virtual exhibition BLE-Spices.
>
> 5.4.5 – Represent BHL-Europe at conferences and special events (responsibilities depend on conference location and theme).
>
> <u>Achievements:</u> We presented BHL-Europe in numerous conferences to the scientific public (see table of attended conferences below) using presentations, posters and flyers.

Deviation from work plan & remedial action

The work plan of the WP5 was modified due to the slow progress in the finalisation of the BHL-Europe portal. The planned Public Core Release Meeting where the portal was planned to be presented was postponed to the end of November and all other activities associated with dissemination and awareness of the project, particularly the BHL-Europe portal, were postponed accordingly (see 3.1.1 CWG meetings).

The slow development progress also affects the finalisation of D5.9 - Second user evaluation report. A user survey is required to prepare the document, which only can be launched if the portal is fully functional. The document is postponed now, but will be available for M36.

The finalisation of the Newsletter issue 9 was delayed. The publication of texts to our social media streams needs to be also improved. RBINS is in the process of hiring a new communications officer to reduce the backlog of texts and information to be published in the various social media and also in the news section of the project web site.

3.3 Deliverables Status

Deliverables List			
Deliverable No[1]	**Deliverable title**	**Delivery due date[2]**	**Actual date of delivery**
D1.6	Progress report 4 including pre-financing request –**final version including resource table and financial statements**	M24	M28, August 29
D2.7	Content analysis and management status report 2 (metadata, page numbers, content providers) – **revised version**	M24	M28, August 29
D1.8	Progress report 5	M30	M31, December 16
D3.8	Sustainability policy for continuation of services e.g. hosting, future development, helpdesk provision for service/ content provider etc.	M30	M31, December 16
D5.9	Second user evaluation report (postponed to M36)	M30	Postponed

---

[1] Deliverable numbers in order of delivery dates: D1 – Dn. Deliverable numbers must indicate which Work Package they relate to, e.g. D2.1 for the first deliverable from Work Package 2).

[2] Month in which the deliverables will be available according to the DoW. Month 0 marking the start of the project, and all delivery dates being relative to this start date.

3.4 Performance Indicators

With the majority of our indicators, we are on track or else have the processes in place to reach the required targets in the very near future. As BHL-Europe is currently in the phase of testing the ingest procedure, a large amount of data and content is in various stages of processing. This cannot be accurately reflected in the indicators but was detailed in M24 in D2.7 and will be detailed again in M36 with D2.8. More details on each indicator are provided in the following table, with notes below.

It should be noted that this is an intermediate report, thus not all details are yet available that are scheduled for M36. In order to provide the number of agreements with Rights Holders, for example, we need to do a survey among the content providers. As such a survey is not in the schedule for M30, no updated figure can be provided here.

Indicators	Expected vs. Actual							
	Year 1 (exp.)	Year 1 (act.)	Year 1.5 (act.)	Year 2 (exp.)	Year 2 (act.)	Year 2.5 (act.)	Year 3 (exp.)	Year 3 (act.)
Number of total accessible pages of biodiversity literature (in million)	17	17.4	35.3	21	40.19	41.0[a]	25	
Number of pages of biodiversity literature from European content providers to BHL-Europe (in million)	3.4	3.4	3.5	4.1	4.19	4.4[b]	4.8	
Percentage of literature available through Europeana	20%	0%	78%	50%	78.3%	81.6%[c]	100%	
Amount of metadata sets to be imported into the GRIB (in million)	0.15	0.15	0.3	2.5	0.3	0.9[d]	2.0	
Number of interconnected repositories	7	14	14	20	16	16[e]	30	
Number of content providers	20	21	19	25	25	28[f]	30	
Number of portal languages	1[1]	1	2	7[2]	7	7[g]	12[3]	
Page views through BHL Portal (in million)	1.0	3.7	4.1	2.0	4.7	5.0[h]	2.5	
Page views through BHL-Europe Portal (in million)	0.0	0.0	0.0	0.5	0	0[i]	0.1	
Digital biodiversity literature usage in Europe (in million page views)	---	1.8	2.0	---	2.1	2.4[i]	2.5	
Page views through Europeana Portal (in million)	0.25	0	0.015	1.0	0.06	0.06[j]	0.1	
Case studies of successful usage of the material by non-scientists	5	5	5	10	10	10[k]	15	
Agreements with Rights Holders / Publishers	2	7	7	4	86	86[l]	6	

