PROJECT DOCUMENTATION

INTERNAL INTERIM PROGRESS REPORT

Project: BHL-Europe

Release: Final

Date: 4 March 2011

Author: Jana Hoffmann

Table of Contents

Internal Interim Progress Report

1   Purpose

To report on a quarterly base the status of work for each member of the consortium. This report will not duplicate the content of the official progress and annual reports requested by the EC every 6 and 12 months, respectively.

2   Follow-ups from previous reports

For further information on past activities we refer to [Progress Reports](#)[1] D1.1, D 1.2 and D 1.4.

3   Activities

General Remark

The following activities reported here are taken from the status report of the following partners:

MfN, NHM, AIT, ATOS, FUB-BGBM, UGOE, NHMW, LANDOE, HNHM, UCPH, NAT, NBGB, RMCA, RBINS, BnF, MNHN, CSIC, RBGE, Sp2000, UH-Viikki, UBER, EDLF. (22 partners!)

Not submitted reports:
MIZPAS

This time, almost all of the partners actively carrying out work for BHL-Europe provided a status report. However, the quality of the status reports has to improve in terms of submission of narrative reports rather than only keyword or bullet points; and writing paragraphs per task rather than months. It would also be helpful if partners would assign their tasks to the task (numbers) described in the Description of Work (DoW).

WP1

1.1.1 - 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.

2nd instalment
In this reporting period the 2nd instalments were transferred to the partners. Some major delays in transferring the budget to individual partners are due to incomplete reporting and unfulfilled tasks on the partner's side.

Recruitment of new BHL-Europe employees
We are welcoming the following new staff members in the BHL-Europe team:
Muriel Amar (BnF) - She is librarian and the right hand of Frédéric Martin, head of the department of Gallica.

---

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

Manuel Sanchez (CSIC) - He is official collaborator of the CSIC team for BHL-Europe. Manuel is actively contributing in the digitisation, building of metadata and other tasks with Marian Ramos and Antonio Valdecasas.
Jiří Frank (NMP) - He is officially working for BHL-Europe since 15 January and was accepted for leading the virtual exhibition subproject.

Kees Hendriks (NAT) left the NCB Naturalis and therefore also the BHL-Europe team 1 December 2010. He started working for the Cultural Heritage Agency of the Netherlands. Tom Gilissen will be the primary NCB Naturalis contact person for BHL-Europe.

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

Consortium agreement
The consortium agreement was revised by several partners and finalised during the final stage of the last reporting period. The meeting in London in December was used to collect and exchange signed documents. A small number of partners have not provided their signatures yet and we follow up with them to get the signatures soon. The final version of the consortium agreement is still available on the BHL Wiki[2].

Project Management Group conference calls (PMG calls)
A series of PMG conference calls were held to communicate project progress and discuss pressing topics in the Project Management Group. The minutes of the calls can be found on the BHL Wiki[3].

Content Provider and Technical meeting, NHM London, 30 November to 3 December 2010
The last meeting of the BHL-Europe consortium was held in London in the first days of December and organised by NHM and MfN. The agenda and relevant minutes of the meeting can be found on the BHL Wiki[4].
For the 30 November, the 2nd intermediate review meeting was scheduled. This review was based on the revised deliverables presented in Vienna in May, the new deliverables for M15 and M18 and the hardware setup at NHM. The WP leader, the project management assistants and few invited experts were representing the project progress and deliverables. Our EC desk officer and the two external reviewers were participating as well. We got a green flag from the EC. The review report is available on the BHL Wiki[5].
After the official review meeting, we discussed financial and administrative questions with our desk officer and then had the BHL-Europe PMG meeting. We discussed the high level project management tasks and identified responsibilities for current and future tasks within the project.
In the morning of 1 December, the BHL-Europe progress report (project progress) was presented to the consortium in a plenary session. Afterwards, we had an update on Europeana and an introduction to the new OpenUp! project that will start in March 2011. During the meeting in London, 6 working groups were scheduled over the two days of the content provider meeting and were introduced in the plenary session. After lunch, the working groups were established and started their meetings:
(1) The **Best Practice Guide Workgroup** meeting was a follow-up of the workshop held in Linz a few weeks before the London meeting (for more information see WP2&3 section of this report).

---

[2] https://bhl.wikispaces.com/BHLE_WP1
[3] https://bhl.wikispaces.com/BHL-E_PMG_calls
[4] https://bhl.wikispaces.com/BHLE_Fall+Meeting_agenda
[5] https://bhl.wikispaces.com/BHLE_review_reports.

(2) The **GRIB Workgroup** meeting was divided into two sessions: BHL-Europe & GRIB data workflow and deduplication, dealing with different scenarios of integrating the GRIB into the BHL-Europe system, and the GRIB Review (for more information see WP2&3 section of this report).
(3) During the **Communication Working Group** (CWG) meeting the organisation of the WP5 work among the two major partners (NMP and RBINS) and current open tasks were discussed. Creating a concept for the virtual exhibitions that BHL-Europe is going to build was the major task for this meeting. Furthermore, a new workflow to get content for BHL-Europe's social media (Weblog, Facebook, Twitter) that can be re-used for the Web site and the quarterly newsletter was established (for more information see WP5 section of this report).
(4) The **BHL-Europe Portal Development Workgroup** meeting was divided into two sessions, the first dealing with taxonomic intelligence; the second focussing on other portal development questions (for more information see WP2&3 section of this report).
(5) The **Pre-Ingest Round Table** was held with BHL-Europe content providers to discuss Pre-Ingest guidelines and procedures implemented by AIT. The meeting provided deeper insights into the Pre-Ingest development process and information about the future roadmap (for more information see WP2&3 section of this report).
(6) The outcomes of the **Business Plan Workgroup** meeting are summarised further below in this report section.
The Content Provider meeting was followed by the TMB and Technical Meeting on 3 December. Major agenda items were the implementation of the BHL-Europe Schema, the handling of the BHL-Europe code, preparation of D3.7 and the continuation of the GUID discussion. It was agreed to continue with the new schema, start with test mappings and investigate further metadata standards for taxonomic metadata. For D3.7, a working group will build the framework and identify responsibilities to deliver the document. We also agreed that BHL-Europe will be open source, but the details need to be worked out (for more information see WP2&3 section of this report).

Email Server
The following mailing lists are running on SYMPA and are in use:
bhl-e@lists.hu-berlin.de - BHL-Europe partners and associates
bhl-e.cp@lists.hu-berlin.de – BHL-Europe content provider

Liason with other EU projects and Europeana
FUB-BGBM communicated with current and future EU projects like OpenUp!, EDIT, SYNTHESYS and LifeWatch for integration of complementary activities.
There is ongoing liaison between BHL-Europe and Europeana and coordination with the Europeana Foundation Office.

