OPEN ACCESS

CrossMark

Project Report

Permits, contracts and their terms for biodiversity specimens

Edmund K. Schiller[‡], Karin Wiltschke-Schrotta[‡], Eva Häffner[§], Jutta Buschbom[|,¶], Frederik Leliaert[#], Breda M. Zimkus[¤], John B. Dickie[«], Suzete R. Gomes[»], Chris H.C. Lyal[¶], Daniel Mulcahy[^], Alan Paton[«], Gabi Droege[§]

‡ Naturhistorisches Museum Wien, Vienna, Austria
§ Freie Universität Berlin, Berlin, Germany
| Statistical Genetics, Ahrensburg, Germany
¶ Natural History Museum, London, United Kingdom
# Meise Botanic Garden, Meise, Belgium
¤ Museum of Comparative Zoology, Harvard University, Cambridge, MA, United States of America
« Royal Botanic Gardens, Kew, London, United Kingdom
» Instituto Oswaldo Cruz, Rio de Janeiro, Brazil
^ Museum für Naturkunde, Berlin, Germany

Abstract

We present two different typologies of legal/contractual information in the context of natural history objects: the Biodiversity Permit/Contract Typology categorises permits and contracts, and the Typology of Legal/Contractual Terms for Biodiversity Specimens categorises the terms within permits and contracts. The Typologies have been developed under the EU-funded SYNTHESYS+ project with the participation of experts from outside the consortium. The document further addresses a possible technical integration of these typologies into the Distributed System of Scientific Collections (DiSSCo). The implementation in the DiSSCo data model is outlined and a concrete use case is presented to show how conditions, e.g. the Typology of Legal/Contractual Terms, can be introduced into the DiSSCo Electronic Loans and Visits System (ElViS). Finally, we give an outlook on the next steps to develop the typologies into a standard that supports compliance with legal and contractual obligations within the wider community of natural science collections.

Keywords

Access and Benefit Sharing, Anthropological collections, Biodiversity collections, Collecting permit, Collection management system, Contracts, Data governance, Data standards, DiSSCo, DNA sequencing, Genetic resources, Geological collections, Herbarium, Legal information, Natural history collections, Open Digital Specimens, Permits, SYNTHESYS+, Zoological collections

1 Acronyms

Acronyms for participating institutions are listed in a subsequent section.

ABS - Access & Benefit-Sharing[*1]

ALA – Atlas of Living Australia

APHIS – the USDA's Animal and Plant Health Inspection Service

BBNJ - Biodiversity Beyond National Jurisdiction[*1]

BCoN - Biodiversity Collections Network (USA)

CARE - Collective benefit, Authority to control, Responsibility and Ethics

CBD - UN Convention on Biological Diversity[*1]

CD - our acronym for indicating that the lawful source of a use-term is a contract or deed (hence CD) under civil law

CETAF - Consortium of European Taxonomic Facilities

CITES - Convention on International Trade in Endangered Species of Wild Fauna and Flora[*1]

D - our acronym for indicating that the lawful source of a use-term is an individual document issued to a specific person or legal entity, based on a public administration's decision (often after examination of a specific project)

DES - Digital Extended Specimen (concept), see Hardisty et al. (2022)

DiSSCo - the Distributed System of Scientific Collections - a long-term European initiative and research infrastructure in preparation

"DSI" - "Digital Sequence Information" as discussed in the CBD context

DOA - Digital Object Architecture

EDP - Electronic data processing

ELViS - European Loans and Visits System (under construction in the framework of DiSSCo)

FAIR - Findable, Accessible, Interoperable and Reusable

GBF - Kunming-Montreal Global Biodiversity Framework of the CBD

GBiOS - Global Biodiversity Observation System

GeoBon - Group on Earth Observations Biodiversity Observation Network

GGBN - Global Genome Biodiversity Network

GMO - Genetically Modified Organisms

GR - Genetic Resources[*1]

iDigBio - Integrated Digitized Biocollections (USA)

INSDC - International Nucleotide Sequence Database Collaboration

IPBES - Intergovernmental Platform on Biodiversity and Ecosystem Services

IPLC - Indigenous Peoples and Local Communities

IPR - Intellectual Property Rights

IRCC - Internationally Recognised Certificate of Compliance

IRI - Internationalised Resource Identifiers

ITPGRFA - International Treaty on Plant Genetic Resources for Food and Agriculture

JSON - JavaScript Object Notation

JSON-LD - JavaScript Object Notation for Linked Data

L - our acronym for indicating that the lawful source of a use-term is a law without individualised document, the law however can be referenced

LN - our acronym for indicating that the lawful source of a use-term is a written evidence that no individual document is necessary for a permission, but the evidence may lack information whether this is based on written law or missing legal provisions

MAT - Mutually Agreed Terms

MBTA - United States of America Migratory Bird Treaty Act

MIDS - Minimum Information about a Digital Specimen in the framework of DiSSCo

MoC - Memorandum of Cooperation

MoU - Memorandum of Understanding

MTA - Material Transfer Agreement

N - our acronym for indicating that the lawful source of a use-term is the fact that no legal provisions exist (hence N) nationally or on the level of applicable supranational (e.g. EU) legislation

NP - Nagoya Protocol[*1]

OA - W3C-recommended Web Annotation Ontology

ODRL - W3C-recommended ontology of the Open Digital Rights Language

openFDO - open FAIR Digital Objects

openDS - the "open Digital Specimen" specification

PIC - Prior Informed Consent

PID - persistent, globally unique and resolvable identifier

PIP - Pandemic Influenza Preparedness Framework[*1]

PROV - W3C-recommended Provenance ontology

REL - Rights Expression Language

ROR - Research Organization Registry

SMTA - Standard Material Transfer Agreement

SPNHC - Society for the Preservation of Natural History Collections

SDR - Specimen Data Refinery in the framework of DiSSCo

TDWG - Biodiversity Information Standards (originally called the Taxonomic Databases Working Group)

TK - Traditional Knowledge in the meaning of the Nagoya Protocol[*1]

UID - unique Identifier

UN - United Nations

UNDP - United Nations Development Program

URL - Uniform Resource Locator

USDA - United States of America Department of Agriculture

W3C - World Wide Web Consortium

Institution acronyms used in this text:

BGM - Agentschap Plantentuin Meise

BGBM - Botanischer Garten und Botanisches Museum, Freie Universität Berlin

CSIC - Agencia Estatal Consejo Superior de Investigaciones Cientificas, Madrid

GBIF - Global Biodiversity Information Facility

HNHM - Hungarian Natural History Museum (Hungarian: Magyar Termeszettudomanyi Muzeum), Budapest

HUJI - The Hebrew University of Jerusalem

LUOMUS - Finnish Museum of Natural History (Finnish: Luonnontieteellinen keskusmuseo), University of Helsinki (Finnish: Helsingin Yliopisto), Helsinki

MfN - Museum für Naturkunde - Leibniz-Institut für Evolutions- und Biodiversitätsforschung, Berlin

MNHN - Muséum national d'Histoire naturelle, Paris

Naturalis - Stichting Naturalis Biodiversity Center, Leiden

NHM - Natural History Museum London

NHMW - Naturhistorisches Museum Wien

NRM - Naturhistoriska Riksmuseet Stockholm

RBGE - Royal Botanic Garden Edinburgh

RBGK - Royal Botanic Gardens Kew

SGN - Senckenberg Gesellschaft für Naturforschung, Frankfurt

SMNS - Staatliches Museum für Naturkunde Stuttgart

UCPH - Kobenhavns Universitet

UGOT-GGBC - Goeteborgs Universitet, Gothenburg Global Biodiversity Centre

ZFMK - Zoologisches Forschungsmuseum Alexander Koenig, Bonn

2 Executive Summary

In the EU-funded SYNTHESYS+ project we addressed the challenge to link legal and contractual information with biodiversity specimens and data. The necessity for linking them arose from events in the years 2009/10 that dramatically changed the conditions for

biodiversity research: Between October 2007 and October 2009 the costs of DNA sequencing dropped from approximately 400 USD to below 1 USD per raw megabase (Wetterstrand 2021) and on 29 October 2010 the *Nagoya Protocol* (NP) *on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization (ABS) to the Convention on Biological Diversity* (CBD) was adopted (it entered into force on 12 October 2014).

The effects were on one hand a massive increase of genetic sequencing in biodiversity research. On the other hand, more and more countries devised their specific national set of regulations for researchers' access to their biodiversity, which several equalise with "genetic resources". Genetic resources*[1] are a key concept in the framework for Access and Benefit Sharing*[1] provided by the CBD and NP. This situation poses considerable challenges for genetic studies in biodiversity research that typically include specimens from different countries. The administrative burden has been increasing enormously.

Our work intends to support an efficient and systematic way of dealing with the multitude of ABS regulations that may be attached to specimens in biodiversity collections (tangible specimens or their electronic representations). Furthermore, the presented resources extend the range beyond ABS regulations and include a wide variety of permits and contracts typically associated with biodiversity collection specimens.

Our work provides the possibility of updating electronic data processing systems with data standards for delivering information not only on permits and contracts, but also on typical terms they include. In this way it becomes possible to flag all specimens in biodiversity collections with their respective legal and contractual information, to automatically display these flags and to search for specimens by selecting legal/contractual terms.

The present document starts with depicting the scientific, infrastructural and policy background of the current work (**chapter 3**), then describes the methodological approach to the problem (**chapter 4**) and the results of the analysis of the community practices and needs regarding the management of legal information in filing and data processing systems (**chapter 5**). The analysis includes a) the results of a survey and detailed interviews with SYNTHESYS+ partners exploring their practices regarding the management of loans and associated documents, and b) a workshop that went beyond the scope of biological collections to get a broader view of policies and regulations present in different communities holding natural science collections and the challenges they pose to documentation and data management. From the analysis, typical categories of use cases were derived that come with different levels of complexity and risk.

In **chapter 6** we present the Biodiversity Permit/Contract Typology which can be used for establishing a comprehensive filing system (physical and/or electronic) for documents on important events in the life-cycle of a collection object. Additionally, it can serve as a revised version of the GGBN data standard's Permit Vocabulary, now also including different contracts, among them, loan contracts.

The Typology consists of seven Document Categories (**1**-Access & Benefit-Sharing*[1] **(ABS) Documents, 2-High Level Arrangements, 3-Permits for Collecting & Related/ Taking/Possessing, 4-Permits for Research, 5-Permits for Special Purposes (excluding ABS), 6-Material Transfer Agreements, Stewardships & Ownership-related Information, and 7-Transport Documents**) with altogether 38 Document Types (Annex 1), enabling flexible use by natural history collections. Each Category and Type respectively is supported by a description of its meaning and scope.

However, the Typology does not include permits/contracts for research with human tissue, human pathogens or genetically modified organisms. Documents focusing on intellectual property rights such as patents are also out of scope of this study.

The two-tiered approach of Document Categories and Document Types provides thematic clusters, graded specificity and the opportunity to flexibly apply these two levels of granularity. It is up to the institutions to use Document Categories and Types according to the needs of their collection or content management systems or document filing system. Aggregating platforms like GGBN might consider implementing the full set of suggested Categories and Types to be prepared for varying contributions from different institutions.

Chapter 7 presents the Typology of Legal/Contractual **Terms** for Biodiversity Specimens, which complements the Biodiversity Permit/Contract Typology that classifies complete documents, but cannot provide reliable information on permitted actions, due to the lack of international standardisation of document contents. The Typology of Legal/Contractual **Terms** for Biodiversity Specimens showcases terms for specific actions performed on and with biodiversity specimens that otherwise would be hidden in multiple hard-to-read documents. It creates a basis for easily and swiftly exchanging information on what can be done with biodiversity specimens, and under which conditions.

For the Typology of Legal/Contractual Terms we identified five types of lawful sources, i.e. *individualised documents issued by authorities, laws, missing legal provisions for certain actions, contracts/deeds, and written evidence that no individualised documents are necessary*. We further identified four characteristics that may be contained in these lawful sources, i.e. *permissions, prohibitions, duties, and restrictions*. Additionally, we identified *expert opinions* as a separate element that is often necessary for fulfilling legal terms and therefore must be included in the Typology (e.g. expert opinions stating that a specimen is free of known pests).

Resulting from this SYNTHESYS+ project 87 different terms have been included in the Typology of Legal/Contractual Terms for Biodiversity Specimens: 50 terms on a general level, 26 specific terms for complementing - or at will replacing - them, and 11 specific terms typically used in loan contracts issued by biodiversity collections (Annex 2). Future work may add more terms to the typology. Since science is highly integrated on an international level, modifying single definitions for the standardised Document Categories, Document Types and Legal/Contractual Terms should involve as many partners from other countries as possible, to maintain consistent understanding.

The next two **chapters 8 & 9** of our work contain information to support the implementation of the data standards for both typologies in electronic data processing systems. Thereby, the proposed implementation approach is independent of the scope of the data infrastructure system into which the typologies and associated functionality will be integrated. It can be applied to infrastructures that support information needs and applications at the international, (supra)national, or institutional level. To be able to show the transition from abstract conceptual work to the integration of programming code into a concrete, existing infrastructure environment, we showcase an exemplary implementation in the European Distributed System of Scientific Collections (DiSSCo) research infrastructure and its ELViS loans and visits module (European Loans and Visits System). The use case we developed consists in the evaluation of a loan transaction for a physical specimen, involving the staffs of the institution holding the specimen and of an institution requesting to borrow it.

Chapter 8 introduces the technical architecture of DiSSCo, its building blocks of open digital objects that are specified as open Digital Specimens (openDS) for the biodiversity domain and an events-based transactional model that forms the foundation of this highly cooperative and dynamic infrastructure.

In the following **chapter 9**, the functionality of the basic transactional model is expanded by integrating functionality for conditions and their machine-actionable assessments as the basis for a final step of human decision-making. Example code for the loan transaction use case is based on vocabularies and information models provided by the Provenance ontology (PROV), the Web Annotation ontology (OA) and the ontology of the Open Digital Rights Language (ODRL).

3 A jungle of juristic requirements attached to biodiversity specimens

For decades or even centuries biodiversity collections have been used to categorize organisms. Accordingly, collection holders and their employees are used to file collecting permits, export- and import permits, and loan forms supporting fundamental scientific research. For most of these decades or even centuries hardly anybody asked for the filed legal documents, which additionallly represented only limited variety. This changed at the beginning of the 21st century.

With the closing of the first decade of the 21st century genetic research became much cheaper. Consequently, the number of sequences deposited in public databases multiplied (Fig. 1). This alone might not have increased the interest in legal documents supporting the collection and holding of biodiversity specimens. But at the same time the international Nagoya Protocol (NP) was negotiated (adopted in 2010), adressing utilisation of genetic resources (the full title of the Protocol is "Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity 2010"). The protocol established a mechanism to enforce compliance with provider countries' regulations by the Parties' jurisdictions. It

applies to a specific set of activities, namely "to conduct research and development on the genetic and/or biochemical composition of genetic resources, including through the application of biotechnology as defined in Article 2 of the Convention" (NP Article 2c). Most likely, the combination of a soaring amount of genetic research, and governmental efforts to claim their vested rights on benefits resulted in a jungle of juristic requirements attached to biodiversity specimens. Until February 3rd 2023 a number of 45 countries submitted national regulations on "Access & Benefit Sharing (ABS)" to the ABS-Clearing House, administered by the Secretariat of the Convention on Biological Diversity. We are aware that more countries have such regulations. Today any genetic research, and this includes the public interest of research for supporting and protecting biodiversity for future generations, is subject to many more regulations, as well as many more associated documents. However, not only the increased number of necessary legal documents adds complexity. In addition a multitude of different national names exists for basically the same type of legal/contractual document. Moreover - and probably worse - between countries identical names are used for legal/contractual documents with *differing* intentions. As a consequence of the current situation, it is not sufficient to just file such legal documents. Instead, it is necessary to be able to quickly and reliably find the correct documents even after decades, to be able to consult them, and most importantly to quickly gain insight into their contents. Overall, to comply with up-to-date legal requirements, as well as to inform planning, reviewing und monitoring as part of the Global Biodiversity Framework (Conference of the Parties to the Convention on Biological Diversity 2022) legal/contractual documents filed by biodiversity collections in the 20th century have become of interest again.

The motivation for the work presented in this publication was to enable easy and transparent lawful handling of biodiversity specimens assembled from various jurisdictions. At present, no standards and no infrastructure systems with the required functionality exist that support traceability of legal/contractual restrictions on engaging with biodiversity specimens. The results presented here, including a range of proposed options for implementation, are the best imaginable practice the authors could think of when debating the needs of the institutions they work for. Nevertheless, it is no "best practice" in the meaning of "a method or technique that has been generally accepted as superior to other known alternatives" (Wikipedia contributors 2022a). However, the authors believe that this publication provides the foundation for such a standard and expedites the selection of a standard that is generally accepted.

The subject of the EU-funded SYNTHESYS+ task and Deliverable 3.3 described in this publication was to "Create a data standard for enabling traceability of restrictions for molecular samples". The motivation for setting this specific SYNTHESYS+ task was to facilitate the handling of biodiversity specimens as they are usually included in taxonomic or phylogenetic research, or in any biodiversity collection, according to the requirements of various jurisdictions. To this end, we have expanded the objective of the work beyond the original scope to cover other characteristics of documents than just restrictions: permissions, prohibitions and duties as well as expert opinions. Since molecular samples cannot be seen in isolation but are linked to other objects such as dead or living collection

specimens, we extended our focus from molecular biology samples to all tangible biodiversity specimens usually housed in natural history museums and botanical gardens and forming the professional background of the authors. Due to the multinational composition of the institutions involved in this task, we could pursue the aim of creating a document that addresses not only decision makers in these institutions, but also a wide range of other entities' decision makers with a biodiversity background from all over the world.
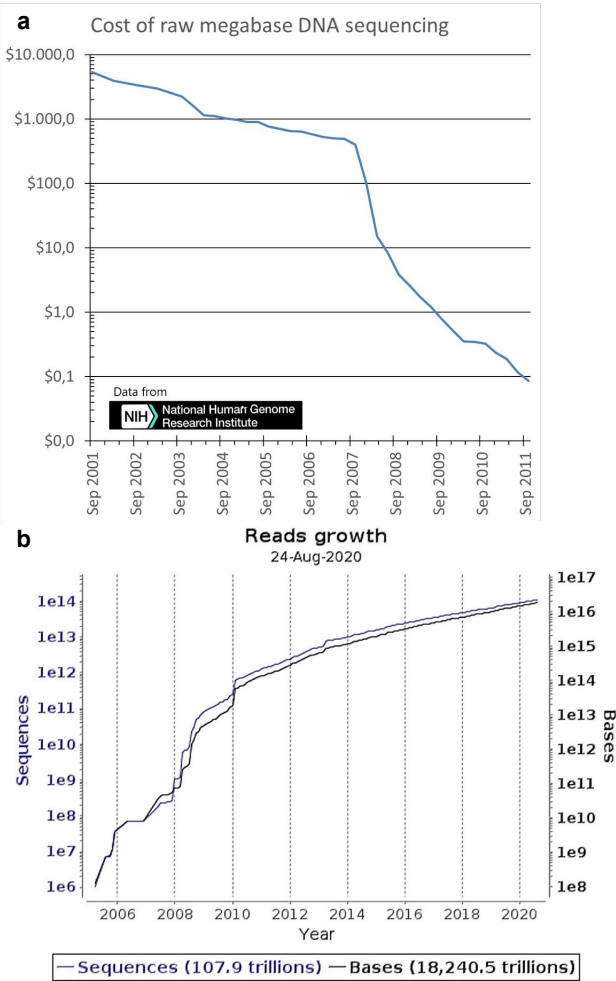
Figure 1.