Note (a): Number of total accessible pages of biodiversity literature (in million)

This figure is based both on the content we have on our servers and on the commitments made by the content providers with the signature of the MoU. At present it is only an estimate as no detailed content analysis has been performed. For more details see also D2.7 - Content analysis and management status report 2 (metadata, page numbers, content providers).

Note (b): Number of pages of biodiversity literature from European content providers to BHL-Europe (in million)

This figure is based both on the content we have on our servers and on the commitments made by the content providers with the signature of the MoU. At present it is only an estimate as no detailed content analysis has been performed. For more details see also D2.7 - Content analysis and management status report 2 (metadata, page numbers, content providers).

Note (c): Percentage of literature available through Europeana

The number of estimated total accessible pages of biodiversity literature (in million) at the end of the project is the 100% level. With the 100,561 objects currently accessible from BHL-Europe, we make 81.6% of 43.9 million pages available through Europeana. For more details of the translation of objects in pages see D2.7 - Content analysis and management status report 2 (metadata, page numbers, content providers). The increase in the last months is due to the increase of BHL-US content, which is reharvested regularly to reflect the up-to-date situation.

Note (d): Amount of metadata sets to be imported into the GRIB (in million)

This is the number of original records in the GRIB from the library catalogues FUB-BGBM, NAT, NHM, NBGB, RBGE, RBINS, UH-Viikki and SIL. The number is lower then originally expected. Two reasons can be given: (1) the number of ingested data is lower than the number of estimated data when we did the survey with the libraries. This is also responsible for the reduced number of GRIB records we expect for M36. (2) The data are provided in very heterogeneous formats causing difficulties during the ingest and de-duplication process. Library catalogues of four further providers are in progress to be ingested over the next weeks to increase the GRIB dataset (RMCA, UBER, HNHM, MOBOT, BHL) and further updates are expected for the data already ingested. For more details refer to D2.7 - Content analysis and management status report 2 (metadata, page numbers, content providers).

Note (e): Number of interconnected repositories

This is the number of BHL-Europe content providers from whom we have harvested metadata into our prototype. As we are currently refactoring the prototype and testing the new Pre-Ingest component, we have not further increased the number of interconnected repositories at present. However, interconnection of repositories currently happens at different levels (upload, mapping, pre-ingest, etc.) due to the implementation and development progress of the various system components. As we finally load the BHL-Europe system, this number will

change continuously to reflect the status of successfully-ingested material that is accessible through our portal.

Note (f): Number of content providers

Since the original submission of D2.7 - Content analysis and management status report 2 (metadata, page numbers, content providers) three more content providers have joined the network: Freie Universität Berlin, Universitätsbibliothek Frankfurt/ Main, Bibliothèque universitaire Sciences & Philosophie Rennes.

Note (g): number of portal languages

Seven languages were implemented in the development version of the portal for M24: Czech, English, French, German, Italian, Portuguese and Spanish. Drupal's localisation and internationalisation modules are used to facilitate the translation of the portal. The full system integration stopped the internationalisation process for now. However, we have set up the technical mechanisms to facilitate the translation of portal terms in the languages we need. For more details on the portal internationalisation refer to D3.7 - Key components documented for output of D3.5.

Note (h): Page views through BHL Portal (in million)

This figure is for the BHL Portal www.biodiversitylibrary.org/ provided by Google Analytics and cover the period from 17 October 2010 to 16 October 2011. This number of page views corresponds to 432,930 absolute unique visitors in this period.

