

FAIR Digital Object and Natural Science Collections data



DiSSCo adopts the RDA Recommendations on Virtual Layers, Basic Vocabularies and Metadata Standards for Attribution of Physical and Digital Collections Stewardship.



Distributed System of Scientific Collections

The Distributed System of Scientific Collections (DiSSCo) is a new world-class Research Infrastructure (RI) for natural science collections - collections hosted in Natural History Museums, Botanic Gardens, Universities and other Research Centres.

The challenge

Approximately 3 billion specimens are stored by the world's natural science collections. These objects have lifespans of decades and often, several hundreds of years and provide essential resources for documenting and understanding our natural world. They are increasingly used by researchers from a variety of disciplines. Despite increasing digitization and accelerated pace of data collection, access to specimen data remains inefficient and inadequate to address the complex research questions. The data derived from and linked to these specimens need to be easily findable, accessible, adhere to open standards, so they are interoperable and marked up with enough additional information to make them widely reusable.

The RDA outputs adopted

One of the key elements of DiSSCo's data infrastructure design is the 'Digital Specimen' - a FAIR Digital Object acting as a digital representation on the Internet for a specific physical specimen in a collection. The ideas for the Digital Object as an element in a data infrastructure was proposed in the output on **virtual layer recommendations** by the **Data Fabric IG** and in the **basic vocabulary** from the **Data Foundation and Terminology WG**.

Through adopting FAIR Digital Objects for Digital Specimens the digital representations of physical specimens can be treated as atomic items that need individual identification to avoid ambiguity and to collect and anchor core information about the specimens in one place. A FAIR Digital Objects architecture also enables to attribute work and track provenance in the curation and maintenance of the individual objects. We have chosen to adopt the output from the joint RDA/TWDG metadata attribution WG because recording actions of human and machine agents during data curation and processing phases is essential to FAIR implementation. The joint RDA/TDWG Attribution Metadata Working Group: Final Recommendations, which uses W3C's PROV data model, makes it easy to implement this in the FAIR Digital Object context as it allows capturing attributions in a standard way to be part of the Digital Specimen data when different operations are performed in multiple contexts.

The impact of the adoption

When digitising natural science collections, a clear link must be made between the physical specimen itself and the information digitally representing the specimen. This representation needs to be uniquely identifiable and findable and also have some descriptive metadata to assist the discovery and verification of the object. DiSSCo expects that adopting the Digital Object Architecture recommended by RDA, and treating Digital Specimens as the core

Find out more at:

www.rd-alliance.org/recommendations-outputs

Visit

rd-alliance.org

or write us at

enquiries@rd-alliance.org

element in the architecture will lead to transformations in working practices of collections-based science. One of these transformations is the implementation of improved attribution and provenance tracking in all parts of the data management cycle for specimens, enabled by the adoption of the metadata standards for attribution of physical and digital collections stewardship recommendation.

The Digital Specimens will act as a digital gateway and allow the researchers to do more than just find specimens. A wide variety of novel services also become possible including for example: harmonizing the arrangement of loans and visits, finding specimens related to one another, linking to third-party information, providing support to Access and Benefit Sharing (ABS), automated approaches involving machine learning and computer vision. The adopted metadata recommendation will also add to this, for example by services that visualise attribution for time spent on data management by researchers.

The adoption process

Taking the RDA's recommendation, the ideas were further explored in the work of the C2CAMP initiative and the RDA-Europe Group of European Data Experts and the GOFAIR initiative. These ideas were then materialized first

in the provisional Data Management Plan for DiSSCo and later in the Conceptual design blueprint for the DiSSCo digitization infrastructure - both are deliverables of the EU-funded ICEDIG project.

Lessons learned

RDA recommendations provided valuable inputs for DiSSCo research infrastructure design. As some of these ideas are still in the conceptual phase, avenues for creating sandbox and prototype environments, global technical hackathons and workshops will help us further clarify and test the concepts, identify potential risks and benefits in the FAIR Digital Object approach.

About DISSCO

The Distributed System of Scientific Collections (DiSSCo) is a new world-class Research Infrastructure (RI) for natural science collections - collections hosted in Natural History Museums, Botanic Gardens, Universities and other Research Centres. DiSSCo is currently in the preparatory phase and working to reach the maturity required to be able to embark on a full-scale implementation phase beginning in 2024. DiSSCo's mission is to transform practices and policies across the institutions that are hosting natural science collections in Europe along with supporting industrial-scale digitisation, community curation and FAIR data.

The logo for DiSSCO, consisting of the letters 'DiSSCO' in a bold, blue, sans-serif font. The 'i' and 'S's are lowercase, while 'D', 'S', 'C', and 'O' are uppercase.

Distributed System of Scientific Collections

Authors:

Sharif Islam (Naturalis)

Wouter Addink,

RDA Europe ambassador for the natural sciences (Naturalis)



Find out more at:

www.rd-alliance.org/recommendations-outputs

Visit

rd-alliance.org

or write us at

enquiries@rd-alliance.org