At the closing of the first decade of the 21st century the costs of genetic sequencing dropped fundamentally, coinciding with soaring numbers of publicly available sequences. After that time the growth rate of sequence numbers decreased.

a: Decreasing sequencing costs, recorded by the US National Human Genome Research Institute NHGRI (data from Wetterstrand 2021) doi
b: European Nucleotide Archive data growth of sequence and bases reads (copied from Harrison et al. (2020)) doi

Data standards are needed to communicate unambiguously information on permits, contracts, agreements or terms relevant to biodiversity collections so that commitments described in such documents can be implemented. We present a standardised system of naming and semantically defining not only such documents, but also terms within them, thereby providing suggestions for extending the scope of the current GGBN data standard. The goal of this work is to create transparency and understanding of the legal terms and conditions attached to a specimen from the time of its collection throughout all uses of the specimen, its derivatives, and related collection data.

3.1 Using biodiversity objects and data for science and societal applications

Scientific collections are profoundly service oriented in their mission for preserving and facilitating research and education on tangible material, i.e. of bacteria, archaea, unicellular eukaryotes, fungi, plants, animals, and human tissue samples (the latter are out of scope of this study). Naturally, this also applies to data and information derived from and associated with this material. This endeavour of scientific collections is embedded in the highly social and dynamic context of science and requires the cooperative contributions of many, who are continuously maintaining, updating, annotating, extending and further developing scientific collections.

This highly cooperative context of scientific collections will need to be reflected and accommodated in present and upcoming biodiversity data infrastructures, intended not only to showcase biodiversity, but also to enable research and education on its components and associated electronic records. These biodiversity data infrastructures excel in fulfilling their purposes when interlinked by a globally shared, harmonised, and interactive infrastructure. Extensive discussion and development for such a global infrastructure are underway (see chapter 8.1). The purpose is to create an extended and powerful, since information-rich, global data network, which requires transparent data governance. Our work contributes to transparency by empowering data providers and users through integrating conditions into data models, in support of rights-based approaches and participatory decision-making.

Natural history museums, botanic gardens, zoos and aquaria serve a societal purpose beyond entertainment. The services that arise from managing, preserving and sharing physical objects, living organisms and digital data maintained in scientific collections form the foundation for sophisticated research. It has the unique feature of relying on biodiversity specimens collected over a long period of time, constituting unique evidence for evolutionary history and the changes in environment over time.

Present biological research relies - at least in part - on analysing the genetic and/or biochemical characteristics of specimens and/or species. This molecular biological research often gains relevance by including samples from many different countries. The resulting large, information-rich data sets can provide the foundation for powerful, parameter-rich analytical approaches, which can produce well-resolved and reliable

results, answering pressing societal questions. Currently, such transnational sampling designs require, e.g., researchers, businesses or public planners, to follow a multitude of different legal or contractual permissions, obligations or prohibitions on collecting, taking, possessing and using biodiversity, adapted to each country's individual needs and legal framework.

The knowledge of especially "foreign" countries' modalities governing accessibility, handling and use of biodiversity is limited not only among biologists, but also among many other professions involved in this topic.

Information on - and traceability of - legal and contractual permissions, prohibitions, duties and restrictions for the use of biodiversity, including for the use of molecular biology samples, is a vital requirement. This applies not only to any kind of biorepository, including natural history collections, living collections and biobanks, but also to scientific research and to international governmental and institutional cooperation for protecting biodiversity. Biodiversity data infrastructures with the ability to inform users of such legal issues can contribute to filling this knowledge gap.

The outcomes of the SYNTHESYS+ Task 3.3 working group reported here, listing and describing the most common permit and contract information related to biodiversity specimens in collections, may also contribute to the development of governance structures and functions for such biodiversity data infrastructures.

Accessibility and the use of physical objects via loan transactions and the investigation of individual collection objects by visitors are of equal importance to the sharing of data. Data include e.g. digital representations of the physical objects and born-digital field observations (directly typed into or saved by a mobile digital device), as well as information derived from and associated with biodiversity specimens and observations (such as taxon identifications, locality information, legal and contractual information, DNA-sequences, 3D-scans, multimedia recordings, metabolomics, traditional knowledge, taxonomic literature, and much more).

The technical functionality that we outline in this paper for integrating conditions into the next generation of biodiversity data infrastructures is forward-looking and generally applicable. We focused our work on legal conditions, including contractual documents. The specifics of categorising and comprehensively representing other conditions, for example arising from social and ethical considerations, intellectual property rights (IPR) and further categories of conditions, will need to be considered and developed in future work.

Our initial motive for working on standardised names and semantic definitions was to alleviate the administrative burden of a special type of regulations, i.e. the huge variety of national "Access and Benefit Sharing"*[1] regulations, based on the Convention on Biological Diversity with Annexes (1992) and subsequent international treaties, especially the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity (2010) that entered into force in 2014. This variety has been a major issue for collection

management and basic scientific research since then (e.g. Zimkus and Ford (2014), Watanabe (2015)).

The extent of our work then expanded to comprise other permit and contract documents related to biological collection objects (including natural history collections, living collections and - in part - biobanks) and researching them based on attached permissions, duties, restrictions and prohibitions.

3.2 Landscape overview: institutional to global infrastructures

Getting information on permits, contracts, and their respective terms governing physical collection objects from biodiversity is easiest with electronic data processing systems (EDP). Only small biodiversity collections can provide it comparably easily with manual filing systems.

Information on the ways of using EDPs, or more precisely, on the software used for managing biodiversity collections on the institutional, European and international level, helps to understand how best to deliver legal/contractual information.

3.2.1 Institutional Collection Management Systems

Natural History Museums use various software for managing their collections, and to manage information about their collection objects. A workshop organised in the framework of SYNTHESYS+ task 3.3 (see chapter 5.2) showed that open source software, proprietary software and individually designed databases are used. Participants listed the proprietary software: Adlib, Axiell and Filemaker. Open source software are Specify (Biological and Earth Sciences), DINA, Diversity Workbench, JACQ (herbaria), Arctos and Symbiota.

Several of these programs provide the feature of attaching or linking documents to records of biodiversity specimens. These documents include permits and contracts, e.g. loan contracts. However, to our knowledge no standardised tags are used for such attached or linked documents. As a result, information from different collections cannot simply be compared and is usually only easily understood by staff if properly trained, but not by other users. By comparing permit/contract information from different collections, we mean general information such as the purpose of a document or the total number of a particular type of permit, not the full contents of a permit, which may include confidential information.

Another issue is the terms themselves contained in permits and contracts applicable to biodiversity specimens in a collection. As far as we could determine, there are no workflows in institutions to systematically extract terms from documents that are relevant for the future, e.g., a term such as the duty to provide a copy of every publication on a specimen. Certainly, this would bind considerable resources if done retroactively, but it could be feasible for new material added to a collection. A reason for the lack of such workflows may be that no standard terms exist that can be used for tagging. It is important to note that such tagging of terms for a certain biodiversity specimen must not replace

necessary legal counsel prior to initiating actions with it, but the tagging could save time when preparing such actions. In this respect, tagging and compiling terms give a quick overview but do not replace the obligation to read the permits and contracts.

3.2.2 An upcoming European research infrastructure: DiSSCo

The mission of the Distributed System of Scientific Collections (**DiSSCo**) is to develop a next-generation research infrastructure for biodiversity data. It is a long-term initiative and infrastructure development process that originated with the Consortium of European Taxonomic Facilities (CETAF).

DiSSCo is to provide the services required for large-scale and transdisciplinary research. Research that in turn is expected to support the continuing development of a wide range of operational applications that address today's societal challenges. Up-to-date information-intense basic research into biodiversity and its transition into operational infrastructure tools are needed to effectively and reliably inform and support, e.g., biodiversity conservation, ecologically sound nature-based solutions and ecosystem-based services, agri- and aquaculture breeding programs, sustainable management of natural populations in fisheries and forestry, One Health approaches, and much more.

A comprehensive approach to protect and conserve the planet's biodiversity is the **Kunming-Montreal Global Biodiversity Framework** adopted on December 19th 2022 (GBF, Conference of the Parties to the Convention on Biological Diversity 2022) at the 15th conference of the parties of the UN Convention on Biological Diversity*[1] (CBD, Convention on Biological Diversity with Annexes 1992). The vision of the GBF is for humans to live in harmony with nature (section F of the GBF). It also contributes to the UN Sustainable Development Goals (SDGs, United Nations General Assembly 2015). In support of the monitoring strategy for tracking the GBF's progress and achievements, powerful information- and communication-technical (ICT) data infrastructures are needed, such as DiSSCo that is further expanded in its applicability to the cross-continental level by its alignment with the **Digital Extended Specimen concept** (DES; Hardisty et al. 2022).

3.2.3 GGBN and other international infrastructures

At least six international data infrastructures provide contents related to biodiversity specimens (compare chapter 4.4.2). They may be interested in making available information on associated permits, contracts and their terms, and hence in the work presented here. These six infrastructures are the Global Genome Biodiversity Network (GGBN), the International Nucleotide Sequence Database Collaboration (INSDC), JACQ (herbaria), the Integrated Digitized Biocollections (iDigBio), the Atlas of Living Australia (ALA, Commonwealth Scientific and Industrial Research Organisation CSIRO 2023), and the Global Biodiversity Information Facility (GBIF).

The Global Genome Biodiversity Network (GGBN) is the only international data infrastructure that already displays some information on permits. The GGBN website includes a searchable catalogue of genomic samples (Global Genome Biodiversity

Network et al. 2023) which not only displays species identification, collection/taking event, extraction and available sequences. It also displays loan information and permit information (Fig. 2), based on the GGBN Permit Vocabulary and GGBN Loan Vocabulary.

Permit(s)

Permit Type:
Permit Status: Permit not required
Permit Status Remarks:no national requirement for a permit at date of access
Permit Remarks:

Figure 2. doi

Sections from the GGBN webpage on a seashell sample (Astarte montagui (Dillwyn, 1817) Catalogue Number ABMBS 172-10 of the Centre of Biodiversity Genomics, Canada), showing how the GGBN Permit Vocabulary and GGBN Loan Vocabulary (from the "GGBN data standard v1") is used by the GGBN data portal.

4 Methodological approach

The scope of documents considered was based on the experience of members of SYNTHESYS+ Task 3.3 in managing biological collections, and encompassed the documents that the group had encountered. We also included previous relevant work such as that of GGBN and SPNHC and information gathered from other collection managers via a survey and workshop as outlined below.

Although we regarded a set of permit templates from the US and the EU (Suppl. material 8) we did not systematically research specific jurisdictions for possible permits and contracts. Only already available documents were included in our scope.

4.1 Starting point: GGBN data standard v1 & SPNHC permitting webpage

We considered a wide range of official notifications and private contracts, and already existing compilations of such documents.

The **Global Genome Biodiversity Network (GGBN)** released the first version of its "GGBN data standard" at the end of 2016 (Droege et al. 2016). This standard comprises 9 different vocabularies:

- for amplification,

- DNA cloning,

- gel imaging,

- loans,

- material samples,

- permits,

- preparation,

- preservation

- and for single reads.

The GGBN vocabulary for permits (Suppl. material 5), but also the vocabulary for loans (Suppl. material 6) served as cornerstones of our work.

For permits the "GGBN data standard" version 1 provides a vocabulary with 30 terms that do not only describe different permit purposes, e.g., Collecting Permit, or Memorandum of Understanding (the GGBN Data Standard uses the word "label" for these terms), but also terms that refer to administrative aspects of the permit, e.g., whether the permit is (publicly) available or not. The data standard then supplements each term ("label") with a normative URI (uniform resource identifier), a definition, information whether it is "required" and/or "repeatable", and in the case of administrative aspects with examples for possible options. We considered only 20 terms and definitions that describe the purpose of the permit as a basis for our work (Table 1) and because of limited resources we discarded five of them (human pathogens, genetically modified organisms, intellectual property rights, patents, copyright).

Table 1.

The first version of the GGBN data standard's Permit Vocabulary contains these terms for describing permit contents (the data standard uses the expression "label").

GGBN Permit Vocabulary: "labels" for permit contents	
Collecting Permit	Material Transfer Agreement
Contract	Memorandum of Cooperation
Copyright	Memorandum of Understanding
Data use	Other
Exemption Permit	Patent
Export Permit	Phytosanitary
Genetically Modified Organism	Salvage
Human Pathogens	Unknown
Import Permit	Veterinary Certificate
Internationally Recognized Certificate of Compliance	Intellectual Property Rights

We did not include the "GGBN data standard v1" Loan Vocabulary in our Biodiversity Permit/Loan Typology, as it only describes administrative elements of loans (date, loan identifier, availability of specimen…). In this way it differs from the GGBN Permit Vocabulary, which contains primarily different types of documents (but only a few of their

administrative elements). Anyway, loans are covered within the "MTA" Documents Types in this project report.

The **Society for the Preservation of Natural History Collections (SPNHC)** runs the website "SPNHC wiki" (SPNHC_Wiki_contributors 2022b), including topics on legislation and regulation, among them a webpage on "permitting" that started in 2016 (SPNHC_Wiki_contributors 2022a, main editor Breda Zimkus). The document categories from the SPNHC permitting webpage formed another cornerstone of our work (Suppl. material 7). These document categories are listed in the webpage section "Categories of Legal/Compliance Documentation" and were first generated by GGBN with a number of additions made by participants of a BCoN-funded workshop "Addressing Legal Issues Involved in Digitized Collections: The Nagoya Protocol as a Test Case" (NSF grant DBI #1441785) held at Harvard University in March 2018. (SPNHC_Wiki_contributors 2022a)

Participants of this workshop used compliance with the Nagoya Protocol[*1] to investigate how US institutions must respond to the need for increased transparency of their biodiversity collections and the required digital tracking. The group recognized the need for standardised definitions across the entire community, nationally and internationally, and they identified a number of permits not included in the GGBN list during working group discussions. Harvard's Museum of Comparative Zoology applied the discussion generated from the workshop and created a controlled vocabulary of Document Categories and specific Document Types (Zimkus et al. 2021).

4.2 Not covered: human tissue & pathogens, GMO, IPR, high granularity

Limited resources led to our decision to exclude the following topics from developing these biodiversity permit/contract typologies: Human tissue & pathogens, genetically modified organisms (GMO), biosafety and biosecurity issues (e.g. Belgian Science Policy Office 2023) and intellectual property rights (IPR). These are extensively regulated topics requiring considerable engagement of legal counsel, at the same time they are not core objectives of biological collections participating in our European Union funded SYNTHESYS+ project.

We also decided not to cover certain topics in a high granularity, e.g. country-specific permits and legislation, such as Australian permits. No member of our group had the necessary comprehensive knowledge of Australian biodiversity legislation (Suppl. material 3). Another example for our decision to dismiss high granularity in featuring permit types is the available information on phytosanitary[*1] requirements: The Food and Agriculture Organization of the United Nations and International Plant Protection Convention Secretariat 2023 lists approximately 700 different documents on that topic, similarly the USDA APHIS United States of America Department of Agriculture Animal and Plant Health Inspection Service 2023 lists approximately 80 different commodity import and export manuals, plus many other manuals on different related topics.

4.3 Creating Document Categories & Document Types, describing Terms

From the sources mentioned in chapter 4.1 we compiled an initial list of 16 Document Categories, intuitively named according to keywords of documents or with seemingly fitting short names (chapter 6.2). In an iterative process of 22 biweekly to monthly meetings we discussed their meaning and what documents they may contain, we adopted their names, devised definitions for them, reviewed them several times, and in this process reduced their number to a final set of 7 (chapter 6.2). In parallel we did the same for the subordinate Document Types (chapter 6.3 and Suppl. material 1).

Creating our collection of Document Types started in June 2021 with an existing list of documents, and with Document Types from the database "MCZbase" of the Museum of Comparative Zoology (Harvard University, Cambridge, MA, USA) provided by Breda Zimkus. In addition, the documents on this list were compared to similar documents from different jurisdictions stored in the other institutions participating in our project. Furthermore, participants added new documents to the list. In cases where no similar documents were available in our institutions, participants consulted other colleagues or checked either European or USA legal information to see if similar documents exist. In the process of researching and comparing documents, new Document Types were created, the names of existing Document Types were changed, a definition was added to each Document Type, and Document Types were assigned to the superior Document Categories.

In the 9th meeting participants decided to separate the description of documents from the description of terms applying to any work with biodiversity specimens. A second and different set of names and descriptions was started, specifically applying on one hand to general terms and on the other hand to very common specific terms extracted from documents (and some that are specific to collection management, e.g. insufficient digitisation).

For the specific terms for collection objects (No. 51-76 in Suppl. material 2) we additionally included publicly available permits associated to Internationally Recognised Certificates of Compliance from Kenya, Peru and India. They have the following ABS Clearing House*[1] Unique Identifiers & reference numbers of the national permit: ABSCH-IRCC-KE-242932-1 & NEMA/AGR/68/2017 (Kenya National Environment Management Authority 2019), ABSCH-IRCC-KE-253235-1 & NEMA/AGR/104/2018 (Kenya National Environment Management Authority 2020b), ABSCH-IRCC-KE-253313-1 & NEMA/AGR/135/2019 (Kenya National Environment Management Authority 2020a), ABSCH-IRCC-PE-249103-1 & 0013PER/SERFOR-2020 (Peru Servicio Nacional Forestal y de Fauna Silvestre 2020a), ABSCH-IRCC-PE-249106-1 & 0016PER/SERFOR-2020 (Peru Servicio Nacional Forestal y de Fauna Silvestre 2020b), ABSCH-IRCC-PE-254865-1 & 00026PER/SERFOR-2021 (Peru Servicio Nacional Forestal y de Fauna Silvestre 2021), ABSCH-IRCC-IN-255561-1 & India/NBA/Appl/9/4124 (India National Biodiversity Authority 2021).

4.4 Questionnaire on using molecular biological collections

To find out critical points in workflow and <u>Nagoya Protocol</u>*[1] compliance with outgoing material requests for DNA studies, and to find out how to overcome insufficient documentation of third party material and ensure good scientific practice with incoming loans, a joint (with SYNTHESYS+ task group T 3.2) questionnaire with 32 questions was sent out to all institutions involved in both tasks on September 4th, 2020. The goal for this survey was to get a first overview on how molecular biological lab data, permit documents, and sampling requests are logged. A distinction between the usage of institutions' own molecular biological collections by internal and external scientists versus using third party molecular biological collections (and receiving third party material for own research) was made. The questions in the questionnaire concerned the handling of these collections of sampled tissue, and the handling of the corresponding permits. This was followed by personal interviews at some institutions. 18 SYNTHESYS+ partners of Networking Activity 3 (NA3) participated and 17 completed this survey: BGM, BGBM, CSIC, HNHM, HUJI, LUOMUS, MfN, MNHN, NHM, NHMW, NRM, SGN, SMNS, UGOT, RBGE, RBGK, UCPH, and ZFMK (see acronym list at the beginning). From the 17 institutions who filled the questionnaire, 8 institutions had a follow up interview.

4.5 Joint MOBILISE-SYNTHESYS+ workshop to explore a cross-community approach

A joint virtual workshop organised by the COST Action MOBILISE (COST Association AISBL 2018) within its Working Group 3 and by SYNTHESYS+ NA3 on "A Loans and Permits Data Standard for Scientific Collections" took place on September 29th and 30th 2021. In total 105 people from 27 countries attended the workshop.

