

Open Science and Collections in Austria
Workshop at the Natural History Museum Vienna
Parts I and II
16.2.2021, 9:00–12:40 Uhr

Speakers (in chronological order):

Dr. Katrin Vohland (Director General NHM Vienna)
Dr. Kostas Glinos (Head of Open Science Unit RTD.G4, Directorate-General for Research and Innovation European Commission)
Dr. Stefan Hanslik (Austrian Federal Ministry of Education, Science and Research)
Prof. Dr. Andreas Rauber (TU Vienna)
Mag. Heimo Rainer (NHM Vienna)
Dr. Olaf Bánki (Naturalis), Dr. Chantal Huijbers (Naturalis) & Markus Döring (GBIF)
Dr. Susanne Blumesberger (Vienna University Library)
Dr. Stefan Eichert (NHM Vienna)
Mag. Iris Ott (NHM Vienna)

Participants (in alphabetical order):

NHM Wien: Nesrine Akkari, Walpurga Antl, Anna Berger, Christian Bräuchler, Sabine Eggers, Anita Eschner, Ludovic Ferriere, Pedro Frade, Sabine Gaal, Karina Grömer, Vera Hammer, Elisabeth Haring, Christoph Hörweg, 2. Zoo Käfer (KAEF), Uwe Kolitsch, Martin Krenn, Andreas Kroh, Luise Kruckenhauser, Alexander Lag, Elke Lhotak, Alexander Lukeneder, Oliver Macek, Oleg Mandic, Maria Marschler, Bettina Müller, Lidia Pittarello, Lukas Plan, Susanne Randolf, Swen Renner, Bettina Riedel, Mario Dominik Riedl, Markus Roboch, Robert Sailer, Edmund Schiller, Martin Schwentner, Tanja Schuster, Andrea Stadlmayr, Gabriel Stöckle, Nikola Szucsich, Wencke Wegner, Anna Weinmann, Karin Wiltschke-Schrotta, Katharina Wölfl, Frank Zachos, Herbert Zettel, Dominique Zimmermann, Eva Zimmermann

DiSSCo Austria: Peter Bilovitz (Uni Graz), Roland Eberwein (Kärntner Landesmuseum), Simon Engelberger (Uni Wien), Robert Lindner (Haus der Natur), Ronald Lintner (Landesmuseen NÖ), Michael Malicky (OÖ Landes-Kultur GmbH), Konrad Pagitz (Uni Innsbruck), Astrid Scharfetter (Uni Graz), Andreas Tribsch (Uni Salzburg), Romed Unterasinger (Tiroler Landesmuseum Ferdinandeum), Kurt Zernig (Universalmuseum Joanneum)

EC: Michel Schouppe

TU-Graz: Ilire Hasani-Mavriqi

TU-Wien: Barbara Sanchez-Solis

Umweltbundesamt: Martin Götzl

Universität Wien: Tereza Kalová

Contents

Introduction.....	3
Dr. Katrin Vohland: Welcome and introduction.....	3
Part I: Open Science on a strategic and policy level.....	3
Dr. Kostas Glinos: Vision of Open Science in Europe and beyond	3
Dr. Stefan Hanslik: Open Science as a vision for Austria?! A regional perspektive on a global Vision 4	
Prof. Dr. Andreas Rauber: Open Science and the EOSC: Building Blocks for FAIRness for Public and Sensitive Data	5
Q&A	5
Part II: Open Science implementation	6
Mag. Heimo Rainer: State of Open Collections at the NHM and in Austria.....	6
Dr. Olaf Bánki, Dr. Chantal Huijbers & Markus Döring: The new Catalogue of Life – A key taxonomic service.....	7
Dr. Susanne Blumesberger: Open Access as a key contribution to Open Science.....	8
Dr. Stefan Eichert: Overcoming Borders between Disciplines – Connecting History, Archaeology, and Sciences by digital Means. The case study of THANADOS	8
Mag. Iris Ott: The Deck 50 as experimental place to engage the public.....	9

Introduction

Dr. Katrin Vohland: Welcome and introduction

[Slides](#)

The NHM was founded in 1889 as a “Cathedral of Science”. It lives from its collections throughout all research areas and performs relevant research. In the last months, we thought about our relevance in the future: The museum really wants to make an impact and would like to support the global development goals, especially by opening up our collections to a wide audience.