Note (i): Page views through BHL-Europe Portal (in million)

As BHL-Europe has no fully functional portal except the prototype, we are not recording the usage. It should be noted that at the time we suggested this indicator, we anticipated usage of prototypes and preliminary portal versions. However, as we understood the risk of losing user commitment when offering a prototype which is not fully functional, we decided to not promote these prototypes beyond the consortium and not to monitor the usage. After a revision of the indicators we now suggest a change of the indicator and the addition of one further indicator.

Usage of BHL-Europe has two aspects, which need to be considered separately. First, the usage of the entire content in Europe is a measure of success for the BHL-Europe project. During the lifetime of the project, the number of European users of the content provided by the BHL-US portal increased significantly. This is shown by the new indicator (Digital biodiversity literature (content) usage in Europe), where the BHL-US content is taken as a proxy. Second, the usage of the BHL-Europe portal shows how quickly users are switching from the US to the European portal. This is a change in user behaviour although users still using the same content. Based on this separation and the expectations and plans we have made to date, the indicators are adapted.

Note (j): Page views through Europeana Portal (in million)

For the page views through Europeana, the Google Analytics of the BHL Portal www.biodiversitylibrary.org/ were evaluated. 42,080 page views of the BHL Portal between 17 October 2010 and 16 October 2011 were generated by Europeana users. Compared to the past values, no increase is obvious. So despite the increase of usage of the BHL-US portal in general, no increase of usage through Europeana is evident.

One reason for this is the stop of indexing of Europeana by Google. From the beginning of 2011, Europeana object pages were indexed by Google and traffic from search engines to the Europeana portal increased rapidly to 400.000 visits monthly (80% of all traffic). However, shortly after it was realised that most of the content displayed by Europeana was indexed already through other aggregators like BHL.

Google stopped the indexing resulting in a dramatic decrease of search engine visits to the Europeana portal from almost 400,000 visits a month in February-March 2011 down to about 120,000 monthly search engine visits in June-July 2011. Google has however started to re-index Europeana from autumn 2011 and since the new Europeana portal launch in October 2011 it has picked up again. This resulted in almost 3 million pages indexed by early December 2011 and this amount is growing with 150-200K per week.

This decrease was also reflected in the usage of the BHL-US content through Europeana. From 1 January to 30 April 2011, BHL-US counted 4741 visits from Europeana users, while from 1 May to 31 August 2011 only 1638 visits were counted, which is about 1/3 of the visits. Having the relationship of content to visits in mind (BHL-Europe is making up about 0.5% of the total Europeana content [20 million objects vs. 100,000 objects] but account for 0.2% of the total Europeana usage), the usage of BHL-Europe is not very bad. Please have in mind, that these numbers account for the people clicking through from Europeana; the usage of Europeana content in the Europeana portal can be expected to be much higher. The current usage statistics of Europeana also show one interesting aspect: the Europeana site only has double the number of visits compared to the BHL-US portals, having in mind that the BHL portal only covers a small fraction of content that is covered by Europeana. The potential for Europeana users should be much higher, but is not realised until now. This makes it difficult for the aggregators to generate usage of the content through Europeana.

Based on the above analysis, the usage of BHL-Europe content through Europeana was highly overestimated at the time we delivered the Description of Work. We anticipated a much earlier increase of the usage by the end users. However, we did not understood the user behaviour of the general public well enough to come up with a realistic estimate. Scientific users are using our content via the BHL-US portal currently very intensively. The attraction of the general interest reader was not successful enough to use the content via Europeana. This is not just a problem of BHL-Europe. Also Europeana is struggling with this topic and therefore intensified the end user marketing significantly. Some important actions Europeana has taken are the imprived search engine optimisation (mentioned above), the introduction and promotion of a search API and widgets (allowing for Europeana content to be searched and used through other websites) and the promotion of content through social media and virtual exhibitions.

However, it is not realistic for BHL-Europe to fulfil the indicator for year 3 as given in the Description of Work and the indicator is corrected based on the estimates we have available to date.

Note (k): Case studies of successful usage of the material by non-scientists

No update available right now.

Note (l): Agreements with Rights Holders / Publishers

No update available right now, but we have fulfilled the requirements for year 3 already.