In this workshop, input was sought from various collection communities regarding their expertise, experiences, needs and challenges in connection with a data standard for loans and permits, with the long-term goal to develop a data standard that can be jointly used by different communities. In the workshop, the implementation of the standard in infrastructures and portals was discussed from a general point of view of what shall be achieved and how.

The workshop contributed to an ongoing cross-community discussion on how to support the adequate handling of legal and ethical requirements in natural history collections and related digital data infrastructures. In this context, the community consultation on converging Digital Specimens and Extended Specimens (GBIF.org 2021), especially part 8 on meeting legal/regulatory, ethical and sensitive data obligations (Hardisty et al. 2021) provided an excellent approach to the topic.

Digital Specimens and Extended Specimens (Webster 2017, Schindel and Cook 2018, Lendemer et al. 2020a, Lendemer et al. 2020b) have slightly different definitions: The DS "represents the sum of information on the Internet about a natural specimen object. The Digital Specimen acts as a processable digital twin on the Internet for the physical specimen in a natural sciences collection" (Hardisty 2020 of Nov. 17th 2020, see a slightly

different definition in Hardisty et al. 2020 of May 18th). It serves as a fundamental element of the European research infrastructure DiSSCo (DiSSCoTech 2020). The "extended specimen" is "a constellation of specimen preparations and data types that, together, capture the broader multidimensional phenotype of an individual, as well as the underlying genotype and biological community context from which they were sampled." (Webster 2017). Combined into a "Digital Extended Specimen" its definition is "the collective representation on the Internet of all digital assets referring to a physical specimen (which can include physical evidence of related observations), that meets the FAIR principles and that is distinguished and linked using globally unique persistent and resolvable identifiers to create an extensive online network of knowledge regarding life and related natural science objects." (Hardisty et al. 2022).

As a starting point, four different communities were invited, representing the main collections of natural history collections: biological collections, palaeontological collections, geological collections, and anthropological collections. DiSSCo and GGBN as data providing infrastructures were also present. Breakout sessions took place corresponding to the different disciplines: Biology Earth Sciences and palaeontology, and anthropology . The fourth breakout session dealt with implementation and infrastructures

5 Community needs and use cases

The ultimate goal of this work and the revision of the GGBN data standard is to create transparency and foster continuity of legal terms and conditions associated with specimens and their linked collection data throughout workflows and different use scenarios. This general goal can be broken down into a number of well-defined use cases, i.e. typical activities associated with the use of the specimens or data, which may come with different levels of complexity and quality requirements regarding e.g. the parties involved, the visibility of the activity, and the risks that non-compliance could entail. Use cases help with defining the scope of a task. In SYNTHESYS+ Task 3.3 the partners have looked into the use cases "outgoing and incoming loans", and assessed the current status and future requirements. In the joint MOBILISE-SYNTHESYS+ workshop on a loans and permits data standard for scientific collections the topic was presented to and discussed with over 100 participants representing the wider community of natural science collections. In this workshop, we looked into the use cases, policies and requirements of different disciplines.

Although the SYNTHESYS+ work package NA3 deals with molecular biological collections predominantly, the standardisation of legal information, which is the objective of the subordinate task 3.3, cannot be restricted to this subgroup of collections - the idea is to foster continuity of associated legal information, and this extends to all materials associated with a molecular biological sample, such as a voucher specimen.

Depending on the scientific discipline there may be very different prerequisites and requirements for the respective collections regarding e.g. the type of material, the legal requirements and other potential restrictions regarding their use, and different best practices followed in the community. This interdisciplinary scope was explored jointly by the

MOBILISE and SYNTHESYS+ working groups. In the long run, an updated GGBN data standard should be powerful enough to serve all scientific communities using natural history collections. Starting out, the project focused on non-human biological specimens and samples, with the option to extend the standard in the future. The present section 5 describes the results of the community consultations and derives typical situations in need of the standard.

5.1 Using molecular biological collections: Survey results

Questions from the joint questionnaire developed in collaboration with SYNTHESYS+ task 3.2 on MTAs (see Suppl. material 10) was one source of information used to get insight in common practice on the use of molecular collections. Outgoing loans for DNA investigations by external and internal researchers are already documented well across the partners, but associated lab data are mostly not linked properly. Associated MTAs, permits and other legal documents are stored centrally in 11 partner institutions, in six institutions they are stored by individual researchers or staff and one organisation doesn't store them at all. 10 partners store the documents in both printed and digital form, seven store them digitally only.

For the question what kind of MTA the institutions employ for sending out samples to be used for molecular biological analysis four partners use the CETAF MTAs, seven partners use individual MTAs based on CETAF or GGBN, four partners use individual MTAs, and two partners have either no MTA in place or gave no clear answer.

When requesting material from third party institutions - incoming loans - eight partners store associated MTAs, permits and other legal documents centrally, at seven partners individual researchers or staff are responsible for their storage, five partners store them both in printed and digital form, while seven partners keep them only in digital format. Some organisations mentioned that many documents are not reported to the management at all.

The question if good policies are in place for documenting incoming loan requests for genetic analysis was answered positively by 12 partners, five partners said no without specifying reasons. The majority of respondents agreed that it would be useful to establish organisation-wide policies for incoming loans, though stated that it often would be very difficult to implement these due to a lack of resources.

As a general result of this questionnaire, workflows for handling voucher specimens and extracted DNA, as well as for maintaining related documentation vary widely and are often not institutionalised. The need for standardisation in the handling of documentation is evident for DNA tissue material transferred both by out- and ingoing loans.

5.2 Use cases in the wider community of natural science collections

This chapter is a summary of the still unpublished detailed report of the joint MOBILISE WG3 and SYNTHESYS+ NA3 -Workshop on a loans and permits data standard for scientific collections (see chapter 4.5).

The workshop explored differences and similarities between the different disciplines of a natural history museum: biology, mineralogy and palaeontology as well as anthropology. The workshop went beyond the scope of SYNTHESYS+ NA3.3 (with a focus on biological collections) to introduce the concept of a loans and permits data standard to the wider community of scientific collections. The main conclusion was that a standardised vocabulary for permits and regulations, in biology for example implementing national ABS legislation arising from the Nagoya Protocol*[1], is relevant for all these disciplines and that this cross-disciplinary approach should be continued in the future. We decided to summarise the comprehensive results of the workshop here, but emphasise that non-biological collections are out of the scope of the current project report. The plan for a next step is to widen the scope again by setting up a TDWG Task Group for continuing the present work and transforming it into a TDWG standard for "Permits & Transactions" (working title).

The goals set out for the joint MOBILISE-SYNTHESYS+ workshop were:

- bringing together different collection communities

- discussing the need and possibilities for a data standard to share permit and loan information in different disciplines

- working on a set of minimum information required to provide permit/loan data

- coming to a decision that it makes sense to include all the invited disciplines

- (finding volunteers to form a working group for the data standard and agreeing on next steps, not relevant for this project report)

Within the scope of this workshop were Natural History Museums, Botanic Gardens including seed banks, Culture Collections, Technical Infrastructures and associated networks like GBIF, GGBN, DiSSCo or iDigBio.

Libraries, medical collections, related biological collections (agriculture and forestry collections, veterinary collections, zoos/aquaria, virus collections) were out of scope for this workshop, but could be added at a later point.

Regarding permit management, linking permits to samples or specimen data is straightforward in most cases, although there is no controlled vocabulary for permits. Post-hoc searching for applicable legal, ethical, and confidentiality restrictions for specimens is a time-consuming task that is currently done manually.

To explore the requirements of the different disciplines that are holding natural science collections, typical use cases and policies of each discipline were collated to form a starting point for a controlled vocabulary for the contents of legal documents. In breakout sessions characteristics and legal issues of different collection types were addressed and discussed.

5.2.1 Framework for use cases from biological collections

Typical objects in biological collections are

- preserved specimens, living specimens in botanic gardens or culture collections, DNA and tissue banks, seed banks
- observation records, collected by professionals or citizen scientists, play an important role

Often these collection objects are closely related to each other (e.g. seeds collected from one or more plants growing in a botanic garden, DNA sampled from a preserved specimen, preserved specimens as vouchers of living collections), resulting in complex data relations.

Use cases that are regulated by permits and policies:

- As a researcher I want to know who is managing specimens in a collection and which associated documents exist.
- As a curator/collection manager, I would like to know which collection items are subject to ABS regulation.
- As a researcher I want to know which analyses are possible and under which conditions or with which obligations, and whether subsampling or destructive sampling is allowed.
- As a researcher or exhibition manager I want to know if the loan of a specimen is possible.
- As a curator I want to know if the exchange of a specific object is possible.
- As a researcher I want to know how I have to reference the collection objects in publications.
- As a researcher or an enterprise, I want to know if specific objects are available for bioprospecting & commercialization.

Regulatory Frameworks:

- Convention on Biological Diversity*[1] and its Nagoya Protocol*[1] (CBD and NP - apply to countries only, national laws oblige their residents)
- Convention on International Trade in Endangered Species of Wild Fauna and Flora*[1] (CITES - applies to countries only, national laws oblige their residents)
- Protected areas (laws on different levels of legislation - e.g. national, subnational)
- Protected species (laws on different levels of legislation - e.g. national, subnational)
- Laws to prevent the spread of pests and diseases (supranational/EU, national)
- Protection of cultural heritage

Handling **sensitive data** applies to:

- Occurrences of endangered species subject to illegal hunting/collecting
- Data indicating commercial potential (e.g. ethnobotany, traditional knowledge)

Handling **sensitive specimens** applies to:

- Type specimens
- Very old, rare, or susceptible specimens

Best practices for publishing information on biological specimens:

- Collaborate and publish together with scientists from provider countries
- Use stable specimen identifiers (such as CETAF stable identifiers) for citation
- Deposit voucher specimens in official and accessible collection-holding institutions
- Cite specimens and samples both in publications and in INSDC databases

Existing **international data infrastructures** (selection, with a focus on specimens):

- Global Biodiversity Information Facility (GBIF)
- International Nucleotide Sequence Database Collaboration (INSDC)
- Global Genome Biodiversity Network (GGBN)
- Distributed System of Scientific Collections (DiSSCo; under development)
- Integrated Digitized Biocollections (iDigBio)
- Atlas of Living Australia (ALA)
- JACQ (herbaria)

Open Source **Collection Management Systems** (selection):

- Specify (Biological and Earth Sciences)
- DINA
- Diversity Workbench
- JACQ (herbaria)
- Arctos
- Symbiota

5.2.2 Framework for use cases from the Earth Sciences

Typical objects in Earth sciences collections are:

- Fossils, rocks, minerals, meteorites, hydrocarbons, gems, models (e.g. physical copies, digital tomography data), moulds (natural casts & embedding made of e.g. wax, gypsum, latex, silicone rubber, gutta percha, epoxy resin), analytical samples (e.g. thin sections, polished sections, SEM stubs, acetate peels)

- observation records are very rare

Some features must be considered as they add value and may affect associated insurance and transport costs:

- large and heavy

- hazardous (toxic, radioactive or asbestiform)

- liquid

- need to be pure (chemically uncontaminated)

- delicate

- brittle

- valuable

- rare (type specimens, from sites that are now depleted or inaccessible)

Use cases that are regulated by permits and policies:

- As a curator I want to know how a specific hazardous specimen has to be handled and stored.
- As a researcher/curator/collection manager I want to know under which conditions a specimen becomes toxic (soluble in acid, soluble in water), poisonous to ingest, or sensitive to temperature and humidity changes -e.g. pyrite decay, sulphur-, sensitive to light e.g. amber, fluorite.

Regulatory Frameworks:

- no international framework exists (to the breakout group's knowledge)

- varies by country, state, county, national park as well as time of collection, size and value

- National or local legislations can vary a lot e.g.

    ◦ Specimens from "sites of special interest (SSSI)" – UK

    ◦ Mines Royal – Scotland & UK (Queen claims ownership of certain metals, fish and birds)

    ◦ Danekræ – fossils are of national importance in Denmark

    ◦ Meteorite – treated differently in every country, land where it falls has ownership

- Museum policies and curators in charge can add additional rules

- Protection of cultural goods

Handling **sensitive data/sensitive specimens** applies to:

- Some locality information may potentially be targeted by (commercial) hunters

Best Practices for publishing:

- Some journals ask for permit descriptions, or a disclaimer that no permits were required
- Types and figured specimens typically require registration numbers
- Stable specimen identifiers (such as CETAF stable identifiers) are not yet established well enough, but should be used in the future

Existing **international data infrastructures** (selection, with a focus on specimens):

- Palaeontology

    ◦ Global Biodiversity Information Facility (GBIF)

    ◦ Integrated Digitized Biocollections (iDigBio)

    ◦ Atlas of Living Australia (ALA)

    ◦ Palaeobiology Database (PBDB, focus on publications)

    ◦ Geobiodiversity Database & Neptune (focus on publications)

- All Earth sciences

    ◦ GeoCASe

- Various national portals are in the pipeline, e.g.

    ◦ SwissCollNet

    ◦ DiSSCo UK

    ◦ Swedish Biodiversity Data Infrastructure (SBDI)

Open Source **Collection Management Systems** (selection):

- Specify (Biological and Earth Sciences)
- Diversity Workbench
- Mostly employed CMS are not open source (Adlib, Axiell, Filemaker etc.)

5.2.3 Framework for use cases from anthropology

Typical objects are:

- Ethnographic artefacts – Cultural objects (including those made e.g. from biological or mineralogical specimens)

- Prehistoric artefacts – Cultural objects
- Human remains – no cultural objects (with exceptions)

Only human remains as collection items were addressed in the following discussion on use cases, regulatory frameworks, sensitive objects, best practices, data infrastructures and open source collection management systems.

Use cases that are regulated by permits and policies:

Different approaches and regulations are in place in different institutions; no standard exists.

- As an exhibition manager I want to know if a specific object is available for loan to be exhibited.
- As a researcher I want to know if a scientific investigation is allowed.
- As a researcher I want to know if destructive sampling is possible.
- As a curator/researcher/collection manager I want to know which specimens are tagged as sensitive or/and have sensitive data (human remains).

Regulatory Frameworks:

- ICOM code of ethics for museums (International Council of Museums 2017)
- Guidelines Care of Human Remains in Museums and Collections (German Museums Association 2021)
- Guidance on Archaeological Measures (Austria, Federal Ministry for Arts Culture the Civil Service and Sport, Bundesdenkmalamt 2023)
- Guidance for the Care of Human Remains in Museums (United Kingdom, His Majesty's Government, Department for Culture Media and Sports 2005)

Many objects from anthropological collections are declared as **sensitive objects** after evaluation of different aspects:

- fossils
- unknown provenance
- sensitive context: colonial context, National Socialism (NS-) context, illegal collection, tribal concerns

Best Practices for publishing:

- Some journals ask for permit descriptions, and/or a disclaimer that no ethical issues are present.
- Pictured objects usually require identifiers in the caption. (Stable specimen identifiers, such as CETAF stable identifiers, should be used in the future).

At present many **internationally available data repositories** are used for storing public shareable anthropological data like GitHub (Wikipedia contributors 2023a), MorphoSource (Duke University Library Digital Repository Software Development and Integration Service 2023), Zenodo (European Organization for Nuclear Research and OpenAIRE 2013), tDAR

(Center for Digital Antiquity 2023) , [Human Fossil Teeth](#) (Voisin et al. 2021); [ki Global Health](#) (Bill & Melinda Gates Foundation 2022); [Stony Brook Primate Locomotion Lab](#) (Stony Brook Primate Locomotion Laboratory 2022); [Primatelocomotion](#) (Stony Brook Primate Locomotion Laboratory 2022); [Ahob-Project](#) (The Ancient Human Occupation of Britain Project 2016); Mulligan et al. 2022, 36.

Furthermore, many institutional portals may be in place, e.g. [THANADOS](#) (Eichert and Richards 2022). At present no Open Source **Collection Management Systems** for Anthropological Collections are in place.

5.3 Synthesis of use cases

Taking together the results of the surveys about managing and using molecular biological collections and of the consultation in the wider community of natural science collections, we derived four levels of activities (transactions) with increasing complexity: 1. Accessioning and (in-house) documentation, 2. Loans and exchange, 3. Use and 4. Submission and publication. With these four levels, the degree of interconnection and visibility of objects and data increases. Accordingly, there is an increasing need to provide standardised information on legal obligations that is transmitted with each step.

5.3.1 Accessioning and in-house documentation

The management of objects and information in collection-holding institutions is the first level of complexity when it comes to documenting and managing legal information. All the collection-holding institutions participating in our endeavour have policies in place for accessioning material, and workflows for managing associated information. The Policies Handbook on Using Molecular Collections, developed by the Synthesys+ project's task group NA3.2, lists best practices for accessioning samples and for managing legal documents related to specimens and samples. The handbook points out the challenge to "record rights, restrictions and obligations related information in a standardised way and unambiguously attribute it to the collection items it refers to" (de Mestier et al. 2023). The Biodiversity Permit/Contract Typology, together with the Typology of Legal/Contractual Terms for Biodiversity Specimens presented here, addresses this challenge by providing the possibility to update the GGBN data standard to form a tool to link legal information to specimens and their derivatives in collection management systems. Such a stable link is a prerequisite for keeping the information through all subsequent transaction levels. Full implementation of the vocabulary is in most cases not a requirement at this first, in-house level in order to achieve coverage of the legal information applying to the objects within an institutional collection.

5.3.2 Loans and exchange

From the perspective of an institution, sharing its resources with other parties constitutes the next level of complexity and responsibility. Each institution that was involved in the current project has a loan policy in place that governs the terms and conditions under

which a specimen or sample is available for loan. That loan policy forms the basis for the material transfer agreement. It is in the vital interest of the institution to include and transfer all relevant legal information that is attached to a specimen sent out for loan. The standardised vocabulary makes it easier for the providing party to apply due diligence and to define the terms and conditions under which the specimen is offered for loan. The receiving party is supported in ensuring compliance and in establishing its own management plan of the received specimen/sample and the related legal information.

5.3.3 Use

A number of activities may be performed on specimens/samples apart from simply holding them in the collection, leading to e.g. identification, exhibition, research, development or commercialisation and may include subsampling and different methods of analysis. Use of the material usually leads to the generation of benefits and is thus the main subject of ABS (Access and Benefit Sharing[*1]) regulations. Independently of whether the material falls under the Nagoya Protocol[*1] or not, collections and users have a vital interest to ensure that the terms for using the material are transparent and complied with. The Biodiversity Permit/Contract Typology and the Typology of Legal/Contractual Terms for Biodiversity Specimens help to flag possible use rules, that is, permissions, prohibitions, restrictions and duties in databases and repositories. In that way, they facilitate locating the sources for these regulations, which may otherwise be hidden in multiple, hard-to-read documents.

5.3.4 Publication and submission to public repositories, databases and infrastructures

The last level of complexity the management of legal information must master is the publication, display or offer of the material and associated data in publicly accessible resources. This may be a public database for specimen information (including institutional collection databases and portals, and data aggregators such as GBIF, GGBN or DiSSCo), a database for research results (as for example the INSDC sequence databases or trait databases), a public repository for reference material, or a scientific journal. This step of opening up data and information about material objects and associated knowledge is essential for openly providing biodiversity knowledge for science and societal applications (see chapter 3.1). At the same time this step can be considered to be most sensitive, specifically for highly interlinked data and digital objects, if legal and ethical requirements have to be fulfilled. The Biodiversity Permit/Contract Typology and the Typology of Legal/ Contractual Terms for Biodiversity Specimens help to create transparency on information related to rights, restrictions and obligations, fosters trust on the side of the providing parties and can contribute to legal certainty of users. Infrastructures become increasingly aware of the responsibility that arises when unprecedented use can be made of their assets with the emergence of semantic technologies, artificial intelligence and advanced data analysis tools. By implementing the data standard, infrastructures take care that legal information is visible to users and providers. Implementing the standard is no guarantee that information on rights, restrictions and obligations is correct or complete. It is still the responsibility of the user to apply due diligence and to take every measure to make sure

that all legal requirements are complied with. Yet, it is an important step towards more transparency and good scientific practice, and can establish increased levels of awareness and public control. The latter are connected with the hope that more biodiversity data are made openly available when there are better and more effective measures in place to respect and comply with legal obligations.

