

Report on Deliverable 3.7

Project:	Synthesis of systematic resources
Project acronym:	SYNTHESYS3
Grant Agreement number:	312253
Workpackage:	3: Sustainability
Deliverable number:	3.7
Deliverable title:	SPNHC Symposium: International symposium to promote SYNTHESYS activities
Deliverable author(s):	Elspeth Haston Saskia Jancke Dimitris Koureas Laurence Livermore Vince Smith Christiane Quaiser Thomas von Rintelen
Date:	21.10.2016

SPNHC Symposium: International symposium to promote SYNTHESYS activities

Enabling Infrastructure: Future Collections, Data & Informatics

The 31st Annual Meeting of the SPNHC (Society for the Preservation of Natural History Collections) took place in Berlin from the 19-26 June 2016. SYNTHESYS together with iDigBio, GBIF and CETAF - the largest biodiversity networks and projects - collaborated to organise three days of talks that focused on coordinating efforts to realise global digitisation and global data access for biological collections. 300 participants originating from 31 countries attended the SPNHC meeting. In parallel to the SPNHC the Global Genome Biodiversity Network (GGBN) meeting took place with another 100 attendees who were also potential audience for the SPNHC meeting.

The collaboration between SYNTHESYS, iDigBio, GBIF and CETAF resulted in three integrated symposia at the SPNHC conference, brought together under the title "One Collection: pathways to integration". The overall goals of the integrated symposia were to discuss:

- Lessons learned so far, including new developments and data use examples;
- Strategies for worldwide access to the means to digitize collections data;
- Guidelines to aid prioritization of collections' digitization and future collecting; and
- The necessary human resources, hardware and software infrastructure for global creation of and access to this much-needed data.

Symposium 1: An international conversation on mobilising natural history collections data and integrating data for research / primary organiser: Digitized Biocollections (iDigBio), SYNTHESYS3, University of Oregon.

Symposium 2: Enabling infrastructure: Future collections, data & informatics / primary organiser: SYNTHESYS3.

Symposium 3: Developing a global research infrastructure framework for bio-collections / primary organiser: SYNTHESYS3.

The first symposium was chaired by Deborah Paul (iDigBio), Elspeth Haston (SYNTHESYS3) and Brian Westra (University of Oregon, USA). Presentations were invited from the collections and biodiversity informatics community involved in the mobilisation and use of NHC data within the following areas:

- Digitisation (including imaging) lessons learned and outstanding questions,
- New and emerging technologies/models, and cross-discipline collaborations for digitisation and georeferencing, and
- Collections data-in-action.

The number of talks submitted resulted in this symposium being split into two sessions, demonstrating the relevance of the topic to collections around the world. The first session started with digitisation lessons learned and outstanding questions with talks on practical lessons in digitisation (Sarah Philips) and evaluation of collections management systems (Erica Kimmel). The next talks looked at new and emerging technologies/models and cross discipline collaborations with talks on alternative approaches to digitising palaeontological collections (Vladimir Blagoderov) and georeferencing collection records (Gabriela Hogue). Finally a session on collections data-in-action included talks on palaeontological collections data within research data systems (David Lazarus),

digital infrastructure and tools for collection discovery and use (Chris Norris), digitisation to help with identification (Paul Mayer), documenting the flora of Melanesia (Shelley James). A final talk on data librarianship and small collections support (Brian Westra) and an opportunity for some discussion concluded this first session. This session was filmed and the talks are being made available by iDigBio.

The second session comprised 18 talks on a range of digitisation projects. These talks provided practical examples of digitisation methods and tools, showing the progress being made across the NH community. The breadth of the talks, and the range of institutes presenting, effectively showed the amount of work and development being carried out.

Laurence Livermore and Vince Smith (NHM) chaired the SYNTHESYS Symposium 2 talks on technical and organisational infrastructure around data mobilisation, data aggregators and data portals. In short these are the tools, systems and platforms that will support the next generation of natural history collections digitisation. In total there were 14 talks representing projects and organisations from six countries.

The initial talks covered large-scale projects such as national NH collections digitisation initiatives (iDigBio), the latest in crowdsourcing ([Notes from Nature2](#),) and progress in digitising legacy biodiversity literature ([BHL](#)), e.g:

- Notes from Nature 2.0: Citizen science at scale - Robert Guralnick*, Michael Denslow & Austin Mast

The main body of the session covered a variety of topics including new research on metadata extraction using OCR and computer vision (SYNTHESYS JRA); how much our natural history data is being used worldwide; new institutional data portals and ways of exploiting linked open data; alongside work to refine and simplify the policies and processes necessary to support mass digitisation projects.