1.1.3 - Production of reports and internal quality assurance
In this reporting period the activity reports M18-20 were collected from the consortium partners. Furthermore the document on quality assurance procedures, M 1.2[6], was revised and adapted to improve the reporting of project's outcome and the communication of the project's progress within the consortium. The new version of M1.2 will be uploaded to the BHL Wiki and distributed to the consortium once finalised.

1.2.1 – Work-package and task coordination
Each work package leader was asked to produce a work plan for upload to the project server. Currently, draft work plans for all WP are available and finalisation of the plans is expected soon. These plans will be made available through the project server and the BHL Wiki.

1.2.3 – Timely production of deliverables

---

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

All deliverables for M18 including the last progress report were available for the 2nd intermediate project review on 30 November. The paper copies were handed over to the EC in London. No deliverables are required for M21 of the project, but preparations are underway to have the nine deliverables for M24 ready in time.

1.4.3 - Business plan development

BHL-Europe has to prepare a first overview of the business plan for long-term sustainability (D1.5) in M24 of the project. Incorporating experiences of other projects, in particular Europeana, a draft document was prepared for the last meeting in London (December 2010) to collect ideas and questions that need to be discussed. As a result of the meeting in London, 5 scenarios for the future of BHL-Europe were defined: stop all, maintain, slow incremental growth, keep going and growing, scale up. A SWOT[7] analysis was performed to compare the scenarios. Now we are in the process of identifying resources for the scenarios and work on a next draft of the document. The document needs to fulfil the requirements of the deliverable, but has to be as unrestrictive as possible to have some flexibility for the final version of the business plan.

WP2

2.1 - IT development

GRIB prototype development and review

A major achievement was the connection of the newly set up Single-Sign-On (SSO) Identity Provider (IdP) for the GRIB to the EDIT SSO Federation. This will allow users to use different online-services with only one login.

The GRIB has been presented at the 3rd EDIT publishing meeting 7-8 October in Copenhagen, Denmark and during the Electronic Publishing Working Group at the Humboldt University Berlin, 29 October (GRIB and UBER).

In preparation of the user review and technical review of the GRIB in London MfN, UBER and the VZG prepared the version 0.7 of the GRIB manual, available on the BHL-Wiki[8].

From 1-3 December the GRIB was discussed during the BHL-Europe Content Provider and Technical Meeting in London. During the plenary on the first day the state-of-the-art of the GRIB was presented to the BHL-Europe consortium, followed by two workgroup sessions. During the first session, BHL-Europe & GRIB data workflow and deduplication*, different scenarios prepared by MfN and BGBM for integration of the GRIB into the BHL-Europe system has been presented and discussed. The integration of the GRIB into the BHL-Europe portal should aim at fulfilling the following three criteria: 1) Grouping duplicates in the result sets in the BHL-Europe portal. 2) Search for either scanned or not-scanned literature as well as combined search of scanned and not-scanned literature. 3) Lookup of digitisation status and nomination of literature for digitisation. Different ways for realising the presented use cases were discussed e.g. ingest of GRIB data into the OAIS or only remote access over the BHL-Europe portal. Their feasibility has been agreed on in principle, but the decision on which way of integration will be used has been postponed. In the second session the review of the GRIB took place. Andreas Krausz (VZG) gave two presentations on 1) GRIB Data Model, 2) GRIB Data Storage Model, as a follow-up of the GRIB meeting during the BHL Global Meeting in Woods Hole. He presented his ideas on how super-record could be realised within the central library system on which the GRIB is based, as well as what implications this would have regarding the Digitisation Management Tool.

A super-record is a bibliographic record that combines data about a unique publication from different sources. It is created by deduplication of a set of bibliographic records from the partners library catalogues. Individual records are matched and merged together and the source information is kept for reference purposes. In addition hyperlinks to the full text of the respective

---

publication will be added if available. Librarians and scientists will be able to identify and mark records where no digitised full text is available for managing their digitisation process.

This aspect has been identified as the major task during the review to be fulfilled with regard to the development of the GRIB. Furthermore, several suggestions on the enhancement of the GRIB have been documented and will be followed up.

A closed follow-up side meeting with colleagues from SIL took place on the second day of the meeting to further discuss the implications of the review and to align the GRIB Digitisation Management Tool concept with BHL-US.

The work on improvement of the GRIB and further development started shortly after the meeting in London. The suggestions on improvement and development of the super-record will be implemented until the end of February; some less pressing suggestions will be implemented until end of April 2011.

Submission of data to the GRIB

HNHM provided their entire Central Library Catalogue (not all the records are available at the moment in digital form) for import to the GRIB. They purchased and installed the Z39.50 interface module (from non BHL-Europe sources) for library systems following a former concept of the GRIB import. This concept changed and a MARC export was requested in the meantime. Submission of data to the GRIB was delayed due to bugs which could be eliminated in the meantime.

Our partner RBINS contacted LIBISNET – the organisation behind the library catalogue – in order to evaluate what has to be done to prepare the library catalogue in a way that it will be possible to send it to the GRIB by the end of January 2011.

As part of the ongoing provision of RBGE library catalogue data for the GRIB, also retrospective cataloguing of monographs is carried out by RBGE.

2.2 - Analysis of domain content

UCPH gathered literature for destructive scanning and analysed for e.g. two other important journals including more than 150,000 pages that will be scanned in the near future. The time for collecting the literature has not been (and will not be) included in the project. UCPH hopes to have about 50,000 pages (two journals) ready scanned by the end of January. Another two journals (one of them > 100,000 pages) will be digitised in the destructive way in future and UCPH hopes to fulfil their commitment of providing 100,000 pages before the summer 2011.

RMCA worked on further content selection for digitisation and updated their common ScanList on the RMCA's African Zoology Department, Biodiversity Information and Cybertaxonomy Group Wiki[9]. Related material to Charles Lemaire expeditions in Congo (1901-1903) have been selected and digitised.