4   Awareness and Dissemination

4.0   Overview of awareness and dissemination activities

In this reporting period the BHL-Europe project Web site[54] has continuously been further improved. The number of languages in which the portal is available could be increased to 10, now also including Finish (UH-Viikki) and Polish (MIZPAS). However, the main focus was on designing and implementing the overall layout for the user interface of the BHL-Europe portal.

The BHL-Europe Newsletter #8[55] was published in June. However, the Newsletter #9 is still in production. The next BHL-Europe Newsletter #10 is in preparation and will be synchronised with the Portal Release in November.

News on BHL-Europe and related projects were spread using the now fully implemented social media streams BHL-Europe Facebook[56], BHL-Europe Twitter[57], BHL-Europe blog[58] and BHL-Europe netvibes[59].

BHL-Europe activities were communicated throughout the Europeana network and Europeana and were promoted via the Europeana social media streams, like Facebook, Twitter and RSS.

BHL-Europe now has a site at ePractice.eu[60]. This is a portal created by the European Commission and jointly managed by the Information Society and Media Directorate-General. It offers a new service for the professional community of eGovernment, eInclusion and eHealth practitioners. It is an interactive initiative that empowers its users to discuss and

---

[54] http://www.bhl-europe.eu/

[55] http://www.bhl-europe.eu/de/newsletter/neue/bhl-europe-newsletters-8

[56] http://www.facebook.com/pages/BHL-Europe/151086001600041

[57] http://twitter.com/BHLEurope

[58] http://bhleurope.blogspot.com/

[59] http://www.netvibes.com/bhl-europe#Projects

[60] http://www.epractice.eu/en/cases/bhl

influence open government, policy-making and the way in which public administrations operate and deliver services. More work is of course necessary to upload more content.

In this reporting period a number of BHL-Europe partners were very active in approaching various target users through official conferences and publications. The list of publications on BHL-Europe written by BHL-Europe partners can be found on the BHL-Europe website[61] or the BHL Wiki[62]. An up to date list of events where BHL-Europe is presented by project partners is available on the BHL Wiki[63]. Presentations and posters[64] were adapted according to audiences. We also prepared small dissemination kits containing folders with annual report and project flyer. The project management group sent the appropriate expert(s) in our consortium to the respective conferences, in order to adequately represent the project.

There were two major dissemination events that took place during this reporting period. The first was the XVIII International Botanical Congress[65] in Melbourne, Australia in 23-30 July 2011. A symposium on Informatics tools for the semantic enhancement of taxonomic literature (Sym136) was held, which was devoted to bioinformatics. Jiří Kvaček presented a poster about BHL-Europe's visions. This conference was particularly important for BHL-Europe because of the unique opportunity to meet with colleagues from three other Global BHL nodes (BHL-US, BHL-Australia, BHL-China).

The second important event was the EuropeanaTech Conference[66] in Vienna, Austria, 3-5 October 2011. Our BHL-Europe partner NHMW organised several activities there: 1) A booth for providing information about the BHL-Europe project; 2) Distribution of dissemination kits including BHL-Europe folders, printed BHL-Europe Newsletters including the Annual report 2011 and the project flyers; 3) A presentation on the first drafts of the User Interface of the BHL-Europe portal during the conference, which was well received among the conference participants.

Furthermore CSIC organised the exhibition "Nombrando Especies" (Naming Species) which included a booth on BHL-Europe. The exhibition opened on the second week of October 2011. It includes: (1) an introduction on "What is in a scientific name?" via a comic like story; (2) a review of the International Institute of Species Exploration (Arizona State University) on the International Commission's Top Ten Species from 2008 to 2011; (3) a dedicated BHL-Europe stand with a short video explaining to the general public how digital libraries support the work on biodiversity.

NBGB published short notes describing the project in the internal newsletter for the staff of the Botanical Garden and in the Botanical Garden electronic newsletter. The BHL-Europe project was presented at the Librarians Day of the Federal Belgian Libraries Forum in Brussels, 17 May 2011, and at the European Botanical and Horticultural Libraries group Annual Meeting in Oslo, 19 May 2011.