5.4 Needs for implementation and infrastructure development

The following aspects were highlighted during the fourth breakout session of the MOBILISE-SYNTHESYS+ workshop with members of the biodiversity informatics community engaged in developing, implementing and maintaining biodiversity data infrastructures:

- The standardisation and harmonisation of legal and contractual information associated with physical specimens, their digital representations, as well as derived and associated data is considered to be much needed and to provide important advantages for future collections-based work.
- Accessibility of legal information is of importance for assessing compliance with, e.g., the Nagoya Protocol[*1].
- Legal conditions apply to a wide range of different transactions, both analogue and digital, including e.g. loans, gifts, accessions.
- Digital specimens have an entirely different legal background as the physical museums' objects. Connecting both has more dimensions than just "linking permits". The relation between legal conditions associated with a specimen and its digital twin needs to be explored and clarified.
- The meaning of "open data" can be very differently understood, and thus be prone to misunderstanding and contention among different sectors and among different rights- and stakeholders.
- Liability issues may arise from transcripts of "closed" contractual agreements to a standardised set of terms (cp. the Typology of Legal/Contractual Terms).
- There are often legal and linguistic difficulties in transcribing and translating legal documents
- Time stamps are needed since the legal landscape is dynamic and laws change.
- Quality flags are required, e.g. in the form of "no permit present" in the absence of permit information.
- Persistent linkage to data derived from a specimen, e.g. during a loan event, is needed to ensure that legal information is appropriately propagated.
- Infrastructure providers need to solve long-term archiving.
- The different types and levels of legal structures need to be represented. For example, the Nagoya Protocol[*1] applies to countries, which then mint national laws, which form the basis for contracts between e.g. providers and users.

The information and experiences provided by infrastructure providers and users in this breakout group became part of the foundation of the work of the SYNTHESYS+ working

group and directly influenced the development of the typologies as well as the implementation model.

6 The Biodiversity Permit/Contract Typology

The Biodiversity Permit/Contract Typology provides a common standard terminology with names and semantic definitions for biodiversity-related permits, as well as contracts. It helps to overcome one of the first obstacles for dealing with these documents: unfamiliar permit/contract titles from foreign jurisdictions have the effect that their significance cannot be easily understood, and familiar permit/contract titles potentially mislead the reader to expect content that is not included in the present document.

In addition, an automatised, swift exchange of information on permits/contracts will be a valuable service that such a standardised typology of permits/contracts potentially facilitates. Likewise, a typology makes sorting and filing documents much easier, be it in physical filing cabinets or in databases.

A common rulebook exists at international level only for a very narrow aspect of configuring biodiversity-related permits and contracts, i.e. for the Convention on International Trade in Endangered Species of Wild Fauna and Flora 1973, that defined a standard for permits in Article VI. The Conference of the Parties to the Convention on International Trade in Endangered Species of Wild Fauna and Flora 2002 further explained this standard.

6.1 A two-tiered system of Document Categories and Document Types for permits and contracts

We collected more than 40 different permit and contract document types for expanding the GGBN permit/loan vocabulary into this Biodiversity Permit/Contract Typology. We selected a two-tiered system where we grouped similar documents (e.g. collecting permits from different countries) in one Document Type (e.g. the Document Type "Collecting Permit"). Then, similar Document Types (e.g. "Collecting Permit", "Salvage permit", "Incidental Take Permit") were pooled to form one Document Category (e.g. the Document Category "Permits for Collecting & Related/Taking/Possessing").

Apart from higher clarity, our two-tiered approach offers additional advantages:

- thematic clusters – Document Categories are grouping permit/contract documents corresponding to the main processes related to biodiversity collection objects, all but one of these processes (and document categories respectively) are repeatable, only "Collecting & Related/Taking/Possessing" represents the first phase of a collection object's life-cycle and is not repeatable for this specific collection object. At the same time most Document Categories contain either only permits or only contracts, while two categories comprise both (permits and contracts are mixed in the document categories ABS and Transport Documents).

- **flexible granularity** – For filing permits and contracts, institutions can choose whether they only use the less granular Document Categories and skip the more granular Document Types classification (e.g. for subject matters where they have few and very different documents), or they use both Document Categories and Types (e.g. for subject matters with a large number of different documents), or whether they use only a specific Document Type without associated Document Category (e.g. for subject matters where they have few very similar documents).

- **graded specificity** – the two-tiered approach of Document Categories and Document Types constitutes two levels of specificity that may help with filing documents. Sometimes it is too time-consuming to determine the correct Document Type for a given document at once, then it may be a compromise solution to allocate the document preliminarily to the (less specific) document category.

6.2 A set of 7 Document Categories

The group established a set of seven Document Categories (DC1-DC7) after a process of classification of existing documents. The process started with 16 clusters of documents that we initially proposed for Document Categories (see chapter 4.1) . Only two of these clusters remained unchanged throughout the classification process, i.e. the Access & Benefit Sharing Documents, and the Permits for Research. Seven clusters were combined into two Document Categories (i.e. "Permits to Collect and Take/Possess" and "Transportation Documents"). These seven may be an indication of how numerous related documents are in our natural history collections, and how highly their importance is estimated. Another three clusters got new names during the classification process: The Category "Agreement Document" became "High-Level Arrangements", "Use permission" changed to "Permits for special purposes", and "Transfer of ownership" to "material transfer agreements, stewardships & ownership-related information". Two initial clusters for Document Categories were transferred to the more granular Document Types. These are on the one hand "Exemption Permit," which we renamed to "Exemption Certificate" and added to all Document Categories for which it was an option based on our experience. On the other hand, we relocated the "Ethical oversight Document" into the "Permits for research" Category. Moreover, the initial cluster for a Document Category with the name "Institutional Permit" was dissolved and the contents (CITES Certificate for Scientific Exchange, US Endangered Species Act Museum permit, Memorandum of Understanding) were moved to the Document Categories "Transportation Documents" and "High Level Arrangements". Finally, the initial cluster for a Document Category with the name "Ownership Document" was excluded because it contained intellectual property rights, and we did not have the resources and expertise to address this topic. An overview of these changes is provided in Table 2.

We realised that our initial drafts for the Biodiversity Permit/Contract Typology were biased by the high number of documents stored in our biodiversity collections for taking objects from nature and transporting them. In the end we needed only two Document Categories for them, and a total number of seven Document Categories.

The Document Category definitions describe the scope of each Category and which Document Types it includes.

Table 2.

Development of the 16 initially used Document Categories (numbers in brackets refer to the final Document Categories-DC) and where applicable the type of change: merged, modified, dissolved, excluded - or downgraded to a Document Type - DT.

Initial Document Categories (DC) and their development

initial DC	decision	initial DC	decision
Access and Benefit-Sharing Document	DC1	Institutional Permission	dissolved
Agreement Document	modified (DC2)	Export Permission	merged (DC7)
Authorization to Possess	merged (DC3)	Import Permission	merged (DC7)
Collecting/Take Permission	merged (DC3)	Ownership Document	excluded
Use Permission	modified (DC4)	Receiving Permission	merged (DC3)
Research Permission	DC5	Salvage Permission	merged (DC3)
Ethical Oversight Document	downgraded (DT)	Transfer of Ownership	modified (DC6)
Exemption Permission	downgraded (DT)	Transport Document	merged (DC7)

6.2.1 Access and Benefit-Sharing Document (ABS)

Document Category definition: *documents permitting the* <u>access to genetic resources</u>*[1] for their utilisation and/or covering the terms of benefit-sharing*

6.2.2 High level arrangements

Document Category definition: *documents that require signature(s) by (a) member(s) of the highest management level of an institution. There may be country-specific arrangements at even higher levels, or higher-level arrangements applicable only to institutions being a government department*

6.2.3 Permits for collecting & related/taking/possessing

Document Category definition: *documents by authorities or private entities allowing collecting live specimens from nature (and entering certain areas) or taking e.g. roadkill, as well as possessing restricted material - this does not include ABS-permits, research permits and permits related to the transport of specimens (see the respective applicable document categories)*

6.2.4 Permits for special purposes (excluding ABS)

Document Category definition: *documents by authorities or private entities allowing use of the collection (objects) for specific, limited purposes (future perspective) - this does not include ABS-permits*

6.2.5 Permits for Research

Document Category definition: *document by an authority allowing basic and/or applied research within its jurisdiction, and legally required committee decisions on research*

6.2.6 Material transfer agreements, stewardships & ownership-related information (excluding ABS)

Document Category definition: *documents demonstrating the will of two parties regarding the transfer of tangible material between them (e.g., an agreement between two institutions that outlines the terms and conditions for transferring specimens or samples; for stewardships the conceding party is a public authority) or which include information on such transfers in the past. This does not include shipping documents, contracts with a carrier (both: see transport documents) or transfer documents under ABS-legislation*

6.2.7 Transport Documents

Document Category definition: *permits, certificates and other documents necessary for the act of sending specimens from one place to another*

6.3 A set of 38 Document Types

We created and allocated Document Types for each Document Category from the previous chapter (Table 3). Each Document Type is characterised by a definition. The 38 Document Types with their definitions, additional information about what a Document Type could include, and examples (e.g. links to publicly available blank forms) are listed in Suppl. material 1. Document Types serve two purposes: The main purpose was to identify and group documents relevant for future use of collection objects. For example, documents related to access to genetic resources & benefit-sharing[*1] based on their utilisation. The second was to create a comprehensive system for filing all documents in the life-cycle of a collection object, e.g. a permission to enter a collecting site by car.

The search for European documents similar to those from the USA quickly became time-consuming, as few are standardised at the European level. Most European documents result from national or even local legislation. As a by-product we compiled a list of permits issued by US or European authorities (completeness was not pursued, Suppl. material 8).

Creating definitions for each Document Type was very labour-intensive because they had to fit all documents we knew that belonged to that Document Type. We intended for them

to be short and concise, but at the same time contain enough information so that even non-experts could correctly assign the documents.

Table 3.

The final seven Document Categories and their 38 associated Document Types.

Document Categories	Corresponding Document Types
1) Access and Benefit-Sharing*[1] Document (ABS)	1-1 Mutually Agreed Terms (MAT) 1-2 Internationally Recognized Certificate of Compliance (IRCC) 1-3 Prior Informed Consent (PIC) 1-4 Specialised standard ABS terms - SMTA 1-5 Other ABS document (e.g. biomedical, BBNJ*[1]) 1-6 Exemption evidence
2) High level arrangements	2-1 Contract (legally binding) 2-2 Memorandum of Cooperation (MoC) 2-3 Memorandum of Understanding (MoU)
3) Permits for collecting & related/taking/possessing	3-1 Authorisation to enter site 3-2 Collecting permit 3-3 Taking: "incidental take" permit 3-4 Taking: Migratory Bird Treaty Act (MBTA) Special Purpose, Salvage Permit 3-5 Taking: Salvage Permit (e.g., Non-US, US federal, state, local) 3-6 Possessing: Receiving permit 3-7 Exemption evidence
4) Permits for special purposes (excluding ABS)	4-1 Permit to reintroduce*[1]/translocate*[1] organism into the wild 4-2 Data use agreement 4-3 Bioprospecting permit
5) Permits for Research	5-1 Research Permit 5-2 Ethical oversight document 5-3 Exemption evidence
6) Material Transfer Agreements, stewardships & ownership-related information (excluding ABS)	6-1 Public law MTA (e.g. acquiring customs' seizures, stewardship agreement) 6-2 Institutional MTA (e.g. loans) 6-3 Individual deeds of transfer (e.g. private gifts) 6-4 Provenance evidence
7) Transport Documents	7-1 (Phyto-)Sanitary/Veterinary Certificate 7-2 Permit to move across boundaries 7-3 CITES export permits & re-export certificates 7-4 CITES import permits 7-5 CITES certificates of scientific exchange (COSE) 7-6 other CITES documents 7-7 Original Export Permit 7-8 Export Permit 7-9 Original Import Permit 7-10 Import Permit 7-11 Exemption evidence 7-12 Other transport documents

7 Typology of Legal/Contractual Terms for Biodiversity Specimens

For the purpose of capturing (not only, but also) "restriction information" attached to biodiversity objects we initially planned to create a list of various biodiversity-related documents including their descriptions; and we wanted to do that by updating the GGBN data standards for the GGBN Permit Vocabulary and the GGBN Loan Vocabulary (Droege et al. 2016).

7.1 Addressing a problem: uniform use-terms are missing in permit vocabulary

Once we began updating core elements of the GGBN Permit Vocabulary and Loan Vocabulary, i.e. names and definitions for a list of different permits and contracts, it soon became clear that uniform names/definitions for similar documents from different jurisdictions do not automatically provide uniform "restriction information" or, more generally, uniform use-terms. Simply because we often could not infer accurate "restriction information" from the name of the permit/contract ("contract" also includes loans of biodiversity collection objects). For instance, two "Prior Informed Consent-PIC" documents for utilising genetic resources*[1] issued by different jurisdictions contained different restrictions, and also different duties. Both are the same type of permit, but containing different types of use-terms. It became clear that the lack of a comprehensive international standard for configuring biodiversity-related permits requires a twofold approach in addressing restrictions attached to biodiversity specimens. We had to create not only a Biodiversity Permit/Contract Typology for sorting the multitude of national permits/contracts, but we also had to complement it with another typology, the Typology of Legal/Contractual Terms for Biodiversity Specimens if we want to arrive at harmonised concepts for "restriction information".

7.2 Identifying a backbone: five types of lawful sources, their characteristics, plus expert opinions

Lawfully handling and using biodiversity collection objects, as well as creating harmonised concepts for delivering "restriction information", in general rests not on permits and contracts alone, there are additional lawful sources to be considered. A recent example is related to "utilisation of genetic resources"*[1], first defined in the international Nagoya Protocol*[1] and subsequently incorporated in several countries' national legal systems. Countries often issue individual permits for "utilisation of genetic resources"*[1] (provided the genetic resource*[1] was taken from that country). Other countries do not regulate their genetic resources, therefore they do not issue such permits and any "utilisation of their genetic resources"*[1] is permitted on the principle that anything is allowed unless it is explicitly forbidden. We identified five lawful sources as part of the backbone of our Typology of Legal/Contractual Terms for Biodiversity Specimens (7.2.1).

Documents related to objects in our collections do not only represent different lawful sources for restrictions on having and using these objects, they may also contain other characteristics than "restriction information", e.g. a permission. These characteristics are compiled further below (7.2.2). Some documents related to collection objects are no judicial texts at all, but opinions of experts in natural sciences (7.2.3).

7.2.1 Lawful sources

The question "How do you know?" that certain "restriction information" exists led us to five types of lawful sources setting conditions for handling biodiversity specimens:

1. an individual **document** issued to a specific person or legal entity, based on a public administration's decision (often after examination of a specific project; abbreviated "D")
2. a written **law** without individualised document, the law however can be referenced (abbreviated "L")
3. **no legal provisions** exist nationally or on the level of applicable supranational (e.g. EU) legislation (abbreviated "N")
4. **a contract/deed** (civil law; abbreviated "CD")
5. evidence that **no individual document** (see No. 1.) is necessary for a permission, but without specifying whether this is based on written law or missing legal provisions (abbreviated "LN")

These five types of lawful sources D, L, N, LN and CD also provide their specific degree of legal certainty. For example, any permission derived from an individually assigned document (or contract) pertaining to a specific collection object provides less room for interpretation than a permission derived from written law - because you need to be certain that the law really applies to the collection object in question. And finally, a permission inferred from the absence of specific legal provisions requires even more research or legal expertise if you want to have legal certainty.

7.2.2 Characteristics of lawful sources

"Restriction information", our starting point, is but one of four characteristics potentially contained in the lawful sources we identified. The lawful sources 1, 2 and 4 may contain all four legal characteristics listed below. We have taken the definitions for two of them, "prohibition" and "duty", from the Open Digital Rights Language ODRL (https://www.w3.org/TR/odrl-vocab/), and slightly adapted ODRL's wording for "permission". We have also added the new legal characteristic „restriction":

- Permission : The CONSENT to perform an action over an asset (we exchanged "ability" for "consent" to emphasise that a person gets the opportunity to perform an action, without testifying that the person is able to do so). *Our example: collecting specimens from nature* .

- Prohibition : "The inability to perform an Action over an Asset." *Our example: Do not damage the specimen's integrity*.

- Duty : "The obligation to perform an Action." *Our example: Deliver a copy of the publication*.

- Restriction : The obligation to refrain from an action for the time being. A restriction can be modified or lifted by subsequent negotiations. *Our example: Only for taxonomic purposes.*

Furthermore, we use "obligation" as an umbrella term for both Restriction and Duty (but not necessarily in the meaning it has in ODRL, where it is used in relation to "policies").

Depending on the jurisdiction or special legal context the lawful sources 3 and 5 express either "not prohibited", or "permitted", for the lawful source 5 the permission may also be based on a law.

7.2.3 Expert opinions

In the context of documents related to collection objects expert opinions that confirm a certain quality, such as "free of known pests", "free of biosecurity risks" or "ethical statement for a research project", also became part of our Typology of Legal/Contractual Terms for Biodiversity Specimens. The source of their content is expert knowledge only, therefore we do not classify them according to our lawful sources or legal characteristics (although we recognize that lawful sources may have initiated writing these expert opinions).

7.3 Describing terms for biodiversity specimens

Users need a specific set of information to work efficiently with terms attached to biodiversity specimens.

1. First of all, they need to know the contents of a term, or in the event of a standardised vocabulary for terms, the definition/description for every item in the vocabulary.
2. Then, most users want to get information on the legal certainty that comes with a specific term, they want to know how much room for interpretation exists, and here the lawful sources mentioned above may serve as first clues.
3. Third, users want to know whether a term potentially applies to them, or has already been fulfilled in the past.

Our Typology of Legal/Contractual Terms for Biodiversity Specimens includes these three elements - definition, lawful source (chapter 7.3.1) , past/future (chapter 7.3.2). The Typology is intended to provide a preliminary, first lead on legal issues, not to replace the necessary thorough legal counsel prior to handling/using biodiversity specimens. Users wanting to implement this Typology in EDP systems (Electronic Data Processing) might also be interested in a classification that specifies whether an item of the Typology is a

permission, prohibition, duty or restriction (or a combination of them) - for this purpose we provide a tabulated version of our Typology (chapter 7.3.3).

7.3.1 Indicating the lawful source for each legal/contractual term

The information on lawful sources that we provide for every term, implying a certain degree of legal certainty, is a judgement call by the authors based on our limited selection of available permits/contracts. The same applies for the information whether a term is only used by public authorities through administrative decisions or laws (e.g. for the term "import permission") or also in contracts under civil law (e.g. for the term "transfer to 3rd parties").