BnF analysed additional already digitised titles for inclusion into BHL-Europe (Revue générale de botanique, Annales des Sciences naturelles). Furthermore a set of monographs has been selected in order to increase the part of monographs in the BHL-Europe portal consisting of 524 directories which mean 8 titles of periodicals, 533 digitised documents and 226,715 digitised pages. A list of the first set of content that will be provided to BHL-Europe can be found in the table below:

TITLES	Number of documents	Number of pages
Annales de la Société linnéenne du département de Maine-et-Loire [Texte imprimé]. 1853-1879 [I-XVII]	13	3,536
Annales des sciences physiques et naturelles, d'agriculture et d'industrie [Texte imprimé] : publiées par la Société d'agriculture de Lyon	73	42,927

---

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

Archives de zoologie expérimentale et générales Paris 1872 et Archives de zoologie expérimentale et générale. Notes et revue	87	35,229
Bulletin de la Societe Linneenne de Normandie 1855	138	29,146
Bulletin de la Société zoologique d'acclimatation- Bulletin de la Société d'acclimatation- Bulletin de la Société nationale d'acclimatation de France- Revue des sciences naturelles appliquées- Bulletin de la Société nationale d'acclimatation de France: revue des sciences naturelles appliquées	82	43,509
Bulletin de la société zoologique de France 1876	81	27,092
Revue zoologique Paris 1838 fusionne avec Magasin de zoologie donne Revue et magasin de zoologie	11	4,704
Revue générale de Botanique Paris 1889, Bonnier, éd.	48	40,572
TOTAL	**533**	**226,715**

LANDOE focused during this working period on improving the calculation on how much content is still available in Europe for digitisation. More precise numbers should be included by the end of April 2011 in the second content analysis and management status report D2.7.

In a collaborative project, UGOE and the University Library of Firenze prepared the coordination of digitisation of Italian books within the UGOE library. 430 volumes (mostly Italian books) were transported from Firenze to Göttingen at the beginning of November and digitisation of the monographs was started immediately after as part of the EZOOLO project[10] (Early Zoological Literature Online). The first steps included unpacking the books and checking for quality by UGOE's restoration department together with Caterina Guiducci from Firenze (MSN) who came for this purpose to Göttingen. The serials will be digitised after the monographs. Since some of the serials contained volumes which can not be digitised due to quality problems with the books, which would not allow them to be digitised without risking damaging these volumes, the BHL ScanList was checked for presence of these serials in other BHL libraries. A scan requests for these serials was sent to BHL-US and the UGOE bids (or scan suggestions) were removed from the BHL ScanList.

For our partner MNHN preliminary work was necessary before starting the outsourced scanning activities. More precisely, they continued and finished the description of bibliographical and physical data of each issue number, for a first batch of 200,000 pages (about 1,800 volumes of MNHN scientific publications). Currently, they are proceeding to receive the files and to make a step-by-step quality control. Further selection and analysis of new content to be digitised has been carried out. In accordance with copyright requirements, a list of estimated content has been prepared. Before deciding to digitise the content, an estimation of the number of pages will be done.

UH-Viikki edited their file naming guidelines for the outsourced scanning. They also investigated how to get relevant metadata from their databases Helka and Arto in XML format using the Voyager Select tool.

They ordered and received 22,700 pages of journals digitised by their digitisation provider Vaasan Micro Copy Oy, performed quality control of these files and uploaded the first items to the BHL-Europe server (Meddelanden af Societas pro Fauna et Flora Fennica 49 and 50). The reprint collections of the Botanical museum of Finnish Museum of Natural History has been studied and evaluated for future digitisation during this reporting period.

RBINS discussed with the library personnel which digital publications can be offered to BHL-Europe and set up priorities.

As a result of working with BHL-Europe, RBINS is now working on a digital library for the institute using the greenstone software. The library has bought an external server for the storage of the digital documents of the institute (no BHL-Europe budget has been used for this purchase), the Greenstone software is installed and for the coming months all existing scanned/digital articles

---

[10] www.animalbase.uni-goettingen.de/

and books that are currently spread at the institute will be identified, collected and imported to this system. This digital library will make it easier to transfer the content to BHL-Europe too.

2.3 - Management of the content acquisition process

Management of current content providers
The management of the content acquisition process was a further task within this reporting period. For this purpose a new [Wiki page](#)[11] has been created within WP2 for the content management and the content acquisition process.
In order to make the content management process more visible to all BHL-Europe partners, the estimations for uploading data by our content providers is published on the above mentioned Wiki page. Furthermore the current status within the data workflow is documented to allow our partners as well as the technical team to catch up at any time in order to facilitate communication and future planning of their own tasks.

Managing data upload to BHL-Europe server
This task has been fulfilled in cooperation with WP3. In order to coordinate the required FTP accounts for our content providers, a [Wiki page](#)[12] has been established:
All content providers have been asked to provide a person in charge and a backup contact for the content upload within their institution. According to the provided information, FTPS accounts have been created for each content provider and logon details have been sent out in November by WP3. As it turned out that some content providers had difficulties using a FTPS connection within their own institution, all content providers have been asked during the Content Provider and Technical meeting in London to check their FTPS connection and report any problems.
The above mentioned Wiki page contains a list of all BHL-Europe content providers, general information about the FTPS account and a 'How to use' section. For those content providers who had problems with the primary FTPS solution a second solution has been provided. Up to date now further problems have been reported by our content providers.

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

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 shall be understandable for every person working with it. They shall guide new and current contributors in a simple and clear way through a digitisation workflow from the either print or 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.
During this reporting period we continued with the work on the first delivery of the Best Practice guidelines and standards. A first meeting was held on 8-9 November 2010, hosted by LANDOE in Linz, Austria with the aim to bring together all participating parties (LANDOE, MfN, Naturalis, NHM, Ernst Mayr Libray – BHL-US, UBER, HNHM) and to discuss existing workflows. The outcome of the first workshop was a first draft of table of contents and assigning contributors to most of the chapters. In order to assist the authors all workshop participants were asked to contribute questions that could be answered within the several chapters. A second and shorter workshop was held one month later during the Content Provider and Technical meeting on 1

---

[11] https://bhl.wikispaces.com/Content+_Management
[12] https://bhl.wikispaces.com/BHL-Europe_Data_Upload

December 2010. During the second workshop in London, the table of contents has been revised by going through each chapter and the related questions provided by the participants in order to make sure that the authors understood their tasks. Furthermore, authors have been assigned for the remaining chapters and the deadline for delivering chapters has been appointed to 5 January 2010.

More details about the Best Practice Guide, workshop, notes, deadlines, participants and a current working document can be found on the [BHL Wiki](https://bhl.wikispaces.com/BHLE_WP2_BPG)[13].

Assist partners in scanning operations

Assisting partners in implementation and evaluation of scanning operations is a task within WP2. For this purpose the scanning workflow management tool GOOBI was taken into account. A first meeting with intranda GmbH - the Goobi software provider - was held beginning of November with members of MfN and UBER. Furthermore a survey has been designed and conducted to better understand the needs and requirements on the digitisation workflow of our partners. The survey has been carried out by our partner UBER and each content provider received an invitation to participate in the survey mid of November 2011. After four weeks the survey has been closed and in totally 13 content providers out of 23 participated in the survey. Following partner participated in the survey: LANDOE, UCPH, NATURALIS, NBGB, RMCA, RBINS, MNHN, CSIC, RBGE, UH-Viikki, UBER and two BHL-US partners (Smithsonian Institution, Harvard University/ Ernst Mayr Library).