The values of collections respectively objects have: they can be linked to more data (genomic, biochemical, morphological data), to photographs, to backbones (geographical and taxonomic data) and also to more research data like CT-scans. So the NHM can become an institution that provides big data, which is important for answering big questions with regard to climate change, ecology or evolution. An example why open data/open collections are so important: We subscribed the European Manifesto for EU Covid-19 research. Providing all data from the bats in our collection (taxonomy, distribution etc.) helps answering questions about this virus.

This workshop is structured in 3 parts:

- I. Open Science on a strategic and policy level (speakers Glinos, Hanslik, Rauber)
- II. Open Science implementation (speakers Rainer, Bánki & Huijbers, Blumesberger, Eichert, Ott)
- III. Practical questions and discussions (in the afternoon, in German)

Part I: Open Science on a strategic and policy level

Dr. Kostas Glinos: Vision of Open Science in Europe and beyond

[Slides](#)

Science is in transition: There are various megatrends in science like the increasing production of knowledge, assimilation of this knowledge, increasing complexity, hyperspecialisation, globalisation/internationalisation and digitalisation of science. In all areas where **digitalisation** has revolutionized, in the beginning it automates the conventional process, but the point comes when the process completely changes and leads to a new process. In the area of science we are in the beginning, the traditional process. What we know is that the new processes are more network-like and will be very data intensive and global. If we want this process to be more efficient and science that is of higher quality, we need to unblock information flows. So the process needs **openness and transparency**, and will so be more trusted by society. So **open science** is defined as sharing knowledge and tools not only between researchers but also between disciplines (including machines) and society at large. Trust of society in science is one of the major challenges of time!

We have the traditional science infrastructures like universities, libraries, scientific societies, museums etc. We need a **new infrastructure to support this transition of science** – and this is what we do with the **European Open Science Cloud (EOSC)**. Main challenges: improve the practice of research and innovation (e. g., open access, reproducible results, citizen science) and develop proper

enablers (skills, technical infrastructure etc.). The current system is dominant (individual competing scientists, publish or perish etc.), whereas open science is collaborative (sharing data and knowledge and so on).

EOSC is the technical infrastructure for open science (see slide 7). There is a number of sectoral data basis foreseen like for health, agriculture, finance, and EOSC as a crosscutting data space for Research and Innovation. EOSC is also foreseen in the new **European Research Area**, it is part one of the key actions in the area roadmap. It is also one of the four pilot actions of the conclusions of the council in December 2020. EOSC is also mentioned in the **Recovery and Resilience Plans (RRP)**, that is the stimulus package of 750 billion € that was brought forward and approved by parliament and council for getting out of the economic depression because of the pandemic. Within these 750 billion € there are 672,5 billion € in loans and grants available to support reforms and investments undertaken by Member States. This is an **opportunity for any institution/stakeholder community in the Member States to bid for this money**. The EOSC is mentioned as one of the examples where the money could be spent on (for hardware, human resources etc.).

EOSC development has been an extensive creation process in the last years (esp. 2019–2020) with an **collaborative approach**: bringing everybody together (researchers, organisations, funders, service providers, European and national research infrastructures) to create a community and agreeing what needs to be done. What needs to be done on a **technical level** goes in steps: developing EOSC core (key technologies to create this web of digital objects) that will work with a federated set of open data in the beginning, and to which progressively services will be added on top.

There was a decision in 2020 to do this in a **European Partnership** approach which is offered in the new European framework programme Horizon Europe.

EOSC governance: EOSC Association (legal body under Belgian law) and EOSC steering board (EC, Member States and Associated Countries). The EOSC Association was created in July 2020, 1st General Assembly in December 2020, currently ca. 140 members, about 50 observers, Memorandum of Understanding planned signature in April 2021.

Three general objectives of the EOSC strategic research and innovation Agenda, 14 action areas (see slides 14–15). Importance of international cooperation see slide 16.

Dr. Stefan Hanslik: Open Science as a vision for Austria?! A regional perspective on a global Vision

[Slides](#)

Links for the definitions of Open Science:

<https://nsf.gov/pubs/2020/nsf20068/nsf20068.jsp>

<https://www.fosteropenscience.eu/node/1420>

OANA – Open Science Network Austria: <https://www.oana.at/>

Infrastructures with Austrian cooperations: <https://www.oana.at/nationale-aktivitaeten/support-von-infrastrukturen/>

Dr. Stefan Hanslik provided two further links in the chat:

<https://www.eoscsecretariat.eu/>

<https://www.tuwien.at/forschung/fti-support/forschungsdaten/forschungsdatenmanagement/fair-prinzipien/fair-data-austria>

Prof. Dr. Andreas Rauber: Open Science and the EOSC: Building Blocks for FAIRness for Public and Sensitive Data

[Slides](#)

Q&A

Q: What is the **function of the ORCID-consortium**?