For implementing the Typology of Legal/Contractual Terms for Biodiversity Specimens in collections, or with international data providers we indicate the lawful source with the relevant capital letter CD (contract/deed), D (individual document), L (law), N (no legal provision), or LN (exemption evidence) or with "not specified", added to the catchphrase for each legal/contractual term. In any case, if no adequate personnel is available when implementing the Typology to specify the lawful sources of terms, the tag "not specified" can be applied for the time being.

> For example: For a given biodiversity specimen the legal term "YES: genetic and biochemical" may come from one of the following lawful sources: D, L, N, LN (or "not specified"). Depending on the contents of the lawful source at hand, a collection holding institution may flag the data record or label of this biodiversity specimen with "YES: genetic and biochemical (not specified)", or "YES: genetic and biochemical (D)", or "YES: genetic and biochemical (L)" etc. Very common will be "YES: genetic and biochemical (N)" for biodiversity specimens from countries that do not regulate their genetic resources[*1], which means that no legal restrictions are in place.

International platforms such as GGBN or the upcoming DiSSCo might want to implement the possibility to use any option from the example above in their systems to be ready for whichever information the participating institutions deliver. Collection holding institutions may want to use only selected lawful sources, e.g. "(D)" and "(CD) for terms based on individual permits or contracts/deeds, and use "not specified" for the rest (in this way discarding "(L)", "(N)" and "(LN)").

7.3.2 Legal/contractual terms for actions in the past and/or future

Some of these legal/contractual terms refer to future use (e.g. "YES: genetic and biochemical"), some only to actions in the past (e.g. "collecting from nature") and some to both (e.g. "under CITES"). When using the Typology to indicate that terms apply to "future use" and "actions in the past" we need a key for a consistent understanding of these two attributes, so that everybody draws the same line between a legal/contractual term that applies only to the past, only to future use or to both. For this purpose, we suggest to select alternately one of two definitions to delimit actions in the past. First, we define

"actions in the past" as any action prior to the end of the project/endeavour bringing a biodiversity object into a biodiversity collection for the first time. For example, if the biodiversity object is collected for a specific research project and then kept in a biodiversity collection, every legal/contractual term applying only to this research project is rated "applies to actions in the past". Any term applicable beyond this first research project gets the attribute "future use". For example, the legal/contractual term demanding to send a taxon list with counts of all collected specimens gets the attribute "past", because taxa are at least roughly determined before specimens are added to a collection (but have in mind, in case the legal/contractual terms are categorised at a very early stage of the research project, e.g. when digitising the documents/terms upon their arrival and therefore prior to the actual collection event, you indicate "actions in the past" although the taxon list has not been sent yet). As a second definition for delimiting actions in the past we suggest that terms always apply to "actions in the past" if they target a single specific point of time, e.g. the time of import into the country where the specimen is added to a collection.

The two attributes "actions in the past" and "future use" provide the following choices for institutions implementing the Typology: Some may choose to implement (e.g. in their collection management routines/systems) only such terms from the Typology that are relevant for "future use", while evaluating terms relevant to the "past" at a specific point of time, e.g. at the time when a biodiversity object is about to be added to the collection at the end of a research project or a collecting trip. If, in this example, the institution refrains from adding biodiversity specimens to the collection for which this evaluation shows that terms relevant to the "past" have not been fulfilled, the institution can provide certainty to staff and users that they do not need to deal with "past" terms (which remain accessible via the evaluation record). Other institutions may want to set up management systems for contracts they enter into, systems that record every single date of fulfilling a certain term and start monitoring all terms right after signing a contract. Or they want to set up management systems for projects they conduct, monitoring (among other project elements like budget) the contractual terms they have to fullfil. Then they may also implement terms the Typology addresses as "action in the past" to have the ability to check off all items that have already happened.

7.3.3 The legal characteristics of every term

For the implementation of the Typology of Legal/Contractual Terms for Biodiversity Specimens in EDP-systems we provide a table allocating every legal/contractual term we describe to the characteristics described in 7.2.2, differentiated by the lawful source from which it originates (Suppl. material 4).

The structure of the large table in Suppl. material 4 is explained in Table 4 showing the heading, and as an example row 12 with the catchphrase "obligations: genetic & biochemical" in the first column. The definition for this catchphrase is provided in Suppl. material 2, it also states that additional "institutional loan conditions may apply"- we provide typical examples for institutional loan conditions in rows 77-87 of the table in Suppl. material 4. The second column of Table 4 shows that the catchphrase in row 12 refers to

future actions with the biodiversity specimen in question, the third column shows that it is no mere expert opinion on the quality of a collection object (such as e.g. "free of familiar pests"). Columns 4-9 show that it is based on public law, and the lawful source may be either an individually assigned document (D) or a law (L) or, indicated by "possible" in the respective columns, a mixture of both. Additionally, they show the characteristics of the lawful source for the specific catchphrase, i.e. permission(s), duty(-ies) and/or restriction(s). The "n/a-not applicable" in the last column "required by contracts/deeds under civil law" indicates that this catchphrase refers only to public law. Loan conditions (e.g. from museums) or other contractual obligations, are covered by other catchphrases.

Table 4.

Explanation how to read Suppl. material 4: Legal characteristics of every term

The "catchphrases" are defined in Suppl. material 2; details on "action in the past/ future" see chapter 7.3.2; information on "expert opinion" see chapter 7.2.3; types of lawful sources are explained in chapter 7.2.1: D=individual document, L=law, N="not subject to national/ supranat. public law", CD= "required by contracts/ deeds under civil law"; characteristics of lawful sources see chapter 7.2.2; n/a=not applicable; possible=the relevant lawful source differs among countries.

Catchphrase ↓	action in the past/ future	expert opinion Y/n	under national (or supranational, e.g. EU) public law						(N) not subject to national/ supranat. public law	(CD) required by contracts/ deeds under civil law
			(D) refers to individually -			(L) refers to written law providing -				
			granted permission	assigned obligation (duty or restriction)	assigned prohibition	permission	obligation (duty or restriction)	prohibi-tion		
12 obligations: genetic & biochemical	future	n	possible	possible	n/a	possible	possible	n/a	n/a	n/a

7.4 The Typology: A set of 87 legal/contractual terms for biodiversity specimens

We identified 87 different legal/contractual terms and characterised each of them, based on documents that we used for setting up the Biodiversity Permit/Contract Typology.

In Table 5 we provide a list of all terms we identified and in Suppl. material 2 we explain these 87 terms in detail, along with their individual characterisation, their possible lawful source (see chapter 7.2.1 - or where applicable expert opinion, chapter 7.2.3) and whether they are relevant for future use, for actions in the past or for both (see chapter 7.3.2). The terms 1-50 are general terms, followed by (optional) specific terms 51-76, each starting

with the word "required", for adding more detail to the general terms. Term No. 77-87 are specific loan terms from contracts governing loans for academic research.

Table 5.

The Typology of legal/contractual terms for biodiversity specimens: we compiled 50 general terms and 37 specific terms from permits and contracts regulating the handling and use of biodiversity specimens (with a few additions from our experience in handling specimens in biodiversity collections, e.g. term #1 "info not (digitally) processed"). The Typology is intended to flexibly label biodiversity specimen records or even the specimens themselves: people in charge of biodiversity specimens may use only general terms, or only specific terms or a combination of both

General terms for biodiversity specimens				
1	Info not (digitally) processed	45		obligations: get ownership
2	no ties	46		stewardship
3	permission & free further use	47		entering permitted
4	free further use + reporting	48		equipment permitted
5	YES: exhibition	49		tagging permitted
6	obligations: exhibition	50		mode(s) permitted
7	NO: exhibition	**Specific terms for biodiversity specimens**		
8	YES: taxonomy	51		Required: for residents only
9	obligations: taxonomy	52		Required: visa
10	YES: physical analysis	53		Required: exit permit
11	YES: genetic & biochemical	54		Required: teaching
12	obligations: genetic & biochemical	55		Required: report progress
13	reporting: genetic & biochemical	56		Required: exit report
14	NO genetic & biochemical	57		Required: cash
15	Benefit-sharing-NP	58		Required: payment
16	Benefit-sharing-CBD	59		Required: send publication
17	Benefit-sharing-ITPGRFA	60		Required: inform public legal entity
18	YES: TK-utilisation	61		Required: send taxa names
19	obligations: TK-utilisation	62		Required: provider uID
20	YES: living organism research	63		Required: holding location
21	obligations: living organism research	64		Required: negotiate 3rd pary
22	NO living organism research	65		Required: work abroad
23	YES: 3rd party	66		Required: negotiate commercial
24	obligations: 3rd party	67		Required: non-commercial
25	NO 3rd pary	68		Required: no economic profit
26	any quality	69		Required: IPR notification
27	ethical research	70		Required: share sales
28	PIC-exempt	71		Required: method

29	MAT-exempt	72	Required: single research
30	pre-act exemption	73	Required: handle duplicates
31	pre-llisted exemption	74	Required: handle left-over
32	not-protected exemption	75	Required: return left-over
33	free of familiar pests	76	Required: handle „DSI"
34	free of biosecurity risks	77	NO destructive sampling
35	import approved	78	destructive sampling restricted
36	YES: collect material	79	dissecting restricted
37	obligations: collect material	80	NO shipping
38	YES: sensitive information	81	preparation restricted
39	export approved	82	share images
40	holding approved	83	NO microwave
41	under CITES	84	storage parameters
42	YES: move countrywide	85	NO sampling
43	YES: setting free	86	negotiate commercial use
44	YES: get ownership	87	not available

The table "Legal characteristics of every use-term" in Suppl. material 4 provides an overview of the 87 terms (without their individual characterisation) and focuses on the "legal background" of each term, i.e. the potential variety of lawful sources and their possible characteristics (see chapters 7.2.1-7.2.3). The table complements Suppl. material 2, and because adding the "legal background" information from Suppl. material 4 to biodiversity specimens is optional, institutions may choose to work with Suppl. material 2 alone. However, the table can also be used to facilitate implementation in EDP systems, an example of how to read the Suppl. material 4 is provided in Table 4. In general, institutions have to spend significant resources if they want to provide information on lawful sources and their characteristics, defining what can be done with a collection object. On the other hand, the benefits of implementing that information appear just as great: First, staff and users quickly receive information on the legal background. For example, if the legal background is a document individually addressing a collection specimen, this usually has a higher clarity than more general statements in a law. Second, implementing the four characteristics - permission, prohibition, duty, restriction - makes it possible to search all specimens according to these characteristics. As a result, e.g. specimens with no restrictions can easily be compiled. Further the implementation of the four characteristics is also a basis for machine-actionability and the automated exchange of information across different databases all over the world.

Flexibility is a main advantage of the Typology of Legal/Contractual Terms for Biodiversity Specimens. If only a selection of terms from the Typology is applicable to specimens of an institution, only this selection can be implemented, not the entire Typology. Likewise, institutions can flexibly choose which information on the terms' lawful sources and their characteristics they want to add, according to their needs and resources: such information

may be implemented only for selected terms, only for selected lawful sources (e.g. all documents individually issued by a public authority), only for selected characteristics (e.g. all duties), only for selected collection items (e.g. newly collected specimens), only for internal use (to avoid liability disputes), or for any mixture of these options.

We are aware that terms in the Typology may overlap content-wise, not only with respect to general terms (No. 1-50) versus specific terms (No. 51-76), but also within the group of general terms. This seems to be an inevitable feature of terms originating from different jurisdictions.

Additional legal/contractual terms will be defined and added in the future, after more extensive consultations with the relevant research communities.

8. The DiSSCo openDS specification and event-based data model

8.1 The open Digital Specimen specification

DiSSCo is currently developing a specification, that is, a technical standard, for digital representations of curated collection specimens, called open Digital Specimen (**openDS**; Addink and Hardisty 2020). Digital representations are the sum of information on the Internet about a natural specimen.

The following description of openDS is based on communication with Wouter Addink, Work package lead SYNTHESYS+ WP6 and DiSSCo CSO Deputy Director in charge of Data and Technology, to closely align the description with the most up-to-date understanding of both groups: "openDS is a specification for open Digital Specimens, which includes three components: a data model, ontology and an API. The openDS data model and ontology aim to describe specimens and their related objects, such as the organisation holding it, a contract or permit dealing with it, images of a preserved specimen, field observation media files, notes in field notebooks, in such a way that they can be related and also interlinked with additional data and information derived from and associated with the specimen, such as traits, taxonomic treatments and genetic sequence data. The different groups of objects within the specification are defined by classes with properties, which can be properties pointing to other classes to relate them, or properties with actual content (data or metadata). The first version of openDS contains the classes which are mandatory for basic interoperability with the DiSSCo Technical Framework and cover use cases such as Minimum Information about a Digital Specimen (MIDS) and the Specimen Data Refinery (SDR). These core groups are: Digital Specimen, Multimedia, Agent, Provenance and Other, where Other is used for concepts that don't properly fit in the other core groups, but are not extensive enough to warrant the creation of another group."

The openDS specification will standardise data and metadata entries for natural specimens and related objects and, thus, provide a foundation for global data harmonisation and aggregation. Enabling the merging of (meta)data of different types and from many different sources for their combined reuse, it will facilitate joint analyses of large datasets and transdisciplinary information. Community agreement on and acceptance of the

specification will form the foundation for an integrated and extensible global data infrastructure providing (meta)data on biodiversity entities preserved and managed by scientific collections.

With this specification, the Distributed System of Scientific Collections (DiSSCo) can be one of the starting points for a globally connected, though distributed and federally governed infrastructure for biodiversity data. It can contribute to the integration of collection- and event-based data from species- to population-level into, for example, the Global Biodiversity Observation System (**GBiOS**) conceptualised by bioDISCOVERY/ Future Earth and GeoBon (Krug et al. 2022), or the Knowledge Centre for Biodiversity (European Commission 2022) of the European Commission. The FAIRness (Wilkinson et al. 2016) of openDS objects and their intrinsic functionality for linking out from discipline-specific platforms beyond the realm of biodiversity and its associated sectors, enables event-based data describing and recording collection objects and field observations to become part of a wide-ranging open and interoperable global data system and network bringing together all kinds of data, information and knowledge.

The foundational architecture for such versatile and highly transdisciplinary next-generation data infrastructures has been laid out in the Digital Object Architecture (**DOA**; Sharp 2016). The DOA aims to enable "an internet of FAIR data and services" (Wittenburg et al. 2019) and conceptualises a system, which will enable the wide-ranging, dense and flexible interlinking of data entities, based on open FAIR Digital Objects (**openFDOs**; Schultes and Wittenburg 2019) as its individual data elements. openFDOs are data packets in which "payload" data are wrapped by several layers of operational metadata that transform data of any type into Findable, Accessible, Interoperable and Reusable (**FAIR**; Wilkinson et al. 2016) entities that can be effectively shared online. Furthermore, as openFDOs data become machine-actionable, this is enabling the automatic updating and extended interlinking of the objects as the basis for building a provenance record and increasing transparency.

When openFDOs and the DOA in addition follow and intrinsically implement within their metadata layers and infrastructure functionality the rights and guidelines associated with Collective benefit, Authority to control, Responsibility and Ethics, together forming the **CARE** principles (Carroll et al. 2021) of data governance, they will empower data providers and promote data and metadata sharing and reuse in a wide range of contexts.

The metadata layers that are wrapping an openFDO and form the foundation of the power and versatility of the digital objects combine both generally applicable operational information common to all openFDOs, as well as discipline- and data type-specific metadata profiles. The openDS specification provides such a discipline-specific metadata profile for the biodiversity sciences and data of the applied biodiversity fields. It is designed to include sub-schemata to represent metadata specific to physical objects of the biological, anthropological and geological domains. Among the disciplines that are in scope, sub-schemata are present for specimens from anthropology, botany, geology, microbiology, palaeontology and zoology.

In conjunction with the DiSSCo technical infrastructure and its openDS specification, which are focused on providing general functionality for work with the digital representations of collection specimens, a module is under development specifically for the management of loans and visits to collections of physical objects, the **European Loans and Visits System** (**ELViS**; Islam and Addink 2023). Supporting the functional focus of this DiSSCo module, here as well (meta)data standardisation and harmonisation will enable the inclusion and handling of all kinds of digital biodiversity records. For example, in ELViS open Digital Specimens can be included that represent physical specimens whose information have been digitised to different extents (cp. MIDS levels; Haston and Chapman 2022). The ability to include objects is independent of their scientific context (e.g. bio-, geo-, anthropological or ethnocultural diversity), object type (e.g. microscope slides, DNA extracts, mounted vertebrates, artist models of organisms), material composition, preparation and preservation (e.g. frozen tissue samples, bones, mineral objects, dried organisms or clay cultural artefacts), as well as their location in the world and their holding organisation.

8.2 Transactional model

Building on the openFDO and biodiversity-focused openDS data specifications, event-based transactional models consider that (meta)data are not static, but are constantly created, accessed (retrieved and read), updated and deleted. That is, they are always changing due to the activities of humans and machines that are interacting with them. Event-based transactional models are able to support and incorporate these interactions and their arising information. They do so by digitally representing and implementing the functionality contained in interactions, thereby building capacity for processes. Hence, they digitally represent the dynamic nature of (meta)data.

Forming the core part of the backbone in the next generation of biodiversity data infrastructures, event-based transactional models enable the **bidirectional sharing** and **reuse** of physical objects and digital data. For example, in event-focused infrastructures not only collection-holding institutions provide data to users, but annotations and extensions with new insights from users flow back to the data providers. From the recording of individual transaction events as well as their persistent linking to the data object arises an object's **provenance** record. This record is created automatically with the help of transactional models. A complete time series of transactional records results in a full chain of custody for the physical object or data entity.

A chain of custody arising from a time series of provenance records that has been accumulating along, e.g., a research workflow can be transformed into an **operation-grade chain of custody**. This requires extensive resources and efforts to be allocated for development and continued maintenance. If required such an operational chain-of-custody can be further developed into supporting **tracking** and **tracing**. However, this would, for example, require extensive and expensive resources for the development and maintenance of such systems. Well-designed checks-and-balance environments will need to accompany them. The implementation of reliable tracking and tracing applications,

specifically those that are expected to scale up to the global level are far out of the range of the resources usually available to natural science collections. Among the technical requirements are that any provenance information is associated with sufficient, reliable and information-rich metadata. This includes metadata that might be legally required for the specific use case at hand. It furthermore requires that such provenance systems are complemented by the implementation, maintenance, and auditing of software applications, operating systems and hardware, as well as comprehensive security measures. These resource-intense requirements and the wide range of consequences associated with implementations of tracking and tracing have to be considered, for example, in the contexts of international benefit sharing, supply chain certification, or verification of geographic origin.

Overall, transactional models with the capabilities of announcing and transmitting events and their results across interlinked infrastructure landscapes form the basis for transparency, attribution and accountability.

The event-based transactional model used in DiSSCo, which therefore can also be used in upcoming versions of its ELViS module, is based on the W3C-recommended **Provenance Ontology** (PROV, Lebo et al. 2013). The functionality of the PROV standard is extended by the W3C-recommended **Web Annotation Ontology** (OA, Sanderson et al. 2016), which allows an infrastructure to capture and record more detailed information about transactions. The OA standard thus fosters the linking of several transactional events, potentially involving different parts of a physical or digital object, into a series. In this way, a sufficiently information-rich provenance record can be accumulated.

The core elements (Fig. 3) of an event in the PROV standard are a **prov:Activity**, associated with or performed by a **prov:Agent**, potentially acting as part of a specific **prov:Role**, on a **prov:Entity**, which can be physical or digital. In a series of activity events on one and the same object or entity, the OA standard provides more specificity by adding functionality for identifying which specific part of the object (the **oa:Target**) was modified in which way (the **oa:Body** of the annotation).