The evaluation of the survey allows UBER to test already available digitisation tools and software in regard to BHL-Europe content providers which can be introduced in the deliveries D2.6 and D2.9. For this purpose UBER is working on a list of test criteria based on the survey.

Analysis and preparation of (meta)data for Pre-Ingest

HNHM puts an emphasis on journal volumes at article level as books are regarded having a lower priority for scanning at the moment. The reason for this is that many of the former Eastern Bloc countries' (like Hungary) literature is handled/regarded as grey literature because of the weak availability and marketing. During this reporting period HNHM worked on the compilation and key-in of the bibliographic metadata sets (xml format). More detailed, HNHM takes all issues in hand and key-in the bibliographic metadata (authors, title, alternative title, volume, issue and page numbers, keywords if available, abstract(s) if available, notes if relevant) into a xml file. This workflow has already been used by HNHM before the scanning itself started and is based on their long working experience.

Following the primary cleaning HNHM converts data into a temporary MySQL data base where they make further cleaning. Another job is to cut scanned material (puff files) to articles and connect the corresponding metadata. After these steps have been finished, metadata is imported into their "final publication data base" and they publish the metadata together with the scans (pdfs) on the HNHM webpage. By completing all the described steps HNHM is able to export the content for the upload to the BHL-Europe server.

The goal of HNHM is to put an emphasis on a high quality instead of using a "quick and dirty" workflow which of course needs manpower. Sometimes it was also necessary to rerun parts of the workflow because of the modifications/refinements of the xml structure in order to fill in of additional data.

Within a frame of a former experiment HNHM also extracted taxonomic metadata of Vertebrata Hungarica (a dead title). All taxon names occurring in the title, abstract and keywords were extracted and inserted into an xml structure at the highest available taxonomic level. For example if only *Turdus merula* (genus and species name) or its common name mentioned in the source, than the higher taxonomic categories are also filled out (Vertebrata, Aves, …). In case of an ecological paper (with a long list of taxa in tables) only the highest relevant categories are used (e.g. singing birds). The granularity of metadata is on the journal article level. The following journals' metadata was completed by HNHM: Fragmenta Botanica Hungarica, Studia Botanica

---

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

Hungarica. Quality assurance and database conversation of the following journal data was completed: Parasitologica Hungarica, Vertebrata Hungarica. The formerly finished datasets were completed with the exact date(s) of publication (both the hardcopy and the digital publication date since they might be different, especially nowadays, and this way digital publications can also be incorporated) since this has a special importance in taxonomy because of the priority rule in nomenclature.

UCPH scanned ca. 50,000 pages, including two journals, Friesia and Botanisk Tidsskrift. The scans are for the time being been quality controlled and will be delivered at the end of January 2011. UCPH had problems with the pages not being horizontal after the scanning and tries to find a tool to correct them. Furthermore, UCPH has been gathering literature for destructive scanning and has e.g. two other important journals including more than 150,000 pages that will be scanned in the near future. The time for collecting the literature has not been (and will not be) included in the project.

RMCA digitised about 14,000 additional pages of biodiversity literature from the zoological department in the framework of other projects and in-house activities. The coordinators of these projects were keen on having those provided to BHL-Europe as well. All files have been controlled and renamed on page level. Further they processed metadata by inserting new data and enriching their catalogue with additional information. Beginning of November they had a meeting with the library staff and Metafro (RMCA catalogue system) developers about unification of their catalogues and future merging with LIBIS group (Leuven). During this meeting OAI-PMH was discussed as a solution for harvesting metadata by BHL-Europe. RMCA has around 20,000 pages within the Public Domain ready for upload and plans to upload this content in January 2011.

BnF and Naturalis sent first metadata sets for Pre-Ingest during this reporting period.

MNHN is currently reviewing their metadata template (.xls document) as they have difficulties regarding the information management (different levels). A first transfer of their metadata and scans has been tried out during this reporting period and difficulties have been realised with the FTPs connection as well as with the backup solution WebDAV. Finally they returned using the FTP connection. The test files used during the test upload refer to 100,000 pages which are already online on their [own portal](#)[14] under the menu "Textes numérisés".

NBGB checked the quality of their bibliographic metadata which will be an ongoing process for the next reporting period. Further they continued searching for content that can be digitised in future.

CSIC digitised Vol.1 and Vol. 2 of Fauna Ibérica and uploaded Vol. 0, Vol.1 and Vol. 2 to the BHL-Europe server. Further, Dirtax – the directory of Spanish Taxonomists – has been digitised and attached with a metadata file.

Work on the quality assurance (QA) and post-scanning processing of the 46 volumes (22,000 pages) of the Notes of the Royal Botanic Garden Edinburgh (NRBGE) and the 8 volumes of the Flora of Bhutan (3,400 pages) was completed by our partner RBGE. Feedback provided after submission of scans and associated metadata of one volume of the NRBGE for pre-ingest testing identified that bibliographic metadata at the title (volume) and item (part) level was required. This metadata was created and will be submitted as part of the ingest process. Therefore, content for ingest will include .tiff files of scanned pages, structural metadata in Excel format and bibliographic metadata at the title, item and article level in .mrc format. A sample set of this data comprising the first three volumes of the NRBGE was transferred to the NHM server in December 2010. Flora of Bhutan volume 3 part 3, scheduled for inclusion in BHL-Europe was blocked because one copyright holder was not prepared to permit digitisation.

Expansion of the content provider network

Attracting new partners until the end of the project is a major task within WP2. For this purpose the "BHL-Europe in a Nutshell" document has been finalised and introduced during the Content Provider and Technical meeting. This document aims to be a very brief but informative document

---

[14] http://mussi.mnhn.fr

including the most important facts about BHL-Europe, benefits and IPR issues and has been created in collaboration with WP4. Several content providers (RBGE, CSIC, RMCA) have been indentified to help in the content acquisition process and were asked to use their connections in order to attract new content providers within their own countries. Further the "BHL-Europe in a Nutshell" document has been translated into Spanish by our partner CSIC during this reporting period, further translations into French and Russian language are planned for the next reporting period. Further information about content acquisition, e.g. the "BHL-Europe in a Nutshell" document can be found on the BHLWiki[15].

During the last board meeting of the GfBS[16] (Gesellschaft für Biologische Systematik,) at the MfN in October it was discussed that the early issues of the societies journal ODE[17] (Organisms, Diversity and Evolution) published by Elsevier can be offered to BHL-Europe. However, it needs to be further investigated if there are any rights issue with the former publisher of the ODE, but as the GfBS is the right holder there should be no problem. Furthermore, the GfBS board will discuss with their current publisher, Springer, if they will let them make issues older than 2 years open access and available in BHL-Europe. Jana Hoffmann as a board member of the GfBS will follow up on this.