A (AR): The consortium was founded by the TU library and the library of the University of Vienna, a lot of institutions are joining up there. The ORCID is evolving to more than just an identifier – this networking shall make people more aware of it. It is now offering additional services like providing additional metadata or linking to publications.

More details: <https://www.tuwien.at/kooperationen/orcid/>

Q: There is some kind of automatism: We open up science and the quality automatically increases. It is not always that easy (e. g., predatory journals). What are the key mechanisms you suggest to **support the quality of science**?

A (KG): What does excellence/quality mean? Funders definition: complies with objectives, has innovation, project goes beyond state-of-the-art, methodology is sound. You could add things like the research process is appropriate for making breakthrough innovations or letting others reuse your results. Shall we add highly reproduceable results too, although a very high percentage of results is not reproduceable? We need a discussion about this, but it is intuitively correct that if science is open, so the data are fully open, so that everybody can check what has been done, then certainly we will have higher quality.

Q: In EOSC, is it the idea that there is one starting point (just like the webpage of GenBank) that makes it possible to **find all this data**?

A (KG): Yes, there is a portal of EOSC – one can go there and find the data, but ideally if EOSC is successful (in 5–10 years) you should use it and don't realise that you're using it (like using the www). So if EOSC is successful, everything is interconnected and the use is unconsciously.

Q: How probable is **data storage or software as a service pool** for the EOSC?

A (AR): We see those things emerging, becoming a reality. It will not cover everything, it will be a continuous process. But you can be sure that there will be such services. They exist now with the only difference that they are isolated, discipline- and region-specific. The transparency of these services will be the big benefit of EOSC.

A (KG): This is a practical examples of the difficulties we have: Fighting the Covid-19 pandemic, a key issue is to accelerate the sharing of all kinds of data: combine genomic data of virus and host, clinical data, socioeconomic data etc. We are not able to do this!

Q: We generate data, but not every institution can create its own data interfaces and so on. Can institutions like the TU support smaller institutions in getting the data out in the cloud? Is there a possibility of **interconnecting or networking with the technical institutions**?

A (AR): It is not necessary, not even useful and desirable that each institution has to provide all the services, repositories, visualisation services, compute services etc. in-house. What we want to achieve is to enable those institutions who want to host services themselves and support them, and at the same time those who don't want to host services themselves, to consume it and offer it via other institutions. So repository as a service or compute as a service will become a reality. Some disciplines have that already. And we at the TU Wien offering such services, building them up and – hopefully via the EOSC – rolling them out on a broader scale. And the same thing does not only apply on technical things, but on legal services or ethical services too. We need to be able to combine and interlink to services, offer them and connect them. And then it really will become an EOSC.

Q: Within a current project, we already collected large data sets from archaeology and anthropology. They are available online, licensed under CC licenses, in machine-readable formats with metadata. Is there already a possibility to **implement or integrate these data in EOSC** and what are the steps to get into EOSC?

A (SH): The easiest way is to get in contact with the EOSC Association, also to apply and to show interest to join the process: <https://www.eoscsecretariat.eu/>

A (AR): Getting data in the EOSC is a rather tricky question. What does it mean to get data in the EOSC? Getting data in repositories that expose the metadata so that it can be indexed by data search engines will be the pragmatic level. There is repositories where you can upload data as file dumps, there is different institution repositories. The actual software of the TU Wien has been set up in the past few weeks and it takes another few months to bring it up and offer it as services and roll it out. So, you can use other repositories out there, you can build your own, or you can team up with other institutions.

A (KG): It is in the guidelines of the Horizon Europe programme that the data has to made FAIR, there are a lot of tools to make this. And the deposit has to be on a trusted repository.

More informations on FAIR data: <https://www.tuwien.at/forschung/fti-support/forschungsdaten/forschungsdatenmanagement/fair-prinzipien/fair-data-austria>

Part II: Open Science implementation

Mag. Heimo Rainer: State of Open Collections at the NHM and in Austria

[Slides](#)

In Austria, we have several holders of collections: federal collections and museums (the oldest one, the national library, dating back to the 14th century), regional museums like the state museums (most of them have been founded in the early 19th century), local museums/collections (e.g. Stadtmuseum Baden or Villach) and collections in monasteries (e. g. Stift Kremsmünster, Admont, Seitenstetten, Melk). Private collections usually end up in public museums.