The session concluded with talks from on high-throughput industrial digitisation services (Witikon and Digitalium) and a view on industry ([Cisco](#)) on the benefits of collaborating with natural history collections, and how the community can better work with innovators, start-ups and SMEs, e.g.:

- Automating Insect Digitisation - Speed and Costs - Riitta Tegelberg, Janne Karppinen, Zhengzhe Wu, Jere Kahanpää & Hannu Saarenmaa*;
- WITIKON: Mass 3D digitisation at a national scale - Mira Silanova;
- A Bridge from Enabling Infrastructure to Digitisation Priorities, a view from Industry - Deb Paul*, Philip van Heerden, Vince Smith, Laurence Livermore & Ehsan Alavi Fazel.

Finally, a closing talk was given by NHM London's Director of Science, Ian Owens, on "The new enlightenment: digital collections and the re-invention of large natural history collections".

Dimitris Koureas (SYNTHESYS3) and Ana Casino (CETAF) chaired the third SYNTHESYS Symposium and started off with an introduction to the topic of developing a global research infrastructure framework for bio-collections. The future of biological collections lies in greater integration and cooperation to develop as a global Research Infrastructure (RI) with common practices, policies and systems. This session brought together key stakeholders to coordinate actions relevant to the development of global RI on bio-collections.

Key topics of the session were:

- Established distributed Research Infrastructures: Lessons learnt;
- Bio-collections in Europe: ESFRI roadmap update preparatory activities;
- Links to related regional initiatives (e.g. iDigBio, ENVRI+, LifeWatch, ELIXIR);
- International collaboration between major programmes;
- Sustainability and business models: What is available and what is possible?
- Training and capacity building programmes;
- Wrap-up discussion: “How do we proceed?” – Action items.

The symposium contributed to a globally coordinated discussion on how national and regional initiatives on digitisation and virtual access can be aligned and linked to develop a global research infrastructure. Talks covered subjects on building a pan European research infrastructure (DISSCO - Distributed System of Scientific Collections), an update on activities of the Biodiversity Collections Network (BCoN) and Research Coordination Network (RCN), the need of sharing resources and expertise at a global scale, integrating e-infrastructures to support environmental research and data exchange via ELIXIR. The symposium ended in a panel and open discussion.

Furthermore, one workshop (Workshop 7) was jointly hosted by SYNTHESYS3 and iDigBio on: Digitisation Software Training Workshop: Inselect, Symbiotica & ABBYY FineReader.

Aim of the workshop was to make these tools more accessible to institutions of all sizes. It was planned as a mix of informative presentations, practical training and open discussions.

In addition to the symposia, there was a promotional SYNTHESYS3 information desk running throughout the conference distributing flyers to promote the project’s activities.

At the GGBN a further symposium was organised by SYNTHESYS3 on “Policies for new physical collections” chaired by Ole Seberg and Thomas von Rintelen with talks on applying best practices to the operation of NH collections and on enhancing the accessibility and conservation of plant tissue samples stored in silica gel. The talks were followed by a general discussion on DNA and tissue collection policy development for GGBN and their application to SYNTHESYS partners joining GGBN. The policies were deemed adequate and ready for publication, they will be made available in the GGBN document library, which is under construction later in 2016 (the policies will also be made available on the SYNTHESYS website, deliverable 2.5).

All these sessions highlighted the relevance of SYNTHESYS results for forming an action plan for the preservation of European and global NH heritage in collections and for identifying future research priorities in this field. The symposia contributed to the following objectives:

NA2, objective 1 – Managing new (virtual and physical) collections;

NA2, objective 2 – Developing strategic priorities for molecular related NH collections;

NA3, objective 2 – Promotion and outreach;

NA3, objective 3 – Sustainability.

Presentations at both conferences were published through short publications in the conference books. They have been provided as hardcopy during the conference and are online available for download [http://www.spnhc2016.berlin/downloads/2016/SPNHC_broschuere_digital.pdf].

Furthermore, the presentations of the symposia were recorded and are online freely available for a wider audience [https://www.idigbio.org/wiki/index.php/SPNHC_2016_Recordings].

The event was covered in a blog post [<https://blog.nhm.ac.uk/2016/06/29/digitisation-futures-in-berlin-digital-collections/>].