MfN was successful in raising funds for a book scanner. BHL-Europe was involved to some extent in the process to date. This scanner will be purchased and installed in early 2011. Plans are not made when and to what extent MfN will contribute content to BHL-Europe. It is anticipated, that MfN can act as a content provider for BHL-Europe before the project ends in April 2012.

RMCA discussed and arranged with other Belgian institutions, societies and other projects to provide additional content to BHL-Europe like for example discussion of collaboration with RMCA's history unit who are hosting archival material.

A meeting has been hold at CSCI with A.G. Valdecasas and the CSIC director of electronic journals to identify the availability of already digitised material for BHL-Europe.

RBGE got involved with the extension of the BHL-Europe content provider network. Approaches for advice and assistance with digitisation of a society journal was received from a committee member of the Forth Naturalist and Historian, and one individual seeking to deposit a scanned public domain biodiversity work (quite scarce) in a suitable digital library. A draft local dissemination plan was commenced, with further work to be done with this from January 2011 onwards.

64 authors who have contributed papers to Notes from the Royal Botanic Garden Edinburgh were contacted. Positive feedback about the digitization of these papers and the work of BHL and BHL-Europe was received from 36 of these authors, 28 did not respond.

Visit of the Berlin State Library
In November 2010, a delegation of BHL-Europe visited the new scanning centre of the Berlin State Library. One goal of the visit was to investigate possible areas of collaboration in digitising book collections. MfN can learn from them in establishing own scanning facilities (see above). MfN also discussed existing scanning projects and possible overlap with BHL-Europe priorities. However, all their scanning projects don't focus on biodiversity literature; although very few books are available that might be considered by BHL-Europe as well (see http://digital.staatsbibliothek-berlin.de/dms/ - in German). Further, MfN also discussed their scanning workflow to identify possible input for the BHL-Europe Best Practice Guide. The State Library is using the scanning workflow software Goobi that we investigate within the project to be used by our partners.

2.3.3 - Coordinate with Europeana

The Europeana Council of Content Providers and Aggregators (CCPA)

---

[15] https://bhl.wikispaces.com/Content+_Management
[16] http://www.gfbs-home.de/
[17] http://www.springer.com/life+sciences/ecology/journal/13127

After the successful plenary meeting of the CCPA in The Hague in October (see D1.4 for details), the five working groups are now officially launched and the chairs of each working group are asked to set up the groups by end of February. Every content provider of BHL-Europe is encouraged to join the CCPA[18] and also sign up for the working group mailing list[19]. More information is available via Europeana[20] and also in the last BHL-Europe newsletter[21].

The Europeana Data Provider Agreement
The MNHN filled the form of Europeana Consultation on New Data Provider and Aggregator Agreements and sent their comments about the main modifications of the document.

Organisational and administrative tasks
Organisational and administrative work has further been done during this reporting period. The coordination of WP2, WP2 project planning and the writing of reports occupied additional time during this working period.

WP3

3.1.3 - Definition of own standards for images, metadata, harvesting according to partner requirements (consensus -building)

Finalisation of the BHL-Europe metadata schema
In October NHMW participated in the metadata discussion around developing our own XML-schema for use within the project. During the meeting in Amsterdam (15/10/2010-16/10/2010) the TMB decided to put forward a draft for a possible XML-schema and continue the discussion afterwards. By the end of October NHMW started to create such a draft. In order to be able to support more than one incoming metadata format from the content providers, NHMW also continued enhancing the schema mapping tool. In November the XML-schema draft NHMW development was continued and a first version was finalised by 10 November, which was published to the Metadata and Tech group (both are Google Groups). Further enhancements were made based on feedback by other participants of the project. As part of the meeting in London (02/12/2010-03/12/2010) NHMW presented the improved XML-schema draft.
During the TMB meeting in December, it was decided to continue with the XML-schema and as a first basis create standard mappings for popular metadata formats, namely MARC21 and DC (Dublin Core). Before Christmas a first mapping of DC sample using the schema mapping tool was done. A sample based on data from MNHN was finished and sent out to the Work Package leaders. In order to further improve the XML-schema it was validated against a list of known use cases by NHMW together with Jana Hoffmann (MfN). The development team is currently in the process of finalising the work to ensure the schema is aligned with use cases and with the BHL-Europe system requirements.

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

BHL-Europe Portal Development Workgroup during Content Provider and Technical meeting, London

[18] http://version1.europeana.eu/web/guest/councilregistration
[19] http://version1.europeana.eu/web/europeana-foundation/workinggroup
[20] http://version1.europeana.eu/web/europeana-foundation/content-council
[21] http://www.bhl-europe.eu/newsletter.php?lang=en

Jana Hoffmann (MfN) organised two sessions during the BHL-Europe Content Provider Meeting in London 1 December 2010 on taxonomic intelligence and the BHL-Europe portal development.
'Taxonomic Intelligence tools for the BHL-Europe Portal': The first session was addressing the need for a taxonomic intelligence tool in the BHL-Europe portal. The consortium partner Sp2000 representing the Catalogue of Life[22] (CoL) was invited to share their experience and knowledge on taxonomic intelligence with the consortium and the technical partner in the BHL-Europe project. Prof F. Bisby (Executive Director of Species 2000) and Dr. Y. Roskov (Catalogue of Life Executive Editor, Species 2000) attended two days of the BHL-Europe meeting in NHM London on 1-2 December 2010. F. Bisby made a presentation "Names, taxa and synonymic indexing" on Taxonomic Intelligence session. Both F. Bisby and Y. Roskov participated in discussions on BHL-Europe Portal Development workgroup meeting and shared their ideas on potential use of the CoL as taxonomic backbone for structuring digitised publications. Latest versions of the CoL database and technical documentation have been delivered to WP3 for analysis of the potential use of the CoL index as taxonomic metadata in the BHL-Europe portal.
It became clear that a precise definition of what we mean by taxonomic intelligence is needed. During this session it was decided that we will use the uBio[23] TaxonFinder to identify potential taxonomic names (contact Chris Freeland, MOBOT). This will be the BHL-Europe automatic 'name finding' component. However, CoL will provide a more elaborated service in terms of 'real' taxonomic intelligence. Identified 'names' in the BHL-Europe content, either by using the Taxon Finder or provided by the content provider in the metadata, will be linked with the CoL and thus a systematic biological classification (hierarchy), synonyms and colloquial names will be accessible and will improve the recall of relevant literature. Technical implementations were already discussed during the meeting and are in progress. Other services like AnimalBase[24], providing species protologues, will be approached in a next step. Furthermore a service to improve the search for author's synonyms, Virtual International Authority File[25] (VIAF), was discussed and will be implemented in the BHL-Europe system. Meanwhile, BHL-Europe has all access data and permission to use the VIAF database. It is suggested to introduce a gateway to access different vocabularies after the inventory of all vocabularies relevant for BHL-Europe is finished.
'BHL-Europe Portal Development': During the second session on the portal development the catalogue of user requirements was briefly presented and discussed with the participants. Furthermore, relevant topics like OCR (tools, crowd sourcing), technical documentation and quality control (test cases), portal testing (usability tests, user survey) and web statistics (Google Analytics, NHM server log files, and user profiles) were discussed and next steps were planned.

