

**WORKING GROUPS REPORT**

**Sofia, BU, 12-13 March 2019**

**WG1/3 WORKSHOP: Data management strategies for scientific collections**

1. Attendance and planning:
   1. WG Leader and Co-Leader

WG1: Patricia Mergen, Quentin Groom, Alexander Weigand

WG3: Eva Häffner, Wouter Addink, Ole Seberg

* 1. Conveners: Wouter Addink and Eva Häffner,

Co-Conveners (unable to attend): Patricia Mergen, Quentin Groom, Alexander Weigand and Ole Seberg

Rapporteur: Wouter Addink

* 1. Number of Participants: 26

Participants: Wouter Addink, Eleftheria Andaloudaki, Torkild Bakken, Olaf Bánki, Dimitri Brosens, Markus Döring, Lorenzo Cecchi, Holger Frick, Valeri Georgiev, Borislav Georguiev, Anton Güntsch, Eva Häffner, Aino Juslén, Danae Karakasi, Colin Kelleher, Dimitris Koureas, Kari Lintulaakso, Kristian Meissner, Aleksandra Nina, Anne Nivart, Lars Petterson, Nikolaus Szucsich, Apostolos Trichas, Claus Weiland, Kaloust Paragamian, Matt Woodburn.

* 1. Inclusiveness rate of the WG participation:
* Gender balance: male: 20, female: 6 (23% female)
* Geographical distribution: Austria: 1 Belgium: 1, Bulgaria: 2 Denmark: 1, Finland: 3, France: 1, Germany: 3 Greece: 4, Italy:1, Ireland: 1, Netherlands: 3, Norway: 1, Serbia: 1, Sweden: 1, Switzerland: 1, UK: 1
* Seniority: 5 Early Career professionals (19%)

1. Objectives of the WG workshop:

* Establish the network of experts across the community to deal with the different data management tasks arising within DiSSCo
* Complement the landscape analysis of current and planned data management activities
* Identify priorities, structure the work program and plan future activities regarding data management in MOBILISE.

1. Topics discussed
2. In reference to completion of Y1 objectives:

**GAPG 1 Synchronise efforts among Working Groups & wider community:**

The workshop was designed to include participants and topics from the scientific collections community as well as from outside the community, with strong links to DiSSCo (Distributed System of Scientific Collections). The wider biosciences community targeted included participants in the European life-sciences infrastructure for biological Information (ELIXIR), the COST Action “Developing new genetic tools for bioassessment of aquatic ecosystems in Europe” (DNAqua-NET), and cultural heritage projects ,e. g. the Digital Research Infrastructure for the Arts and Humanities (DARIAH), “Linking Biodiversity and Culture Information” (LINBI).

For the synchronisation of efforts especially within the collections community and MOBILISE, the workshop built on the results of the MOBILISE Align meeting that was held on November, 16th 2018 in Bratislava. That meeting resulted in the formation of 5 thematic synchronisation groups, see also <http://bit.ly/sync_groups>. The workshop focused on the thematic synchronisation groups 1 (international coordination and policy harmonisation), 2 (standards and other common resources) and 3 (data models, management, publishing pipelines & services). The respective experts have been invited and the breakout groups focused largely on analysing these fields further in terms of gaps and challenges and concrete actions.

**GAPG 6 Identify and activate experts, from inside and outside the collection community**

See also GAPG1 regarding inclusion of respective experts. Expertise and interests/ideas of experts have been identified in an interactive workshop session and have been documented (see Annex). Experts have been invited to name further experts who should be invited to contribute towards the goals of WG 1 and 3 and MOBILISE in general.

**GAPG 8 Data models, management, publishing pipelines & services**

In the breakout sessions of the workshop, two groups identified challenges and possible activities and solutions to problems regarding standards, data models and data management, see also the extended report in the Annex.

1. Towards the definition of Y2 activities and actions

Experts have identified action points and fields of activity that contribute to further analysing the landscape of standards and other resources and to address gaps and challenges. For year 2, it was agreed that WG1 and WG3 activities should contribute to the completion of the landscape analysis of standards and other common resources in 1-2 workshops. WG3 has identified digital specimen specification (Open DS specification) and linked open data as priority topics for year 2 and intends to support respective workshops. All groups have furthermore been invited to discuss possible contributions to the biodiversity\_next symposium ([https://biodiversitynext.org](https://biodiversitynext.org/)), especially in the symposia with reference to MOBILISE. An informal meeting of the Biodiversity Services and Clients Interest Group of TDWG will be organised during biodiversity\_next by Anton Güntsch to align Mobilise/DiSSCo activities with the Biodiversity Services and Clients (BSC) Interest Group.

1. Others

The workshop also addressed work and activities beyond year 2. The task was to structure the broad field of data management in general. Especially, topics related to policy issues will have to be addressed in the further course of the Action. Policy issues have to be reflected in all steps of data management. A policy graph has to be established for DiSSCo and will also be part of DiSSCo Prepare (i. e. lining policy components to activities and services, discussed in breakout group 1).

1. Major outcomes (also see Annex)

* A list of experts with specification of their expertise and interests
* An analysis of gaps and challenges in the landscape of standards and services (to be continued)
* A list of concrete action points and activities for the coming years of MOBILISE
* Ideas for concrete products to be generated out of MOBILISE and related activities within and beyond DiSSCo

1. Challenges faced for Y2

To use funds as efficiently as possible to achieve impact and to be excellent and inclusive at the same time. There are many established experts in the field who are indispensable for developing the concepts needed for DiSSCo, but it is equally important to include young scientists and people from as many collections as possible to keep the flow of ideas and expertise alive within the wider community (and in all directions). It is therefore advisable to concentrate activities.

1. Next steps

Organise community (*via* Open Science Framework, <https://osf.io/j6psx/>)

Determine Y2 program

Work on standards landscape analysis (deliverable), use FAIRsharing, <https://fairsharing.org> to compile the landscape analysis from a services perspective.

Discuss and build community consensus in a few concrete topics (e.g. linking biodiversity data openly, Open DS specification, Access and Benefit Sharing).

ANNEXES:

1. List of Participants: Participants name, affiliation, expertise, contact
2. Extended Report
3. Other material