During the German Librarians Day, 100. Bibliothekartag 2011 in Berlin 6-8 June 2011, Dennis Zielke (UBER) presented the BHL-Europe project at the booth of the German Initiative for Networked Information (DINI e.V.) within the exhibition area. During the

---

[61] http://www.bhl-europe.eu/de/publikationen/veroeffentlichungen

[62] https://bhl.wikispaces.com/BHL-Europe-publications

[63] https://bhl.wikispaces.com/dissemination+of+project+by+partners

[64] https://bhl.wikispaces.com/BHL-Europe-posters

[65] http://www.ibc2011.com/

[66] http://www.europeanaconnect.eu/europeanatech/

Workshop on virtual research environments (VRE) at the Humboldt-University 21 June 2011 Dennis Zielke presented the BHL-Europe project to a staff member of the Konstanz University Library. The University of Konstanz is working on a new DFG-project called movebank (documenting moving traces of animals worldwide).

Together with Nicole Hanquart from NBGB, Tine Mallaerts (RBINS) prepared a Dutch and French presentation[67] about the project for the "Dag van de bibliothecaris" in Brussels, Belgium. On this day about 60 librarians from different Belgian federal institutes came together and discussed different subjects.

Walter Berendsohn (BGBM) gave details on the aims and the infrastructural setup of BHL-Europe during a talk titled "Sustainable infrastructures for e-taxonomy in Europe" for the Royal Danish Academy of Sciences and Letters during the Symposium on "Biodiversity in the Silicon Age", Copenhagen 17-18 May 2011.

On 28 October, NHM London organised a symposium in the honour of the 150th year of the birth of Charles Davies Sherborn: "Anchoring Biodiversity Information: From Sherborn to the 21st century and beyond"[68]. This symposium brought together an international panel of experts on bibliography and biodiversity bioinformatics, linking a view of the past with an active debate on the future of the related fields. BHL-Europe co-sponsored this event and was represented with one talk and three posters. Charles Davies Sherborn (1861-1942) was an English zoologist and provided the bibliographic foundation for zoological nomenclature with his magnum opus Index Animalium, an extremely important multivolume work of 9200 pages which appeared in 11 volumes between 1902 and 1933.

The Annual Conference of the "Paläontologische Gesellschaft" was held from 12 to 15 September 2011 at the University of Vienna, Austria. Henning Scholz prepared and held a workshop for the participants to introduce and demonstrate operation of the variety of online tools available to taxonomists for finding literature. In about three hours, the various BHL portals were introduced with detailed and interactive demonstrations.

RMCA finalised the integration of the Europeana API into the RMCA's website at http://bit.ly/AfricamuseumEuropeanaAPI. An article will appear in the next museum's newsletter. RMCA also discussed with social media experts from BHL-US on how to feature RMCA's scientific and BHL-Europe project related activities in the BHL blog. Furthermore, RMCA also created a proof of concept for a BHL-Europe virtual exhibition around the theme of Lemaire's expedition to Katanga[69] showing connections between the published part of the expedition literature with other RMCA content and content in Europeana. In the frame of this activity, RMCA staff attended the E-corpus training, where an introduction to the creation of virtual libraries and satellite-websites within the e-corpus platform was presented.

The 2011/12 intake of postgraduate and diploma students starting courses at RBGE were informed about BHL and BHL-Europe as part of student library induction tour.

---

[67] http://www.slideshare.net/Bibforum/bhl-europe-biodiversity-heritage-library

[68] http://www.shnh.org.uk/meetings/future-meetings/anchoring-biodiversity-information.html

[69] http://cybertaxonomy.africamuseum.be/drupal7/

Figure 4: Henning (left) and Dennis (right) during the individual time trial in Halle.

A few BHL-Europe staff members were equipped with bike jerseys to promote BHL-Europe to the general public during sports events. The first event we used the jerseys was a cycling time trial in Halle[70] on 29 May 2011. The second event was the Berlin Triathlon[71] on 19 June 2011. Dennis Zielke and Henning Scholz were the two starters in both events for BHL-Europe. The success of both starts justifies the extension of these activities. Two more jerseys were ordered for the two NHMW colleagues in Austria for disseminating BHL-Europe to other people at other sport events.