EDIT Platform for Cybertaxonomy
Service methods and a web service have been implemented into the CDM library which allows using the OpenURL resolvers of BHL directly from within the Taxonomic Editor. Initially the Open URL service from BHL-US will be used until the according service in BHL-Europe becomes available. According user interface elements which would make this feature available are subject of further implementation tasks.

Create Pre-Ingest Tool Alpha 1
Pre-Ingest guidelines have been further refined and Pre-Ingest round table meeting was prepared to present the Pre-Ingest guidelines in London (December). The intended audience were mainly all content providers who need to see and understand the concepts of Pre-Ingest and Ingest so that upload and ingest of content and metadata can proceed quickly.

FTP upload site for Content Providers

---

[22] http://www.catalogueoflife.org/
[23] http://www.ubio.org/
[24] http://www.animalbase.uni-goettingen.de/
[25] http://viaf.org/

During the TMB meeting in Amsterdam the need for content providers to be able to upload already scanned materials was becoming very pressing. ATOS gathered functional and technical requirements and then discussed with the NHM the feasibility given their infrastructure and security constraints.

Once the functional and technical specifications were decided NHM installed the FTPS server and created the accounts.

Finally, the different functions on the test site have been tested and now all content providers are able to upload data to the NHM Servers. Additionally a User Guide for the content providers has been produced and distributed via email and on the [BHLWiki](https://bhl.wikispaces.com/BHL-Europe_Data_Upload)[26].

OCR

NHMW did some more OCR researching, especially in cooperation with BHL-US.

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

Installation and Configuration of Components on BHL Europe Servers at NHM

As part of the ongoing technical development work for the BHL-Europe system, ATOS have continued to progress with the work of installing major components on the BHL-Europe servers at NHM (including: Fedora Commons and Drupal for the portal). There were also many sub component and library dependencies that were installed such as Java, PHP, MySQL, etc.

3.4.2 - German Prototype

Preparation of the German Prototype (documentation including use cases)

The documentation of the German Prototype was created so that a first version could be discussed during the TMB Amsterdam meeting. AIT created a German Prototype System Description document and also a Use Case Document, which describes the various use cases for the BHL-Europe German Prototype front end & the administrative area. Both documents describe the used components and functionality in details. The documentation was further refined for the intermediary review in London in collaboration with the project coordinator (multiple iterations).

Mapping MODS/METS/ESE

Additional content provider's metadata has been mapped to ESE for presentation in the German prototype: UBBI, UBER, UGOE. The metadata was first mapped to METS/MODS and further mapped to ESE for ingest into Europeana. This step is required since the BHL-Europe schema also includes MODS for bibliographic metadata.

Release of the German Prototype (D3.6, M18 Deliverable)

Deliverable 3.6 Release of the German Prototype (M18) was submitted to the EC in November 2010. A demonstration and presentation was given at the 2nd Intermediate Review meeting by WP3 Leader/AIT and also by AIT during the BHL-Europe Content Provider & Technical meeting which took place at NHM, London (30 Nov - 3 Dec 2010).

The German Prototype was prepared for review according to the Amsterdam meeting. This includes proper documentation, bug fixes, user interface changes and inclusion of new metadata. Also a preliminary content viewer was integrated for demonstration purposes only.

The Prototype was created purely as a demonstrator to show that we can deliver the technology required for the final system. It also demonstrates that we can actually ingest and harvest data.

Finally, the German Prototype will be re-factorised and integrated into the community portal. Hence, future development will focus on other core features such as Taxonomic Intelligence Tools, OCR and Content Viewer.

---

[26] https://bhl.wikispaces.com/BHL-Europe_Data_Upload

3.4.3. - Revision of user requirements list for the Portal.

In this reporting period the description of the user requirements or 'catalogue of user requirements' was finalised (MfN, AIT). The user requirements were defined in more detail including a description, pre-conditions / requirements and a priority for development (based upon user survey, contribution from the consortium members, individual input). They were assigned to a feature category and the feature categories were grouped into four core features: search and retrieval, content viewer, indexing services and multilingual community portal. However, the individual features need further discussion and need to be wrapped up in a few use cases eventually. Furthermore, the user requirements were presented during the London meeting (which included the consortium members with taxonomic expertise) and based on the feedback provided additional requirements have been formalised and prioritised. The development team is currently in the process of translating these requirements into deliverables within the Scrum Backlog to ensure that the priority 1 features are realised for the M28 public core release and for the final BHL-Europe system/community portal.

WP4

4.1 – IPR Framework, and D 4.3 – IPR Agreements

The activities in WP4 continue to centre around the project's relationship with Europeana. The WP4 leader attended Europeana working groups in The Hague on data sharing, Linked Open Data, and the continued development of their Data Provider and Data Aggregator agreements. The workshop was used to discuss the perceived and actual risks and rewards of more generous terms of re-use of data than the existing non-commercial provisions of the data provider agreement. The results of this workshop are to be analysed alongside the possible rewards of freer re-use and form part of a second round of consultation on the Agreements with the providers and aggregators for the Danube release of Europeana in the first half of 2011.

The WP4 leader also attended the Europeana plenary[27] in Amsterdam where the value of data sharing and open access was discussed broadly. Among those speaking and sharing their experiences were the National Portrait Gallery London, Wikipedia, legal experts and Google Books. These discussions impact D4.1 agreements with BHL and Europeana for reciprocal access, Rights metadata and the sustainability issues to be considered for D4.3. Following the Europeana plenary some clarity was required from WP4 leader for BHL-Europe partners regarding the adoption of the Creative Common Public Domain Mark and the appropriate use of Europeana's bespoke 'Free Access' Mark. Clarity was provided to Naturalis and AIT in regard how this distinction affects licensing metadata.