9 Extending DiSSCo's basic model by proposed components for conditions

In the context of DiSSCo's initial framework of considered use cases, its basic model (Islam et al. 2020), which is built based on the PROV ontology and extended by OA functionality as mentioned above, is fully adequate for describing events and transactions, as well as for building a provenance record. This is sufficient for the open and free sharing of digitised representations of physical objects and digital (meta)data that have no conditions associated. Originating in the basic biodiversity sciences, the open and free sharing of (meta)data that do not need special protection and do not come with specific considerations or concerns currently is DiSSCo's starting point.

Nevertheless, oriented towards the future, DiSSCo has the goal to develop and implement additional capacity. On one hand, DiSSCo's basic model's functionality needs to be expanded for the implementation of its ELViS module. The module will enable the DiSSCo technical infrastructure to act as an agent that mediates loans of unique physical objects. These are susceptible to damage and decay, at the same time there are no identical backup copies, in contrast to the digital world, in which such copies are possible. On the other hand, DiSSCo plans to be able in the future to support the sharing of data that have conditions attached. Such data can include, for example, sensitive (meta)data (cp. the exploitation of endangered natural populations), and information derived from physical specimens, as well as (meta)data that are governed by decisions of different rights holders and stakeholders. For example, science-informed policy processes by UN bodies can give rise to conditions, for example, to consider the Access and Benefit Sharing (ABS) regulations implementing the Nagoya Protocol[*1], the International Treaty on Plant Genetic Resources for Food and Agriculture 2001 (ITPGRFA) and the UN treaty on marine Biodiversity Beyond National Jurisdiction (BBNJ)[*1]. However, none of these conditions are considered by the basic PROV and OA ontologies.
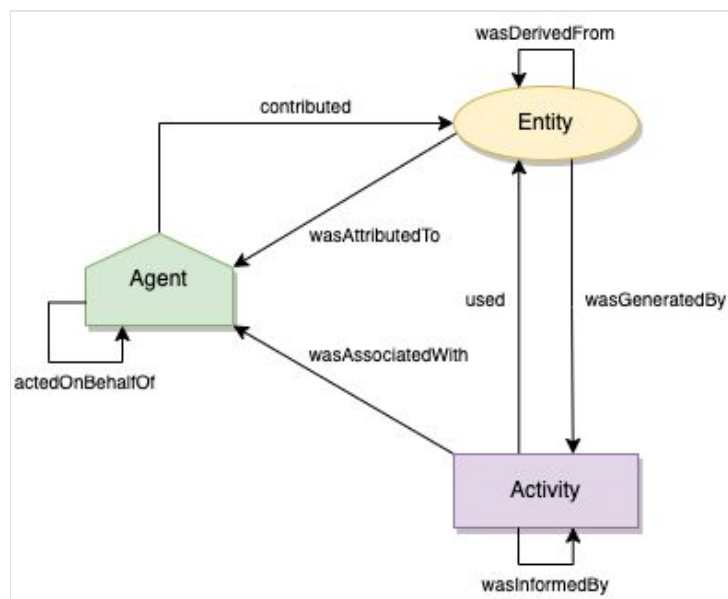

Figure 3. doi

The core elements of the PROV standard and their structure (recreated from Thessen et al. 2019 Appendix A, fig. S1, CC-BY 4.0; based on the PROV Model Primer, Copyright ©2013 W3C® (MIT, ERCIM, Keio, Beihang), All Rights Reserved. W3C liability rules, trademark rules and document use rules apply.)

A wide range of conditions has to be considered that are associated with biodiversity data. This publication deals with conditions recorded in legal, including contractual, documents

and policies that are associated with physical objects preserved in natural science collections. Generally, conditions may, for example,

1. concern the protection of sensitive (meta)data (e.g. informing on endangered populations or vulnerable data providers; also, GDPR-relevant information),

2. arise from specific considerations regarding data's social contexts and ethical uses, e.g. data and knowledge (including traditional knowledge) from indigenous peoples and local communities (IPLCs) flagged by Local Contexts labels (Local Contexts 2023),

3. govern the subsequent sharing of arising benefits with the original provider in return (cp. ABS regulations), and/or

4. have specific contractual conditions attached (cp. MTAs, collection policies, licence information).

While the technical functionality that we propose for extending the basic model should be forward-looking and generally applicable to any conditions, the specifics of categorising and representing social and ethical considerations, intellectual property rights (IPR) and further categories of conditions will need to be considered and developed in future work.

Even in the context of DiSSCo's current implementation focus on data without conditions, as already mentioned, procedures are different for physical objects and the data representing them, even if the underlying physical object in general can be openly shared and exchanged. Practical management tasks associated with physical objects, e.g. loan transactions and visits to collections, most often involve a careful consideration of appropriate conditions. These considerations include, for example, a review of available information that is attached to each object, and informs on e.g. concrete access to, required handling, packaging, and the use of the object. Who can receive the object? How should or does it need to be handled (e.g. hazadous materials)? Which ethical and legal, as well as collection policy information needs to be made available with the object? What actions and investigations are allowed on and with the object by the receiving party?

Hence, concerns for the preservation of the physical object and the need to consider its manifold contexts lead to concrete conditions that need to be incorporated and evaluated during both types of "events", that is, loan transactions and on-site visits. For the purpose of an integrated comprehensive, efficient and effective digital loans and visits system, the events, the involved physical specimens and agents, as well as the associated conditions must be digitally represented within DiSSCo's technical infrastructure. This digital representation has the effect that any information, including about conditions, is available and as needed can be shared with both the holding organisation's staff preparing the loan or visit, as well as the receiver of the physical object or visitor. Hence, easily accessible information on conditions associated with each physical and digital object enables informed decisions by providers, mediators and users as the cornerstones of potential transactions.

9.1 Outline of a technical integration of conditions into the transactional model

Certain functionality needs to be implemented, for conditions to take the step from an attached information resource to being functionally incorporated and forming an integral part of (automated) online transactions. All information, first, needs to be reliably and persistently **linked** to the digital representation of the physical object or digital data. Second, information needs to be made **machine-actionable** as a prerequisite for implementation and, third, all information will need to be embedded in an information model or architecture that provides the structure and semantics supporting **automated functionality**, specifically **reasoning**. These prerequisites are essential for empowering complex online tools such as, for example, DiSSCo and its ELViS module.

Therefore, information on conditions that are specific to certain events, agents, objects and data (e.g. contracts) or inherited from contexts (e.g. national laws and regulations, or collection policies) will need to be stored with or linked to the digital objects as part of their metadata. For this purpose, the openDS specification will need to be extended. A general "slot" for conditions will allow the integration of information about conditions into the openDS specification. To design the data structure for this slot and for identifying the metadata details in this slot, the working group developed, as described in Chapters 6 and 7, a two-level classification. The first level, **Document Categories and Types** (see Chapter 6) provides an initial, high-level orientation to users, as well as a scaffolding for the structuring of the second level, the **Typology of Legal/Contractual Terms for Biodiversity Specimens** (see Chapter 7). The Typology of Legal/Contractual Terms directly contributes to the foundation for machine-actionability.

Once information on conditions associated with a collection specimen, e.g. in the form of governmental or contractual documents, has been digitised and integrated into the metadata of the openDS object (Document Categories and Types), the contents of such information and documents, which describe concrete rules for behaviours and activities, need to be identified, extracted and represented as machine-actionable commands (see the Typology of Legal/Contractual Terms for Biodiversity Specimens as a development step towards this goal).

Finally, the basic event-based transactional model built using the PROV and OA ontologies will need to be extended to allow the incorporation of functionality that transforms such document-derived, machine-actionable information on conditions into corresponding, appropriate and well-designed computational commands, steps and routines.

9.2 Transforming linked legal documents into machine-actionable information

The Typology of Legal/Contractual Terms for Biodiversity Specimens (Chapter 7) has been developed primarily with the goal to document, standardise and harmonise the action-

oriented contents stated within and between Document Types, as well as found associated with the diverse contexts in which physical and digital objects are embedded.

At the same time, the standardised Typology with its "catchphrases" and their definitions takes an important step towards a "vocabulary". A vocabulary provides human-readable labels that can act as codified condition statements (see e.g. the labels of boxes and connections in Fig. 4). These labels are associated with corresponding machine-readable terms, which can act as commands in algorithms and computational routines (see e.g. the code examples below) that underlie and give rise to automated conditional actions.
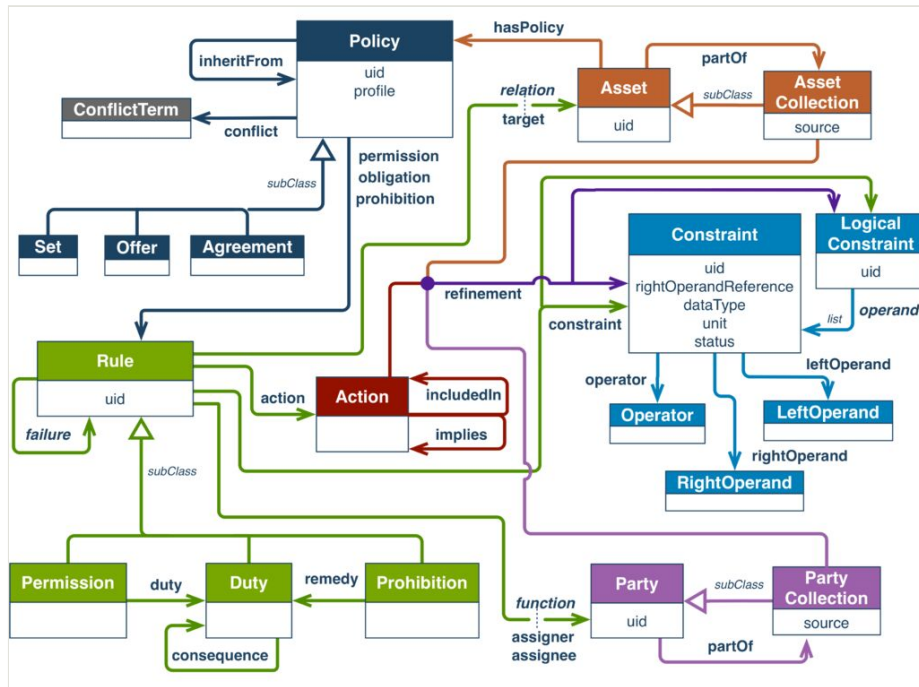

Figure 4. doi

ODRL information model (Diagram copied from "ODRL Information Model 2.2", Figure 1, https://www.w3.org/TR/odrl-model/; Copyright © 2018 W3C® (MIT, ERCIM, Keio, Beihang). W3C liability, trademark and permissive document license rules apply, see https://www.w3.org/Consortium/Legal/2015/copyright-software-and-document).

Such condition statements specific to the bio-, geo- and anthropodiversity sector, and the computational workflows that implement arising domain-specific computational decisions and consequences, can be integrated into an existing **Rights Expression Language** (REL) and, thus, take advantage of its already developed functions (i.e. expressions). The vocabulary and functionality of a REL expanded for the natural sciences domain will complement the PROV and OA ontologies that are currently forming the basis of DiSSCo's technical architecture. In the following, we are exploring the W3C-recommended **Open Digital Rights Language** (ODRL) ontology for use within the biodiversity sector and more specifically for an ELViS use case.

The ODRL ontology provides a general structure, as well as functions and workflows for transforming predefined condition statements into actions, which can be machine- and human-mediated, as appropriate (Fig. 4). In future work, the integration of biodiversity-relevant condition statements ("contents") and actions into the already existing information model of the ODRL ontology can be further explored, necessary extensions developed and their limits identified. In this way, conditions can be made functional, resulting in automated events and transactions. In addition, biodiversity-specific condition statements can form the basis for developing, implementing and populating informative user interfaces. These enable human decisions and subsequent manual input. The incorporation of human decisions into workflows might be implemented as, e.g., a click on a "yes" or "no" button by a user after they made a decision that was informed by a machine-generated assessment provided by the user interface, as well as, if available and disclosable, attached documentation materials.

When developing and implementing machine-actionable and automated assessments of legal, including contractual, and other conditions, the following statement from the "ODRL Implementation Best Practices - Draft Community Group Report 10 January 2023" needs to be kept in mind and considered:

"*While ODRL can represent elements of a license or regulation for machine consumption, it cannot replace them in court! It is best practice to explicitly point to the license or regulation that a policy models using the dc:terms property provided by the Dublin Core Metadata Initiative.*"

(Cited from Smith et al. 2023, Section 2.3. © Copyright 2023 the Contributors to the ODRL Implementation Best Practices Specification, published by the ODRL Community Group under the W3C Community Contributor License Agreement (CLA). A human-readable summary is available.)

Therefore, the final decision on an assessment of conditions associated with a physical specimen, openDS object or dataset always remains with the human user and requires in a final manual step the input of the human agent.

Fig. 5 provides a high-level, abstract view of the integration of ODRL functionality into the PROV-based transactional model employed by DiSSCo. The integration of an element representing "**Conditions**" (e.g. the terms from the typology) into the basic event-based transactional model (Fig. 4) fundamentally changes the structure of the data model. The model transitions from a two-dimensional triangular structure with bilateral links between all elements into a three-dimensional tetrahedron, in which for each node several bilateral links might be activated in parallel, the information of each of them being required for the event to be performed appropriately and as expected. Represented as a two-dimensional structure, a central junction arises (cp. the centroid of a tetrahedron) with connections to all of the four primary elements of the model, i.e. entity, agent, activity and condition. In this "**Decision**" hub simultaneous input from all four model elements is needed for the formulation of a valid outcome.
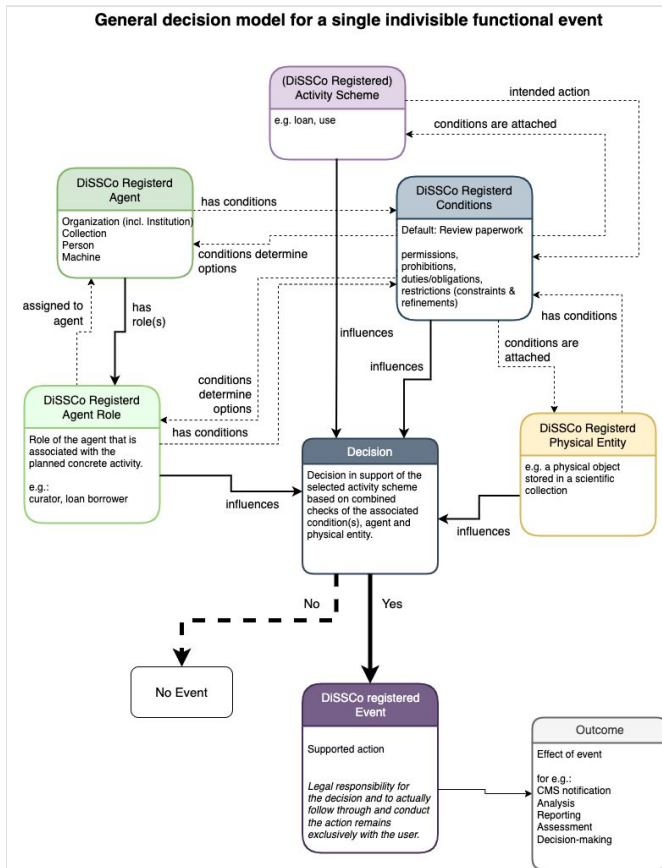
Figure 5. doi

Generalised decision model for a single atomized step, that is, a functional event that cannot be further divided. The decision model arises by integrating conditions (e.g. the terms from the typology) into the basic transactional model. The model can be set into the background of a physical or digital infrastructure (e.g. a collection institution or digital platform) that publicly presents open data to the world.

For example, a decision about a loan will only be possible if a user (manually) and/or the model (automatically) considers at the same time the involved elements: the **prov:Agents** (**odrl:Parties**; Who are the holding institution and the loaner? Where are they located, also relative to each other?), the **prov:Entity** (**odrl:Asset**), that is, the physical object that is requested for loan (e.g. is it currently available and can it be sent out by mail or is it too fragile?), the **prov:Activity** (**odrl:Action**; here: a loan event, including an access event, which in addition might be associated with further requested use activities, e.g. a scan by computer tomography/magnetic resonance tomography, removal of tissue for DNA extraction, etc.) and specific **Conditions** (cp. **odrl:Policy**) arising from the contexts associated with the object, including its geographic origin and history (e.g. is it a biological type specimen and is there an institutional or collection policy for sending out types?;

Where does the object come from and does it thus have legal documents attached, e.g. a MAT, governing access, use and subsequent duties of benefit sharing?).

In the ODRL ontology three categories of concrete conditions are defined, called **odrl:Rules** that can modify and thus have an impact on a decision, these are **odrl:Permission**, **odrl:Prohibition**, and **odrl:Duty** (for rules) or **odrl:Obligation** (for policies). **Restrictions** can be applied to these conditions in the form of **odrl:Constraints** modifying rules in general and/or **odrl:Refinements** of agents, activities and entities that enter the joint evaluation and decision process.

The overall structural system for conditions in ODRL is that all conditions are placed in the context of **odrl:Policies** at the highest level. Thereby, policies present the envelopes for **sets of odrl:Rules** (i.e. permissions, prohibitions, duties), these rules in turn can be modified by restrictions (i.e. constraints and refinements).

More fine-grained detail for an exploration of and insights into the applicability of the ODRL standard and its functionality to common tasks and use cases within biodiversity contexts is provided by considering the three types of policies defined within ODRL. These three policy categories are namely **odrl:Agreement**, **odrl:Offer** and **odrl:Set**. In a biodiversity context these can correspond to and be used for, for example, contracts between two parties, a licence attached to an odrl:Asset or prov:Entity by a rights-holder, and a governmental law or regulation, respectively.

Within **odrl:Actions** (cp. **prov:Activities**), the classes **odrl:Use** and **odrl:Transfer** of ownership form the two highest level subclasses. "Use" and "transfer" in general seem to be both pivotal to events and transactions, as well as to the conditions themselves that are associated with activities performed on and with (meta)data and digital representations of physical objects in the biodiversity sector. However, they are not only of interest in connection with ownership, but can interact with a wide range of conditions. Of interest are within the context of biodiversity data infrastructures, for example, the delegation of actions and their associated rules to third parties, or actions that involve multiple, that is more than two, parties each with different rights. This is an area that will need to be explored further, see Chapter 9.4 on next steps and open tasks.

9.3 Transforming concepts into code: an ELViS use case

In the previous sections a conceptual approach to the integration of conditions into the next-generation of biodiversity data infrastructures, based e.g. on the Digital Extended Specimen concept (DES; Hardisty et al. 2022) and represented by e.g. DiSSCo's event-based transactional data model, has been explored and outlined. In this section we provide an example of how these abstract considerations might be transformed into a concrete, code-based implementation. The chosen use case for the example is a loan transaction represented in, mediated by and made machine-actionable by the planned ELViS module of the DiSSCo technical infrastructure.

The **use case** expressed in human-readable prose can be described as follows: Curator Hortensia (names were chosen in honour of the taxon group used in the example) of Institution 2 requests the loan of a specific physical specimen identified as common earthworm from Institution 1 holding the specimen. Curator Annelie at collection institution 1 prepares the loan and assesses if the specimen can be sent out to curator Hortensia. Once curator Annelie decides that the loan can move forward, she links associated legal/ contractual documents with the loan transaction for online access by curator Hortensia. Finally, after the specimen has been sent out, the collection management record of the specimen gets updated with the new location information during the loan period.