We started working together on these collections and biodiversity data with the erection of GBIF (Global Biodiversity Information Facility) in 2001, chaired by the Umweltbundesamt. In 2007, we formulated a project called ODINS where we captured all physical objects and the degree of digitisation to calculate the costs for digitisation of these collections. The data were updated 2012 but got stuck at this point. In 2019, we had a presentation of the European DiSSCo initiative where we invited all collecting holding stakeholders, representatives from the ministries and so on.

In the GBIF initiative, we talked about defining standards for our data to being able to share our data right from the start (www.tdwg.org, www.biocase.org). We started digitising objects not only by taking metadata from labels but also 2D-scans and provide them by the internet (under FAIR principles).

In the last months, we updated the data concerning number of objects and the degree of digitisation in the institutions taking part in the DiSSCo Austria Consortium: Overall there are ca. 51 million objects, around 30 % of them digitised in a digital catalogue, and ca. 1 % with digitised media (2D-images).

Outlook: In the DiSSCo Prepare Project (2019–2023), **MIDS** (Minimum Information about a Digital Specimen) and **OpenDS** (open digital specimen) standards are developed, which should enable institutions to share their data on different levels of data depth:

MIDS level 0: a skeletal record, just an identifier and the institution

MIDS level 1: basic record, includes also name and material type

MIDS level 2: all information that is available from the label(s)

MIDS level 3: includes also derivatives like genetic data or literature

(MIDS levels – especially 2 and 3 – are still in discussion!)

DiSSCo readiness: In the DiSSCo Austria group we are currently discussing 4 dimensions of readiness (scientific, data, technological and organisational readiness). Financial readiness is still pending topic.

We had a discussion with **FAIR office Austria** and were ventilating possibilities of a FAIR implementation network (i. e. forming a network providing data under the FAIR principles, setting up services or repositories that are trustworthy over time: stable, maintained, financially supported, technically stable).

We have connections to **RDA Austria** (Research Data Alliance Austria, <https://www.rd-alliance.org/Austria>, head: A. Rauber) with ongoing discussion about integration of all research and object data in EOSC.

We could also think of having an **open collections group inside EOSC**. Working groups are now re-set, so this is a good time to bring up this idea.

We should also think of the **correlation of physical objects to a research object**.

Link to EOSC: <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>

The end-phase is the participation in the **LOD cloud** (Linked Open Data cloud, <https://www.lod-cloud.net/>).

Dr. Olaf Bánki, Dr. Chantal Huijbers & Markus Döring: The new Catalogue of Life – A key taxonomic service

[Slides](#)

The Catalogue of Life (COL) is one of the oldest biodiversity informatics initiatives that are still running (since 20–25 years): nearing 2 million species globally, mediated by 163 taxonomic databases, maintained by hundreds of global taxonomists, large group of users (scientists, institutes, higher policy initiatives...), annual editions.

In December 2020, the new COL infrastructure was launched (powered and hosted by GBIF). It consists of three parts:

- new portal (<https://www.catalogueoflife.org/>) with tree browser, new dynamic data components (e. g. for picking up a part of a tree), taxon and higher taxon pages, contributing datasets. Data sources can be properly cited. There is also access to previous editions.
- COL ChecklistBank (<https://data.catalogueoflife.org/>): built on original checklistbank from GBIF.
- new checklist bank API (<https://api.catalogueoflife.org/>): with stable name/taxon IDs. Legacy API is ported, but recommend everybody to switch to new API service.

Alliance for biodiversity knowledge (www.biodiversityinformatics.org): a global collaboration to create a shared infrastructure and data ecosystem for biodiversity knowledge. The aim is not only to share data openly between initiatives, but also share, maintain and develop tooling and infrastructure jointly. The initiatives Catalogue of Life, Global Biodiversity Information Facility, Biodiversity Heritage Library, Barcode of Life, and Encyclopaedia of Life have formed together with the COL partners ITIS, Natural Biodiversity Center, Illinois Natural History Survey are part of the Steering Committee of the new Catalogue of Life infrastructure project. In 2019, the LifeWatch ERIC and DiSSCo have joined the Steering Committee. Each of these partners commit themselves in using the new Catalogue of Life infrastructure for their taxonomic backbone services. The next step for the new Catalogue of Life infrastructure is to build an Extended Catalogue Of Life as a 1st candidate for the replacement of the GBIF Backbone Taxonomy. The Catalogue of Life is also in close discussion with the DiSSCo initiative in delivering taxonomic backbone services for DiSSCo.