The WP4 leader and RMCA have also reviewed and provided feedback to Europeana on the latest iteration of the Data Aggregator Agreement. It is not likely that this agreement will prove problematic for BHL-Europe to sign although further investigation into the possible confusion between Moral Rights and the desired acknowledgement of the content provider of digital works is necessary. Further feedback needs to be provided in due course which will inform the data sharing agreement between the two projects in the longer term.

During the December Project Review meeting in London the WP4 leader contributed to discussions and content provision for the WP2 Best Practice Guide regarding rights clearance, metadata standards for rights information, licensing and Open Access terms of participation in the project for prospective content providers.

Partners including, UGOE, MNHN, RMCA and NBGB have been working on their internal rights clearance process and prioritisation to ensure the maximum amount of content can be provided to BHL-Europe in the shortest time.

WP5

---

[27] http://version1.europeana.eu/web/europeana-plenary-2010/

5.1 - CWG meetings

During the Content Provider and Technical meeting in London in December, the [CWG](#)[28] had a separate dissemination workshop. Several points were discussed:

- The transfer of the website to a CMS, with a database of events and a general update of the website, for example: new images on the homepage.
- We also discussed the possibilities of new promotion material such as fact sheets, new flyers, new demonstrator, and others
- To improve the internal communication the project coordinator will organise a general conference call open for all partners, once a month. This will give the possibility for the partners to be informed about the latest technical work, and to have a regular direct contact with the work package leaders.
- In the future we will share the events and conferences were BHL-Europe will be presented with BHL-US via a general agenda. Contacts were made.
- We went through the WP5 work plan and made some adaptations.
- A first work plan was created for the virtual exhibitions, several ideas were mentioned, and Europeana will help with guidelines and technical possibilities.

This workshop was followed by a small CWG meeting (only RBINS and NMP) in Prague at the end of December 2010. Here we discussed the news-workflow between events, the blog, Facebook, Twitter and the newsletter, a little further about the virtual exhibitions, and we finalised the WP5 work plan.

5.2.1 – BHL-Europe Website - http://www.bhl-europe.eu/

The project website has been updated. New news and events were added and the website "Home page" and "About us page" are now available in six different languages (English, German, Italian, Spanish, Slovakian and Czech). A French translation will be added in the next weeks. Several partners were contacted to help with these translations.

The successful setup of Drupal for our website will facilitate the regular update of the site by the CWG and the management of the different languages.

5.3.1 – Develop, prepare and update promotion kits and related printed matters

Conference bags for Biosystematics Berlin 2011

Pavel Šťastný (NMP) produced a draft of BHL-Europe conference bags, which were prepared for the Biosystematics conference organised in Berlin in February 2011. The type of the bag and logo design was discussed and agreed with MfN. The production of 600 bags is in process. The conference bags contain all relevant information (conference proceedings, promotion material (also BHL-Europe flyers), city/ travel information) for the conference. They will be delivered to MfN and packed by volunteers.

Folders

Pavel Šťastný (NMP) produced a draft and the final design of the BHL-Europe folder and provided also calculation for its printing costs. It was not yet decided how many folders will be prepared due to shortage of finances and other high priority tasks.

Christmas Card

The production of a BHL-Europe e-Christmas card was suggested by the WP2 leader and a layout was decided in cooperation with the NMP artists. This was sent to the consortium and external contacts together with the newsletter at the end of December 2010.

---

[28] https://bhl.wikispaces.com/file/view/CWG_Minutes_101201.pdf

Virtual exhibition

First ideas and concepts for our virtual exhibition were discussed during the conference call on 4 January 2011 with partners from HNM (Graham Higley) and the European Library (Aubéry Escande). The European Library allows us to use their virtual exhibition (Reading Europe) as a model. It was realised that we need more than Reading Europe provides. We decided to contact a company IT4CARE for construction of the engine for the virtual exhibition.

Basal structure of the virtual exhibition was prepared and agreed that the data of the virtual exhibition webpage will be in three layers: home page, topic page and book annotation card.

Two topics - spices and expeditions - were selected for the pilot project. We also decided to expand the virtual exhibition in the way of informative questions "Did you know that…?" and several other new items: photo of each spice, short story, one page comics in some cases etc, new search options via names of plants and spices. RMCA contributed to the virtual exhibition in working on Charles Lemaire expeditions in Congo (1901-1903), selecting and digitising of related material. This material will be used as a basis for the virtual exhibition on expeditions.

Logo branding

Pavel Šťastný produced first draft of 13 logos for BHL-Europe branding. These were discussed within the communication and management group during conference calls.

5.3.2 – BHL-Europe Newsletter

Newsletter 5 and 6[29] were combined and were sent to the consortium and external contacts at the end of December 2010. The work on the newsletter 5 took considerably amount of time in this reporting period. In July two articles were received about vifabio, one was not acceptable for publication in the format presented in summer 2010. After the delay with the newsletter publication, Henning Scholz revised the article in collaboration with the author to incorporate all information needed for the newsletter. A final version was agreed that was acceptable for the newsletter. Furthermore, Henning Scholz provided the project news for the newsletter and revised the newsletter several times to bring it in its final format. An additional review by Jana Hoffmann (layout check) resulted in minor changes of the newsletter layout.

---

[29] http://www.bhl-europe.eu/newsletter.php?lang=en

The work on Newsletter No. 7 was started and partners were asked for contributions. A paper about the new project OpenUp will be prepared for the next newsletter.

5.3.3 - Facebook

The BHL-Europe Facebook[30] page has been created halfway in October, and at the end of December, we had 42 people who liked the page, and this number is steadily growing. It is clear in the insights of the page that after sending around the newsletter and the announcement that BHL-Europe is on Facebook, more people became a 'member' (after Christmas and right after New Year) so normally it will be more popular if we use it often and together with the other social networks.

5.4.5 - Dissemination activities of partners

Several partners participated in various institutional or project meetings and conferences to represent BHL-Europe. All information collected with the dissemination questionnaire can be found on the BHLWiki[31].

BHL-Europe is organising a symposium and a workshop during the BioSystematics Berlin 2011. The symposium 'World wide Access to Biodiversity Literature: Digital and Virtual Libraries for Taxonomists.' is organised and will be chaired by Boris Jacob (MfN) and Jana Hoffmann (MfN). The workshop 'Biodiversity Heritage Library: Hands-on training on search and retrieval of digital literature' is organised by Henning Scholz (MfN). Jana Hoffmann is part of the organising committee for this conference and represents BHL-Europe in regular committee meetings either at MfN or BGBM.

RMCA participated in local and international meetings, events and discussion groups in order to elucidate and promote BHL-Europe and BHL Global (see also WP 2&3 above and dissemination table of events). They have decided to publish a short overview in the next issue of JAZ (Journal of Afrotropical Zoology) about the project, an article is in preparation.