Starting out with a minimal loan scenario the code base of the example is introduced and subsequently expanded by stepwise integrating additional information and functionality, so that an increasingly complex and thus more realistic and common loan scenario will be represented by machine-actionable code. The following **outline** provides an overview over the main stages of developing the loan transaction used in the example:

1.   Introduction to the JSON-LD serialisation.

2.   Digital representation of the elements involved in and necessary for the loan transaction using PROV.

3.   A simple loan transaction without associated conditions, represented by PROV functionality. 3a Including the delegation of the activity from the responsible institution to a curator. 3b Communicating legal documents that are associated with and inform the loan transaction.

4.   Updating the CMS or DiSSCo record to reflect the new locality of the physical specimen during the loan period using OA functionality. 4a Attaching information about rights to the annotation of the digital record.

5.   Integrating conditions governing the loan transaction using ODRL. 5a Adding global YES/NO conditions governing shipping and handling of the physical specimen. 5b Setting loan decisions that are context-dependent: To which countries and partners can the specimen be shipped?

6.   Moving towards an operational framework: exploring the policy life-cycle with inheritance between policies

A compilation of the example code provided for the loan use case can be found in Suppl. material 9.

JSON (JavaScript Object Notation, Wikipedia contributors 2023b) is a widely used, human-readable data exchange format that enables interoperability between independent infrastructure platforms. Its extension for linked data, **JSON-LD** (JavaScript Object Notation for Linked Data, Wikipedia contributors 2023c) opens up the opportunity of integrating efficiently the latest versions of distributed, independently maintained, often information-rich and standardised information resources, as for example ontologies, data files, database records or open FAIR Digital Objects (openFDOs), into data transactions

between federated digital infrastructure systems by linking to these resources. The links should be based on persistent, globally unique and resolvable identifiers (**PID**s), or more generally on Internationalised Resource Identifiers (IRIs, Wikipedia contributors 2022b). In the following example Uniform Resource Locators (URLs, Wikipedia contributors 2023d) are used, colloquially known as "web addresses" for locating and retrieving information on networks.

The following JSON-LD implementation of the ELViS loan use case was inspired by the examples and JSON-LD serialisations that have been developed and published for PROV (Provenance ontology), OA (Web Annotation ontology) and JSON-LD itself, which are provided by the World Wide Web Consortium, Inc., Delaware: Dover (W3C) at https://www.w3.org/ . More information, examples, explanations and background can be found at these links for

- PROV: https://openprovenance.org/prov-jsonld/

- OA: https://www.w3.org/TR/annotation-model/#complete-example

- JSON-LD: https://www.w3.org/TR/json-ld11/#data-model

This online validator can provide support for writing JSON-LD serialisation:

- JSON-LD Playground: https://json-ld.org/playground/

Step 1: Introduction to the structure of a JSON-LD serialisation

In the example code developed for the loan use case, the top-level of a JSON-LD serialisation (Fig. 6) is structured into two parts or entries: **@context** and **@graph.**

```
1  {
2      "@context" : [ {  } ],
3      "@graph" :   [ {  } ]
4  }
```

Figure 6. `doi`

EXAMPLE CODE 1: Top-level framework of the example JSON-LD serialisation

In **@context** a shared digital context is defined for the overall JSON-LD object, in which terms are mapped to full IRIs (Fig. 7). This mapping enables the use of a short-hand notation throughout the object that is better human-readable. Thus, the term "http://example.com/physicalSpecimen1" in the following code can be written as " **ex:physicalSpecimen1**" and the full namespace address for the PROV term "**Entity**", that is, "https://www.w3.org/ns/prov#Entity" can be shortened and unambiguously identified by using the short-hand "**prov:Entity**".

```
 1  {
 2      "@context" : [ {
 3          "xsd" : "http://www.w3.org/2001/XMLSchema#", // date & time
 4          "dcterms" : "http://purl.org/dc/terms/",      // titles
 5          "ex" : "http://example/",                     // example prefix
 6          "foaf" : "http://xmlns.com/foaf/0.1/",        // names
 7          "prov" : "http://www.w3.org/ns/prov#",        // provenance
 8          "oa": "http://www.w3.org/ns/anno.jsonld",     // annotations
 9          "odrl": "http://www.w3.org/ns/odrl.jsonld",   // conditions
10      }, "https://openprovenance.org/prov-jsonld/context.json"
11      ],
12      "@graph" : [ {  } ]
13  }
```

Figure 7. doi

EXAMPLE CODE 2: Top-level framework: @context

In **@graph** the data structure and data of the actual transaction are described as entries in a list of data structures that are called maps or dictionaries, and which are **key : value** pairs (https://infra.spec.whatwg.org/#data-structures). These entries are each independent, but interlinked and interacting JSON objects themselves that might be, e.g., openFDOs. The objects require at least an **@type** property describing them. They describe an expression (i.e. a snippet of code) as a resource, which can be explicitly named by the use of an identifier (**@id)** or remain anonymous, in which case no identifier is associated with the resource or expression. The following graph in EXAMPLE CODE 3 includes two anonymous objects within **@graph**, one a **prov:Entity**, the other a **prov:Agent** (Fig. 8).

```
 1  {
 2      "@context" : [ {
 3          "xsd" : "http://www.w3.org/2001/XMLSchema#", // date & time
 4          "dcterms" : "http://purl.org/dc/terms/",      // titles
 5          "ex" : "http://example/",                     // example prefix
 6          "foaf" : "http://xmlns.com/foaf/0.1/",        // names
 7          "prov" : "http://www.w3.org/ns/prov#",        // provenance
 8          "oa": "http://www.w3.org/ns/anno.jsonld",     // annotations
 9          "odrl": "http://www.w3.org/ns/odrl.jsonld",   // conditions
10      }, "https://openprovenance.org/prov-jsonld/context.json"
11      ],
12      "@graph" : [ {
13          "@type": "prov:Entity",
14      },{
15          "@type" : "prov:Agent",
16      } ]
17  }
```

Figure 8. doi

EXAMPLE CODE 3: Top-level framework: @graph

In the following, the context defined in **@context** remains the same, so that the subsequent example code snippets focus only on the development of the expressions in the data structure part under **@graph**, and the **@context** part will be omitted.

Step 2: Digital representation of involved objects

At the core of a loan transaction lies a collection specimen that is sent to a different locality for a certain time. To digitally represent this loan transaction, the digital representation of the specimen, its openDS twin, needs to be identified.

In the PROV ontology, the specimen's representation as an openDS instance takes the role of a **prov:Entity**, thus, the JSON object is of **"@type": "prov:Entity"**. The expression **"@id" : "ex:physicalSpecimen1"** links the JSON object to the openDS resource (that is, the digital twin of the physical specimen stored in the DiSSCo technical infrastructure) and, thus, uniquely identifies it (Fig. 9).

```
1 {
2     "@graph" : [ {
3         "@type": "prov:Entity",
4         "@id" : "ex:physicalSpecimen1",      // the openDS PID
5         "prov:label" : "Common earthworm"
6     } ]
7 }
```

Figure 9. doi

EXAMPLE CODE 4: Digital representation of the collection specimen

To improve human readability, a **prov:label** is set for the **prov:Entity**, informing us that the physical specimen in the focus of the loan transaction of this use case had been identified as a common earthworm *Lumbricus terrestris* Linnaeus, 1758 (Fig. 9).

A loan transaction has two agents, first, the institution holding the physical specimen, which is preparing the loan. Second, the institution that requested the loan and will receive it once the loan has been approved and sent out.

Both institutions are digitally represented as agents using **"@type": "prov:Agent"**. The PROV standard allows a classification of agents into agent subcategories, we choose the subcategory **prov:Organization** for the description of the institutional agents and set it using the expression **"prov:type" : ["prov:Organization"]** . The institutions in this example are uniquely identified by their Research Organization Registry (ROR) identifiers, which are incorporated into the script as the values of **@id** (Fig. 10).

We know based on the given human-readable ID to which of the two institutions the specimen belongs. However, this information is not machine-readable and -actionable. Therefore, it is necessary to explicitly link the specimen and Institution 1 that is holding it by employing an object of **"@type": "prov:Attribution"**. The **prov:Attribution**-object links an **prov:Entity** to its associated **prov:Agent** (Fig. 11).

```
1  {
2      "@graph" : [ {
3          "@type" : "prov:Entity",
4          "@id" : "ex:physicalSpecimen1",
5          "prov:label" : "Common earthworm"
6      },{
7          "@type" : "prov:Agent",
8          "prov:type" : [ "prov:Organization" ],       // agent subcategory
9          "@id" : "ex:institution1_holding",           // ROR PID
10         "prov:location" : "http://www.wikidata.org/entity/Q40",     // AT
11         "prov:label" : "NHMW, Vienna, Austria"
12     },{
13         "@type" : "prov:Agent",
14         "prov:type" : [ "prov:Organization" ],       // agent subcategory
15         "@id" : "ex:institution2_requesting",        // ROR PID
16         "prov:location" : "http://www.wikidata.org/entity/Q27",     // IE
17         "prov:label" : "National Museum of Ireland, Dublin, Ireland"
18     } ]
19 }
```

Figure 10. doi

EXAMPLE CODE 5: Adding the legal agents of a loan transaction

```
1  {
2      "@graph" : [ {
3          "@type" : "prov:Entity",
4          "@id" : "ex:physicalSpecimen1",
5          "prov:label" : "Common earthworm"
6      },{
7          "@type" : "prov:Agent",
8          "prov:type" : [ "prov:Organization" ],
9          "@id" : "ex:institution1_holding",
10         "prov:location" : "http://www.wikidata.org/entity/Q40",
11         "prov:label" : "NHMW, Vienna, Austria"
12     },{
13         "@type" : "prov:Agent",
14         "prov:type" : [ "prov:Organization" ],
15         "@id" : "ex:institution2_requesting",
16         "prov:location" : "http://www.wikidata.org/entity/Q27",
17         "prov:label" : "National Museum of Ireland, Dublin, Ireland"
18     },{
19         "@type" : "prov:Attribution",
20         "@id" : "ex:institution1_accession",         // (P)ID
21         "entity" : "ex:physicalSpecimen1",           // specimen
22         "agent" : "ex:institution1_holding"          // holding institution
23     } ]
24 }
```

Figure 11. doi

EXAMPLE CODE 6: Linking the specimen to its holding institution

Step 3: A simple loan transaction without associated conditions

Now the loan event itself needs to be defined and represented as JSON objects. Two objects are defined to capture the fundamental information of the loan process. These

objects are the complementary entities of "**@type**" : "**prov:Activity**" and "**@type**": "**prov:Usage**".

The object of type **prov:Activity** defines the outline of the activity, that is, it sets the stage by defining its **@type** as an activity and minting an **@id** that uniquely identifies this concrete activity, this loan transaction. Furthermore, this object represents the (planned) duration of the loan (given by start and end time).

Due to the structure of the standard and its JSON-LD implementation, a second object seems to be needed that "qualifies" the activity as being of type **prov:Usage**. Usage describes an interaction between two resources that are in this case the interaction of a **prov:Activity** with a **prov:Entity**, in which the **prov:Activity prov:used** the **prov:Entity** (see diagram and examples at https://openprovenance.org/prov-jsonld/#introduction-qualification-pattern as well as https://www.w3.org/TR/2013/REC-prov-o-20130430/#description-qualified-terms). The **activity** and the **entity** of **prov:Usage** are unambiguously identified by 1) the (P)ID of this specific loan transaction as defined in the object of type **prov:Activity** and 2) the (P)ID of the specimen that is requested and prepared for loan, defined in the object of type **prov:Entity**, respectively. Both involved objects were defined earlier in the script.

After the loan activity defined as **prov:Activity** with the "**@id**" : "**ex:loan2023_1**" has been linked to the specimen defined as "**entity**" : "**ex:physicalSpecimen1**" , it needs to be associated with the two involved agents, **ex:institution1_holding** and **ex:institution2_requesting**. The object **prov:Association** makes it possible to do so. In addition, the **prov:Association** is further qualified by providing metadata on the plan or template that the loan activity follows ("**plan**" : "**ex:elvis_standard_loan_routine**") as well as the role of the two agents, stored as **prov:role** property.

Finally, after the loan transaction has been recorded and the physical specimen has been sent out to Institution 2, the specimen needs to be associated with the receiving institution for the duration of the loan. Thus, the attribution of the specimen needs to be updated, using again **prov:Attribution** for this purpose (Fig. 12).

Obviously, this is a very rudimentary way of recording the change in association for the specimen during the loan that doesn't differentiate between properties that change (e.g. locality and associated curator) and those properties that remain associated with the original holding institution (e.g. ownership and accession information). Here, expressions need to be found or developed in future work that enable the system (e.g. the DiSSCo technical infrastructure) to update and record only those properties that require updating

Step 3a: Delegation of the activity from the responsible institution to a curator

While loan transactions officially (legally) often are transactions between institutions, in reality, it is collection staff and scientists who act on behalf of institutions. Organisations' concrete activities are delegated to individual persons. The PROV ontology can represent patterns of delegation.

To represent a delegation, two additional entities are defined for each delegation relationship. First, a **prov:Person** is defined as **prov:Agent** in addition to the already defined **prov:Organization**-agent object. This **prov:Person** object is uniquely identified by the **@id** of the involved person (here referred to as **ex:curator1** or **ex:curator2**), which might be, for example, an ORCID. For improved human readability, the name of the person (here Annelie or Hortensia) is added to the JSON object representing the involved person.

```
 1 {
 2     "@graph" : [ {
 3         "@type" : "prov:Entity",
 4         "@id" : "ex:physicalSpecimen1",
 5         "prov:label" : "Common earthworm"
 6     },{
 7         "@type" : "prov:Agent",
 8         "prov:type" : [ "prov:Organization" ],
 9         "@id" : "ex:institution1_holding",
10         "prov:location" : "http://www.wikidata.org/entity/Q40",
11         "prov:label" : "NHMW, Vienna, Austria"
12     },{
13         "@type" : "prov:Agent",
14         "prov:type" : [ "prov:Organization" ],
15         "@id" : "ex:institution2_requesting",
16         "prov:location" : "http://www.wikidata.org/entity/Q27",
17         "prov:label" : "National Museum of Ireland, Dublin, Ireland"
18     },{
19         "@type" : "prov:Attribution",
20         "@id" : "ex:institution1_accession",
21         "entity" : "ex:physicalSpecimen1",
22         "agent" : "ex:institution1_holding"
23     },{
24         "@type" : "prov:Activity",              // loan activity
25         "@id" : "ex:loan2023_1",                // unique loan identifier
26         "startTime": "2023-03-31T09:21:00.000+01:00",   // start
27         "endTime": "2025-04-01T15:21:00.000+01:00"      // end
28     },{
29         "@type" : "prov:Usage",                 // loan details
30         "activity" : "ex:loan2023_1",           // id of involved activity
31         "entity" : "ex:physicalSpecimen1"       // id of involved entity
32     },{
33         "@type" : "prov:Association",           // loan details
34         "plan" : "ex:elvis_standard_loan_routine",  // id of ELViS routine
35         "activity" : "ex:loan2023_1",           // id of involved activity
36         "agent" : "ex:institution1_holding"     // id of involved agent
37         "prov:role" : "ex:loan_provider"        // role of involved agent
38     },{
39         "@type" : "prov:Association",           // loan details
40         "plan" : "ex:elvis_standard_loan_routine",  // id of ELViS routine
41         "activity" : "ex:loan2023_1",           // id of involved activity
42         "agent" : "ex:institution2_requesting"  // id of involved agent
43         "prov:role" : "ex:loan_recipient"       // role of involved agent
44     },{
45         "@type" : "prov:Attribution",           // Updating of new status
46         "@id" : "ex:institution2_loanReceived",
47         "entity" : "ex:physicalSpecimen1",
48         "agent" : "ex:institution2_requesting"  // current association
49     } ]
50 }
```

Figure 12. doi

EXAMPLE CODE 7: Defining the loan activity of the transaction

The second object defines the delegation itself ("**@type": "prov:Delegation**") and its unique **@id** ("**@id" : "ex:toCurator1**" or "**@id" : "ex:toCurator2**"). The following line "**responsible" : "ex:institution1_holding**" (or **"responsible" : "ex:institution2_requesting**") identifies the delegating organisation, while the next line "**delegate" : "ex:curator1**" (or **delegate" : "ex:curator2**") identifies the person that performs the activity in delegation for the organisation. The last line **"activity" : "ex:loan2023_1**" identifies the activity that gets delegated, that is, to which the delegation refers. A more realistic implementation might divide this activity further into institution-

specific sub-activities, that is, e.g. the sub-activities **"ex:preparingLoan"** and **"ex:receivingLoan"** that can inherit the delegation (Fig. 13).

```
1  {
2      "@graph" : [ {
3          "@type" : "prov:Entity",
4          "@id" : "ex:physicalSpecimen1",
5          "prov:label" : "Common earthworm"
6      },{
7          "@type" : "prov:Agent",
8          "prov:type" : [ "prov:Organization" ],
9          "@id" : "ex:institution1_holding",
10         "prov:location" : "http://www.wikidata.org/entity/Q40",
11         "prov:label" : "NHMW, Vienna, Austria"
12     },{
13         "@type" : "prov:Agent",                    // curator handling loan
14         "prov:type" : [ "prov:Person" ],
15         "@id" : "ex:curator1",                     // e.g. ORCID
16         "foaf:givenName" : [{ "@value" : "Annelie" }]
17     },{
18         "@type": "prov:Delegation",                // delegation
19         "@id" : "ex:toCurator1",                   // PID
20         "responsible" : "ex:institution1_holding", // defined above
21         "delegate" : "ex:curator1",                // defined above
22         "activity" : "ex:loan2023_1"               // defined below
23     },{
24         "@type" : "prov:Agent",
25         "prov:type" : [ "prov:Organization" ],
26         "@id" : "ex:institution2_requesting",
27         "prov:location" : "http://www.wikidata.org/entity/Q27",
28         "prov:label" : "National Museum of Ireland, Dublin, Ireland"
29     },{
30         "@type" : "prov:Agent",                    // curator receiving loan
31         "prov:type" : [ "prov:Person" ],
32         "@id" : "ex:curator2",                     // e.g. ORCID
33         "foaf:givenName" : [{ "@value" : "Hortensia" }]
34     },{
35         "@type": "prov:Delegation",                // delegation
36         "@id" : "ex:toCurator2",                   // PID
37         "responsible" : "ex:institution2_requesting", // defined above
38         "delegate" : "ex:curator2",                // defined above
39         "activity" : "ex:loan2023_1"               // defined below
40     },{
41         "@type" : "prov:Attribution",
42         "@id" : "ex:institution1_accession",
43         "entity" : "ex:physicalSpecimen1",
44         "agent" : "ex:institution1_holding"
45     },{
46         "@type" : "prov:Activity",
47         "@id" : "ex:loan2023_1",
48         "startTime": "2023-03-31T09:21:00.000+01:00",
49         "endTime": "2025-04-01T15:21:00.000+01:00"
50     },{
51         "@type" : "prov:Usage",
52         "activity" : "ex:loan2023_1",
53         "entity" : "ex:physicalSpecimen1"
54     },{
55         "@type" : "prov:Association",
56         "plan" : "ex:elvis_standard_loan_routine",
57         "activity" : "ex:loan2023_1",
58         "agent" : "ex:institution1_holding"
59         "prov:role" : "ex:loan_provider"
60     },{
61         "@type" : "prov:Association",
62         "plan" : "ex:elvis_standard_loan_routine",
63         "activity" : "ex:loan2023_1",
64         "agent" : "ex:institution2_requesting"
65         "prov:role" : "ex:loan_recipient"
66     },{
67         "@type" : "prov:Attribution",
68         "@id" : "ex:institution2_loanReceived",
69         "entity" : "ex:physicalSpecimen1",
70         "agent" : "ex:institution2_requesting"
71     } ]
72 }
```

Figure 13. doi

EXAMPLE CODE 8: Delegation of the activity

Step 3b: Communicating legal documents that are associated with and inform the loan transaction

In the ELViS use case it is important that documents associated with and governing the loan transaction are transferred from the holding institution to the institution that receives the physical specimen as loan.