Dr. Susanne Blumesberger: Open Access as a key contribution to Open Science

[Slides](#)

General information about Open Access <https://open-access.net/en/information-on-open-access/what-is-open-access>

Golden OA: publishing in OA journals

Green OA: providing OA to articles that have appeared in (closed-access) journals by depositing a version of the work in an OA repository

Phaidra (repository for digital objects at the University of Vienna): <https://phaidra.univie.ac.at/>

Dr. Stefan Eichert: Overcoming Borders between Disciplines – Connecting History, Archaeology, and Sciences by digital Means. The case study of THANADOS

[Slides](#)

Thanados: <https://thanados.net/>

Stefan Eichert also provided a document with the terms/definitions of Open Source, Open Data and Linked Open data: [Link](#)

Mag. Iris Ott: The Deck 50 as experimental place to engage the public

[Slides](#)

According to the Open Innovation Strategy for Austria opening up the research process rests on four key pillars, each of which contributes added value to both the scientific community and civil society: Adaptivity, anticipation, openness and inclusion. Establishing open innovation and experimental places to incorporating citizens into the scientific process is one of the main aims – Deck 50 is such a place.

Deck 50 is the **new room for science communication at the NHM**. Our partner in this project is the Ars Electronica Futurelab. Together we created a multifunctional communication platform offering new forms of participative communication by combining science, art and new technologies forming a professional framework for Citizen Science and dialogue. Although many citizen research projects are based on digital platforms, this exchange of ideas and knowledge between researchers and members of the public also needs physical spaces to convey the importance of Citizen Science and open science.

The name Deck 50 is based on the old numbering system used for the rooms at the museum. The word “Deck” also symbolises a platform or promenade, a runway or landing – a ramp from which to launch ideas, research and discussions. Designed as a multi-perspectival communication platform, Deck 50 sees itself as an **open space for innovation and experimentation** where visitors can take part in social discourse through action and new forms of participative communication.

- Deck 50 makes the connection between research and society visible and tangible for different age groups.
- Deck 50 is a space for research, experimentation, exchange with experts, interactive exhibits and sustainable participation processes.
- The focus is on mutual enrichment through new perspectives in which the opinions and views of individuals become more important.
- The framework for the content is provided by a wide range of educational activities, innovative dialogue and close cooperation with the museum's researchers.
- The spatial elements are attractively designed in terms of content, aesthetics and media, and also attract casual museum visitors who are less actively involved.

Principles and recommendations for action have been defined in order to achieve these objectives:

Identify topics

- Involving scientists
- Integrating Citizen Science projects
- Seeking potential benefits for science

Create flexibility

- Design topics of participative formats to be easily interchangeable / easily adaptable
- Allow failure (of experiments, contents, visitor interaction) and include it in planning
- Offer different levels of participation
- Diversity as a basic principle

Access for all

- Events and activities should be open and accessible to everyone
- Events and activities should be linked to the realities of life

Benefits for research

- Public participation using different formats / multi-perspectivity
- Visitor input should be evaluable and able to flow into research
- Data must be made easily accessible and usable for museum employees

Sustainability for museum and visitors

- Visitor input is visibly and attractively curated for subsequent visitors
- Echo of participation within the institution
- Visitors find out what happens with their contribution

Following these principles, different forms of participation were considered and examined for their applicability in a museum context. A distinction was made between the intensity of the participation level and the general requirements (infrastructure, spatial requirements, material, etc.). The result is a matrix with desired participation formats, which provides the spatial design for Deck 50. This makes it possible to divide Deck 50 into **three areas**:

The **Laboratory** is an experimental field and knowledge workshop. Here, visitors can experience how science works, observe for themselves and draw confident conclusions. The focus is on experiencing research as a process and on gaining scientific knowledge.

The **Meeting Zone** is a recreation and discussion room. Comfortable seating invites participants to linger and offers space for interaction between visitors. The content and design of the Meeting Zone follows the "three pillars of participation", which were developed in a co-design process by museum staff, visitors and external experts from the Natural History Museum Berlin: 1. Create spaces for participation which are characterised by a welcoming atmosphere and openness for different uses. 2. Enable multi-perspectivity. 3. Bring debates on social topics into the museum.

The museum as a **stage** on which scientific connections can be experienced and understood. The stage as a communication booster, where socially relevant topics of our time find their place.

The aim of the Stage is not only to communicate information grounded in science and stimulate research-based learning, but also to leave room for personal perspectives. The Stage strengthens and highlights the social and cultural competence and relevance of the NHMV. It ensures the continued interest of the visitors and broadens ownership of the museum.

Opening in September – come and see! There is a short video to give you a first impression of the room: [Link](#).

Protokoll: 23.2.2021, MM, HR