The list of publications on BHL-Europe written by BHL-Europe partners can be found on the BHL-Europe website[72] or the BHL Wiki[73].

In this reporting period, Henning Scholz finalised the BHL-Europe contribution for the International Innovation Report in collaboration with ResearchMedia. This report was published in May 2011 and has a 3-page publication on BHL-Europe. In the following, ResearchMedia also produced a brochure with additional material. Brochure and article is available for dissemination activities of BHL-Europe. We have the full copyright for distribution of the material in printed form or through the project website.

---

[70] http://www.velo-ev.de/index.php?page=135

[71] http://www.berlintriathlon.de/

[72] http://www.bhl-europe.eu/de/publikationen/veroeffentlichungen

[73] https://bhl.wikispaces.com/BHL-Europe-publications

4.1   Events and meetings

List of Events & Meetings

Below is a list of contractual meetings and workshops of members of the BHL-Europe consortium to discuss and work on various aspects of the project.

Name	Location	Start Date	End Date
BHL-Europe Annual Review	Tervuren, Belgium	07/06/2011	10/06/2011
Public Core Review Meeting	The Hague, the Netherlands	29/08/2011	03/09/2011

Below is a list of conferences and events attended by members of the BHL-Europe consortium in order to disseminate the project to the various audiences or to collect important information for the progress of the project.

Partner	Title	Start date	End date	Location	Level of dissemination
NBGB	Forum des bibliothèques des services publics et établissements scientifiques fédéraux de Belgique	17/5/2011	17/5/2011	Brussels, Belgium	talk
NBGB	European Botanical and Horticultural Libraries	17/5/2011	21/05/2011	Oslo, Norway	flyers
RMCA	EU infoday "Conflict and peace in a changing international context: funding opportunities, matchmaking, networking in FP7"	06/07/2011	06/07/2011	Brussels, Belgium	flyers
RMCA	Food, Agriculture and Fisheries, and Biotechnology Info Day and Brokerage Event on Call FP7-KBBE-6-2012	15/07/2011	15/07/2011	Brussels, Belgium	talk
NMP	XVIII International Botanical Congress 2011	23/07/2011	30/07/2011	Melbourne, Australia	poster
RMCA	BELSPO (Belgian Science Policy Office)	19/08/2011	23/08/2011	Brussels, Belgium	talk
RMCA	Seminar for leading staff of the RMCA on EU project management	01/09/2011	01/09/2011	Brussels, Belgium	talk
MfN	Jahrestagung der Paläontologischen Gesellschaft	12/09/2011	16/09/2011	Vienna, Austria	talk, workshop
RMCA, NMP	CETAF general meeting	12/09/2011	14/09/2011	Madrid, Spain	talk
NHMW	Europeana tech	03/10/2011	05/10/2011	Vienna,	flyers,

				Austria	dissemination kit, poster
RMCA	TDWG 2011 Annual Conference	16/10/2011	21/10/2011	New Orleans, USA	poster
NMP	Czech-Slovak-Polish Palaeontological Conference	27/10/2011	28/10/2011	Bratislava, Slovakia	poster
RMCA	2011 DLF Forum	31/10/2011	01/11/2011	Baltimore, USA	poster

5 Conclusions

BHL-Europe is entering the final half-year of the project and is about to launch the first public versions of the BHL-Europe system. Many components have been developed and integrated, and more features are still to come. The ingest process is still in the reviewing and testing phase, but bulk ingests will start soon in order to have all BHL-Europe content available for the final release in spring 2012. New resources were found to finalise the Pre-Ingest tool, and the newly-created Ingest Manager position will support our content providers with any ingestion problems. Dissemination work is progressing well and BLE should significantly assist in reaching non-scientific target users. Meanwhile we have mechanisms in place to keep BHL-Europe operational after the end of the project in April 2012. Content and services will also be available through Europeana and the Global BHL network with BHL-Europe being an important player.