RMCA and NBGB were present at the 3rd publishing meeting EDIT (Copenhagen, 06-07.10.2010), to present a speech on the cooperation between the Belgian partners, NBGB, RBINS and RMCA, for sharing human and material resources for the BHL-Europe project.

NBGB participated at the Linnaeus Link Union Catalogue (LLUC) partners meeting in Geneva (01-02.10.2010) and presented the BHL-Europe project to the attendees. One of the aims of the LLUC is to provide links to Linnaean digitised books, namely on BHL.

AIT gave a presentation about the BHL-Europe project at the Austrian EuropeanaLocal National event on the 19th of November at the University of Graz. This presentation was two-fold: preparation and compilation of promotion material in German for distribution and a short PowerPoint presentation given during a speech.

CSIC was one of the partners who helped with the translation of the project website and presented BHL-Europe to students of a Master in Biodiversity, Universidad Autónoma de Madrid.

---

[30] http://www.facebook.com/pages/BHL-Europe/151086001600041
[31] https://bhl.wikispaces.com/dissemination+of+project+by+partners

UH-Viikki wrote an article for the journal of the Finnish Biological Society Vanamo, "Luonnon tutkija" (hopefully to be published in the last issue of 2010).

RBGE made sure to advertise BHL-Europe and BHL generally at the Annual Meeting of the Scottish section of the Botanical Society of the British Isles, held on 6th November 2010 via a poster.

UBER sent a submission of the abstract for the event: "Vernetzungstage 2011" in Osnabrück.

BHL-US – Extract from the Program Director's Report

Chris Freeland and Mike Lichtenberg attended the Darwin's Library Project Meeting in October in London at the National History Museum.

In October, while in London, Chris Freeland had a meeting in Oxford with Ed Pentz and Geoffrey Bilder from CrossRef to discuss assigning DOIs to BHL content. The outcomes of this meeting helped out clarifying some doubts about DOIs and sparked a fruitful discussion through the BHL-Europe list and call to help decide later on how to move forward assigning GUIDs for the project elements.

A test set of BHL texts, including examples of main issues encountered, was developed for evaluation in EU-funded IMPACT project, focusing on OCR improvement evaluation and comparison from different initiatives.

Chris Freeland, Martin Kalfatovic and Keri Thomson from BHL visited the Chinese Academy of Science in November to share their experiences, review issues and comment on optimising the scanning process.

BHL soft launches Citebank v.1.0[32], an open access platform to aggregate citations for biodiversity publications and deliver access to biodiversity related articles, using Drupal and Fedora.

4   Deliverables

No deliverables were scheduled for M21 of the project.

5   Schedule status/problems/issues/deviations

Naturalis
Kees Hendriks from NCB Naturalis (Leiden) has left his position as head of the department Library and Information Service of NCB Naturalis. A temporary replacement has been found, but he is not available for any work in BHL Europe. It is expected that a new head will be in function from July 1, 2011. This should only have a minor impact on the project (WP1 and WP5). In the meantime Tom Gilissen will be the main contact for Naturalis.

Reporting
As mentioned above, the reporting still must be improved to be effective for all partners. Every status report must be narrative, in indirect speech, and on task level (including task number according to DoW). This way it is easier to compile a high quality progress report.

Project Server
Unfortunately, the user accounts for the Citrix-based Project Server access are still not set up by NHM, so that the WP leaders outside the NHM are still unable to upload and modify their reports using the project server. There is ongoing communication and hopefully this issue will be solved in the next reporting period.

---

[32] http://citebank.org/

Portal development and ingest workflow

Over the last weeks we received some external feedback concerning the German Prototype. This feedback will be taken into account in the ongoing planning of the BHL-Europe portal to develop a user friendly portal that is actually addressing user needs. In addition to our own perception, this feedback also showed that some key components (as described above) of the future BHL-Europe portal were not addressed and discussed satisfactorily in the portal development in WP3, e.g. taxonomic intelligence, OCR, portal testing. Therefore, a Workgroup on portal development was organised during the London meeting in December. In this group the portal development progress itself, user requirements, taxonomic intelligence, OCR and future tasks were discussed. As an extension of the London meeting the group will have regular calls with the goal to discuss and bring forward the portal development. This includes that technical issues and solutions will be discussed and user requirements will be prioritised. It is a major goal for the next reporting period to create a detailed plan for the portal development, assign responsibilities, identifying adequate technical solutions and implementing high priority features of the BHL-Europe portal prototype.

During this reporting period we have also seen the requirement to manage the existing dependencies of WP2 and WP3 effectively. The Ingest and data workflow needs to be defined and communicated with the partners. We need clear roles and responsibilities within WP3 (data mapping, Pre-Ingest, QA) to give our content providers a clear direction. This requires very effective, open and regular communication between the parties involved and a high quality project management. This is necessary to not delay our deliverables for M24 and to not delay getting content into Europeana as well as into our portal. Regular and extensive conference calls are scheduled to solve the existing issues and internal documents are in progress to define the details of the work to be carried out.

6   Work planned

The work planned for the next phase is explained in more detail in Annex 1 (DoW).

We open now one PMG call every month for the consortium to allow all partners to call in, get the most recent update and discuss urgent questions with the relevant WP leader.

Further to discussions held during the Global BHL Technical meeting in September, several members of the technical group will be attending the BHL-US/EU & Bibliotheca Alexandrina meeting (24-28 January 2011) to discuss such topics as Data Sharing, OCR Processes, BHL Services & Interfaces etc. During the visit the technical group will carry out further research & investigation required for the GUID mint solution. The requirements that BHL-Europe has identified and which need to be developed are very similar to those already deployed at the Bibliotheca Alexandrina. The BHL-Europe Architecture & Development schedule will also be discussed during the visit on the 28th January, 2011.

The BHL-Europe Annual Review Meeting will take place 6-10 June 2011 in Belgium (Tervuren, Meise and Brussels) and will be organised by RMCA, RBINS and NBGB. Further information will be available and updated on the BHL Wiki[33].

7   Products to be completed in the next phase

24	D1.5 First overview of the business plan for long-term sustainability – M24
25	D1.6 Progress report 4 including pre-financing request
26	D1.7 Annual report

---

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

27	D2.5 Final and enhanced Web-database for content management and collection analysis
28	D2.6 Delivery of the first version of the approved best practice guidelines and standards
29	D2.7 Content analysis and management status report 2 …
30	D3.7 Key components documented for output of D3.5 e.g. BHL-Europe Portal, OCR demonstrators, distributed storage model, etc.
31	D4.2 Complete signed agreements with first group of rights holders to enable material to be used in the BHL Europeana, and established process for addition of further material.
32	D4.3 Deliver IPR framework to support long-term access and sustainability of the digitise material.