

Workshop in Sofia
(March 2019):

“Data Storage and Archiving Strategies - Towards a Guideline for Long term storage and archiving of digital objects from the collection data domain (i.e. digital specimens and digital collections)”

Organizers: Dagmar Triebel and
Gila Kahila Bar-Gal

No. of participants: 23

Workshop in Biodiversity Next, SI55
(October 2019):

“Federated Infrastructures for Sustainable Biodiversity Data Management”

Organizer: Birgitta König-Ries
Co- organizer: Dagmar Triebel
(among others)

No. of participants: 45-50

“Data Storage and Archiving Strategies”

- Important topic with low profile/awareness
- Limited number of experts in the field of long term storage and archiving of digital objects from scientific collection(s)
- Involvement of experts from different disciplines
- Concepts and applied standards of data archiving in relation to digitisation efforts at natural history collections - first consultation with experts

- Definition of “data archiving” and related terms
- Stressing the difference between “data archiving” and data backup.
- Establishment of a first draft list of terms to be used by the community
- Identifying existing communities and expertise – organization of a section at the Biodiversity_Next meeting.

“Federated Infrastructures for Sustainable Biodiversity Data Management”

- Setting Up a Trusted Network for German Botanic Gardens Using Open Source Technologies
- Use of European Open Science Cloud and National e-Infrastructures for the Long-Term Storage of Digitised Assets from Natural History Collections *
- NFDI4BioDiversity: Biodiversity, ecology and environmental data
- The Freshwater Information Platform: An online network supporting freshwater biodiversity research and data publishing *
- Long-Term Reusability of Biodiversity and Collection Data using a National Federated Data Infrastructure
- SEINet: A Centralized Specimen Resource Managed by a Distributed Network of Researchers #

* European project involves several countries

American collaborative project

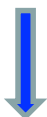
- Increase number of collaborative digitized projects worldwide.
- Projects have an impact in various scientific fields of taxonomy, ecology, evolution, biodiversity, environment, conservation etc.
- Projects needs to expand/engage to other communities by including farm animal, crop sciences, citizen science etc. establishing science-society-policy interface

- Focus on data collection – developing various platforms/
software's **less awareness of archiving the data!**
- The challenge for making archived data reusable is the **integration of heterogeneity data**: often implicit but differing semantics.

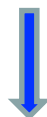
- Rapid increase of digitised assets (6,000 herbarium specimens per day)
- Lacking long-term storage support
- National and European e- infrastructures can provide an alternative solution by supporting different parts of the digitisation workflows.



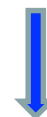
e-infrastructures providing long-term storage - pilot studies:
EUDAT- CINES, Zenodo, and National Infrastructures.



Herbarium - National
 Museum of Natural
 History France



Herbarium -
 Meise Botanic
 Garden (APM)



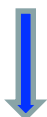
Herbarium & entomology -
 Finnish Museum of Natural
 History (LUOMUS) into the
 Finnish Biodiversity Information
 Facility (FinBIF 2019)

EUDAT- CINES

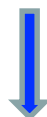
Zenodo

National Infrastructures.

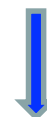
Source:



Herbarium - National
Museum of Natural
History France

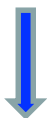


Herbarium -
Meise Botanic
Garden (APM)

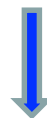


Herbarium & entomology - Finnish
Museum of Natural History (LUOMUS)

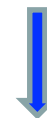
Composed of :



Specimen data & its
business metadata

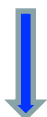


Specimen data &
metadata supporting
indexing & access to
resources



Link of specimen data with other
information

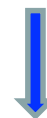
Use of:



ePIC identifiers



**DataCite Digital
Object Identifiers
(DOI)**

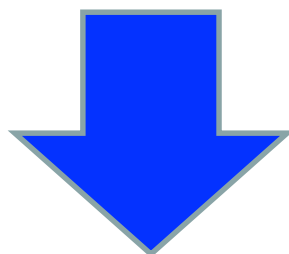


HTTP URI based identifiers

EUDAT- CINES

Zenodo

National Infrastructures.



The three specific contexts are relevant for implementation for long-term storage and archiving digitised specimen data