Such legal and ethical documents and information can be directly associated with the physical specimen (the **prov:Entity**, in that case likely as part of a **prov:Collection**). In addition, certain documents, e.g. defining institutional and collection policies, might be associated with and govern the loan event itself (not the specimen directly). In that case, a set of documents will be assembled in preparation of the loan that is loan-specific.

In general, documents that will be attached to and transferred with the loan will be focused on the physical specimen and the tangible, offline loan process, in accordance with the main perspective of our use case. They might include selected and/or abridged versions of documents or their extracted information contents, whereby all three will be associated with the physical specimen.

Information that governs the loan transaction can be incorporated into the digital representation of the loan event itself, using objects of "**@type" : "prov:Communication**". The functionality of **prov:Communication** makes it possible to represent that one activity informs another activity. In the example, the **prov:Communication** object has the (P)ID **ex:transferLegalInfo**.

The activity that is **informed** by the **prov:Communication** is the earlier defined loan activity (**ex:loan2023_1**). The activity that informs the loan event is the **informant**. To define the **informant** activity, again a **prov:Activity** entity is generated (with the (P)ID **ex:createdLegalInfoCollection**) and further qualified by a **prov:Usage** object. The **prov:Usage** expression associates a collection of legal documents and information (a **prov:Entity**) with the **ex:createdLegalInfoCollection** activity. The entity representing the collection remains anonymous in the example, that is, without (P)ID. It is defined by a list of two elements that have the IDs **ex:assocLegalDoc1** and **ex:assocLegalDoc2** (Fig. 14).

```
 1  {
 2    "@graph" : [ {
 3      "@type" : "prov:Activity",              // providing legal info
 4      "@id" : "ex:createdLegalInfoCollection"         //PID collection
 5    },{
 6      "@type" : "prov:Usage",                 // compilation of infos
 7      "activity" : "ex:createdLegalInfoCollection",   // see above
 8      "entity" : [{
 9          "@type" : "prov:Entity",          // associated docs & info
10          "@id" : "ex:assocLegalDoc1"
11        },{
12          "@type" : "prov:Entity",
13          "@id" : "ex:assocLegalDoc2"
14      }]
15    },{
16      "@type" : "prov:Communication",         // information transfer
17      "@id" : "ex:transferLegalInfo",         // PID of the transfer
18      "informant" : "ex:createdLegalInfoCollection",  // ID of the info
19      "informed" : "ex:loan2023_1"            // ID of the loan event
20    } ]
21  }
```

Figure 14. doi

EXAMPLE CODE 9: Communicating legal information associated with the loan transaction

The information encoded by using PROV, so far represents some basic building blocks of a loan transaction. This digital representation of the loan transaction in the ELViS module of the DiSSCo technical infrastructure can be made accessible to both (or all) partners of the loan transaction. Shared online via the ELViS system, digital representations of loan transactions enhance communication between the involved partners and in this way can support and mediate interactions associated with loans of physical specimens.

Step 4: Updating of the CMS or DiSSCo record to reflect the new locality of the physical specimen during the loan period

Supported by the digitised information available in the online ELViS system, once the decision has been made to send out a physical specimen, it can be of interest to collection managers to update the record of the physical specimen with some or all of the information associated with the loan. Such an update can make sense for the record of the physical specimen in an institution's local collection management system (CMS) or in the globally accessible openDS twin of the physical specimen stored in the DiSSCo technical infrastructure. We will use the example of updating the locality information associated with the physical specimen to introduce functionality provided by the Web Annotation Ontology (OA).

In the following code examples the **@context** entity will be added with one line defining the specific namespace that is the focus of this part of the example code. This is to highlight that the examples are now using expressions of additional, more specialised ontologies.

EXAMPLE CODE 10 shows an encoding of an annotation event that modifies the local CMS record or a globally shared openDS twin of the physical specimens with information taken from the loan transaction. In this case, the locality information is updated to reflect the location of the physical specimen during the loan period. If set up in advance by the holding institution, it is assumed that such an annotation event is automatically triggered by the ELViS system at the time the loan is sent out, as well as when it is returned and the physical specimen is reinserted into the collection of its steward institution. In this case, no human involvement is needed, since the transactional model of ELViS will support machine-actionability.

The JSON object of **"@type" : "oa:Annotation"** is identified by a (P)ID, here represented by **ex:update_spec1_location**. The keywords **body** and **target** are functionality provided by the OA ontology. **body** identifies the information (e.g. provided by a standardised locality string) that will be inserted into the CMS record or openDS metadata, and thus update the locality information there. The **target** identifies and links to the CMS record or openDS object that represents the physical specimen and needs updating (Fig. 15).

Step 4a: Attaching information about rights to the annotation of the digital record

The OA ontology defines a key - value pair (**rights**) to capture and store rights information that is associated with an annotation event itself, as well as the information given in the body and/or the target object (here: the CMS record or openDS). EXAMPLE CODE 11 shows how rights information can be integrated at the level of the annotation event as a

whole (**ex:AnnotationRightsInfo**) and at the level of the target object. The expression "**rights": "ex:RecordRightsInfo**" is only changing and setting the rights for the target context (Fig. 16).

```
1  {
2      "@context" : [{ "oa": "http://www.w3.org/ns/anno.jsonld" }],
3      "@graph" : [ {
4          "@type" : "oa:Annotation",          // of digital representation
5          "@id" : "ex:update_spec1_location",       // PID of annotation
6          "body": "ex:loanLocation",            // new location info
7          "target": [{
8              "@id" : "ex:CmsRecordSpec1",        // PID CMS record
9          }]
10     } ]
11 }
```

Figure 15. `doi`

EXAMPLE CODE 10: Annotation of a corresponding CMS record (or openDS twin)

```
1  {
2      "@context" : [{ "oa": "http://www.w3.org/ns/anno.jsonld" }],
3      "@graph" : [ {
4          "@type" : "oa:Annotation",
5          "@id" : "ex:update_spec1_location",
6          "rights": "ex:AnnotationRightsInfo",        // assoc. w/ annot.
7          "body": "ex:loanLocation",
8          "target": [{
9              "@id" : "ex:CmsRecordSpec1",
10             "rights": "ex:RecordRightsInfo"        // assoc. w/ record
11         }]
12     } ]
13 }
```

Figure 16. `doi`

EXAMPLE CODE 11: Adding information on rights associated with the CMS record and annotation itself

Step 5: Integrating conditions governing the loan transaction

The digital representation of the loan transaction at this point provides a sufficiently developed context for introducing conditions into the code of the example. Functionality provided by the ODRL ontology is used as an example of how legal information provided by documents can be transformed into and implemented as machine-actionable commands.

Step 5a: Adding global YES/NO conditions governing shipping and handling of the physical specimen

In preparation of a loan transaction the conditions attached to the physical specimen need to be clarified. This involves, for example, an assessment if the specimen can be shipped

to a different collection or research institution, and which actions on and with the specimen are allowed by the receiving partner.

Chapter 4.6 provides proposals for standardised legal/contractual terms that can be used to set conditions that might apply in the context of a loan transaction. In future work, these terms can be integrated into the functionality provided by the ODRL ontology (using e.g. the "profile" functionality, see https://www.w3.org/TR/odrl-model/#profile) or other rights expression languages. The W3C recommendation webpage for the "ODRL Information Model 2.2" (https://www.w3.org/TR/2018/REC-odrl-model-20180215) and the "ODRL Implementation Best Practices" webpage (https://w3c.github.io/odrl/bp/) offer an introduction to and guidelines for implementing the ODRL ontology in JSON-LD.

In the following code examples, a globally applicable (that is, no constraints or refinements apply) permission is encoded that allows unconditional shipping to 3rd parties. A second global permission subsequently is added that provides explicit consent to physical analyses of the material specimen.

In EXAMPLE CODE 12 a contractual agreement is digitally represented by using the ODRL policy type **odrl:Agreement**. This loan contract is unambiguously identified by and has the address of **ex:inst1_inst2_loan_contract**. The next line of code links to a (future) ODRL profile ("**profile": "ex:biodivDomain_profile**") that defines legal/contractual (cp. Chapter 4.6.4) - and at that point potentially also social and ethical (see CARE principles) - terms that are specific to the biodiversity domain.

The agreement explicitly provides a **permission** that allows the shipping of the physical specimen to third parties ("**action" : "ex:shipping_to_3rdParty**"), for example implementing term **23_YES_3rd party** of the Typology of Legal/Contractual Terms. This **odrl:Rule** has been set by the institution holding the physical specimen (the "**assigner" : "institution1_holding**"). It governs the rights and conditions associated with the physical specimen ("**target" : "ex:physicalSpecimen1**"), for which conditions are evaluated for the concrete loan transaction in question. The loan requesting institution is identified by "**assignee" : "ex:institution2_requesting**".

As before, the informative elements of this policy are defined by links to PID/ID-identified information (which might be formatted and deposited as openDS or openFDOs objects with other specifications) that is stored independently within the Distributed System of Scientific Collections (DiSSCo) or the wider network of a decentral, federated digital object architecture (Fig. 17).

The same contractual agreement might further set a second permission ("**action" : "ex:physical_analysis**") and allow anatomical-morphological analyses on the physical specimen (Fig. 18), compare **10_YES_physical_analysis** of the Typology of Legal/Contractual Terms (Annex 2).

EXAMPLE CODE 14 alternatively implements a globally valid **prohibition** that applies to the physical specimen and prohibits all shipping of the specimen ("**action" : "ex:noShipping**") implementing **80_NO_shipping** of the Typology (Fig. 19).

```
1  {
2      "@context" : [{ "odrl": "http://www.w3.org/ns/odrl.jsonld" }],
3      "@graph" : [ {
4          "@type" : "odrl:Agreement",
5          "@id" : "ex:inst1_inst2_loan_contract",
6          "profile" : "ex:biodivDomain_profile",
7          "permission" : [{
8              "target" : "ex:physicalSpecimen1",
9              "assigner" : "ex:institution1_holding",
10             "assignee" : "ex:institution2_requesting",
11             "action" : "ex:shipping_to_3rdParty"
12         }]
13     } ]
14 }
```

Figure 17. doi

EXAMPLE CODE 12: Applying a permission to the loan transaction

```
1  {
2      "@context" : [{ "odrl": "http://www.w3.org/ns/odrl.jsonld" }],
3      "@graph" : [ {
4          "@type" : "odrl:Agreement",
5          "@id" : "ex:inst1_inst2_loan_contract",
6          "profile" : "ex:biodivDomain_profile",
7          "permission" : [{
8              "target" : "ex:physicalSpecimen1",
9              "assigner" : "ex:institution1_holding",
10             "assignee" : "ex:institution2_requesting",
11             "action" : "ex:shipping_to_3rdParty"
12         }],
13         "permission" : [{
14             "target" : "ex:physicalSpecimen1",
15             "assigner" : "ex:institution1_holding",
16             "assignee" : "ex:institution2_requesting",
17             "action" : "ex:physical_analysis"
18         }]
19     } ]
20 }
```

Figure 18. doi

EXAMPLE CODE 13: Adding a second permission to the loan transaction

Step 5b: Representing information for loan decisions that is context-dependent

Often decisions associated with loan transactions are context-dependent. In the following we provide example code for implementing rules representing answers to the question to which countries and partners a physical specimen can be sent out.

The basis again is an explicit **permission** with its associated information on the **target**, the **assigner** and the **assignee**. The **action** now implements a different predefined legal/ contractual term, that is **24_obligations_3rd_party**, with the outcome after a successful evaluation however remaining the same **ex:shipping_to_3rdParty**.

```
 1  {
 2      "@context" : [{ "odrl": "http://www.w3.org/ns/odrl.jsonld" }],
 3      "@graph" : [ {
 4          "@type" : "odrl:Agreement",
 5          "@id" : "ex:inst1_inst2_loan_contract",
 6          "profile" : "ex:biodivDomain_profile",
 7          "prohibition" : [{
 8              "target" : "ex:physicalSpecimen1",
 9              "assigner" : "ex:institution1_holding",
10              "assignee" : "ex:institution2_requesting",
11              "action" : "ex:noShipping"
12          }]
13      } ]
14  }
```

Figure 19. doi

EXAMPLE CODE 14: Applying a globally applicable prohibition to the loan transaction

The permission is modified by a **constraint** expression that limits the permission for shipping of the physical specimen to third parties located in countries within the European Union. If that constraint is evaluated and the country, in which the receiving institution is located, satisfies the constraint (i.e. it is a member of the EU for this example), then an **action** appropriate to fulfilling the term **24_obligations_3rd party** from Annex 2 can be conducted, i.e. the loan transaction can go through and the physical specimen can be prepared for shipping to this institution.

The second global permission (allowing anatomical-morphological analyses) is not impacted by the constraint to the first permission and remains as before (Fig. 20).

```
 1  {
 2      "@context" : [{ "odrl": "http://www.w3.org/ns/odrl.jsonld" }],
 3      "@graph" : [ {
 4          "@type" : "odrl:Agreement",
 5          "@id" : "ex:inst1_inst2_loan_contract",
 6          "profile" : "ex:biodivDomain_profile",
 7          "permission" : [{
 8              "target" : "ex:physicalSpecimen1",
 9              "assigner" : "ex:institution1_holding",
10              "assignee" : "ex:institution2_requesting",
11              "constraint" : [{            // define allowed countries
12                  "leftOperand" : "spatial",
13                  "operator" : "isPartOf",
14                  "rightOperand": "http://www.wikidata.org/entity/Q458",
15                  "comment" : "European Union"
16              }],
17              "action" : "ex:shipping_to_3rdParty"
18          }],
19          "permission" : [{
20              "target" : "ex:physicalSpecimen1",
21              "assigner" : "ex:institution1_holding",
22              "assignee" : "ex:institution2_requesting",
23              "action" : "ex:physical_analysis"
24          }]
25      } ]
26  }
```

Figure 20. doi

EXAMPLE CODE 15: Defining a context-dependent decision - based on country

constraint properties modify specifically **prov:Rules**, which include **prov:Permissions**, **prov:Prohibitions** and **prov:Duty**'s. The **refinement** property is used to modify (restrict) the scope of **prov:Actions**, **prov:Assets** and **prov:Party**'s.

In the following EXAMPLE CODE 16 a **refinement** is introduced to the **assignee** of the first **permission**, which requires loan recipients to be of age (i.e. 18 years or older). To be able to do so, the **assignee** is defined as "**@type": "foaf:Person**". In the following line the person is identified with "**source" : "ex:curator2**" to be curator Hortensia at Institution 2. The **refinement** checks the age of Hortensia by defining an operation that states that a **foaf:Person**'s age (**foaf:age**) is required to be greater than the integer number "17", that is, it needs to be "18" or higher. If that equation is true, then the geographic constraint will be evaluated. If both restrictions are positively evaluated, then the **action** of **ex:shipping_to_3rdParty** corresponding to the linked resource **24_obligations_3rd party** can be performed (Fig. 21).

```
1  {
2      "@context" : [{ "odrl": "http://www.w3.org/ns/odrl.jsonld" }],
3      "@graph" : [ {
4          "@type" : "odrl:Agreement",
5          "@id" : "ex:inst1_inst2_loan_contract",
6          "profile" : "ex:biodivDomain_profile",
7          "permission" : [{
8              "target" : "ex:physicalSpecimen1",
9              "assigner" : "ex:institution1_holding",
10             "assignee" : [{
11                 "@type" : "foaf:Person",
12                 "source" : "ex:curator2",        // specifying the delegate
13                 "refinement" : [{                // restricting age to 18+
14                     "leftOperand" : "foaf:age",
15                     "operator" : "gt",           // greater than
16                     "rightOperand" : {
17                         "@value" : "17",
18                         "@type" : "xsd:integer"
19                     }
20                 }]
21             }],
22             "constraint" : [{
23                 "leftOperand" : "spatial",
24                 "operator" : "isPartOf",
25                 "rightOperand": "http://www.wikidata.org/entity/Q458",
26                 "comment" : "European Union"
27             }],
28             "action" : "ex:shipping_to_3rdParty"
29         }],
30         "permission" : [{
31             "target" : "ex:physicalSpecimen1",
32             "assigner" : "ex:institution1_holding",
33             "assignee" : "ex:institution2_requesting",
34             "action" : "ex:physical_analysis"
35         }]
36     } ]
37  }
```

Figure 21. doi

EXAMPLE CODE 16: Defining a context-dependent decision - adding an age restriction

Step 6: Moving towards an operational framework: exploring the policy life-cycle with inheritance between policies

In an operational setting, as for example ELViS and the DiSSCo technical infrastructure, policies and events do not exist in isolation. Hence, their machine-actionable representation as program code will not be called and run in isolation, too.

In the here explored use case of an ELViS loan transaction, the code describing **odrl:Policies** is populated with links to information that has been captured by PROV-constructs, which in their first function are used to record all activities and lay down a

provenance trail over time. At the same time, due to this function they therefore are the go-to resources for linking to the most up-to-date versions of digital objects and (meta)data. These up-to-date versions can then be used to construct conditions in the form of **odrl:Policies**.

Furthermore, **odrl:Policies** complement each other, exchange information and function in interaction. Already within the considered loan use case they evolve and introduce a life-cycle perspective. A loan environment can be considered to start with a **prov:Agent** designing a loan policy in the form of an **odrl:Offer** and offer **prov:Entities** (physical collection specimens) on the ELViS platform as available for loan. Another **prov:Agent** interested in receiving a specimen on loan, can design an **odrl:Request** (a subtype of **odrl:Policy**). Generally, this will mean that they fill out a standard request form on the ELViS system with the specifics of their loan request. The loan system will then bring together both the **odrl:Offer** and the **odrl:Request** policies and create from them a concrete loan contract, an **odrl:Agreement**. This newly minted **odrl:Agreement** policy will inherit its information from the two parental policies (**odrl:Offer** and **odrl:Request**). Based on the information and conditions inherited, the ELViS loan system will evaluate if the loan request fulfills the requirements associated with the requested specimen(s). If it does, the contract can be manually accepted by the involved **prov:Agents** and the loan be realised. In this way, **odrl:Offer** and **odrl:Request** policies evolved into an **odrl:Agreement** contract.

The following code example (EXAMPLE CODE 17) provides an overview over the elements of the policy life-cycle (Fig. 22). Several open questions currently remain regarding the details of inheritance, for example when information and conditions overlap and need to be merged, as well as concerning approaches for identifying and accessing specific properties within the policy objects. Thus, the example code exemplifies a general framework that can be further developed.

9.4 Next steps: Further development of the openDS specification and adopting ODRL for the biodiversity domain

The Kunming-Montreal Global Biodiversity Framework (GBF; Conference of the Parties to the Convention on Biological Diversity 2022) adopted in December 2022 by the parties to the UN Convention on Biological Diversity (CBD) sets global biodiversity conservation into an overarching context of ethical and socio-political principles that provide conditions for human actions and interactions (CBD 2022, Section C). Most of these principles are of direct application and impact to biodiversity data, data infrastructure design and maintenance, as well as knowledge management. For example, biodiversity data and knowledge services are called to consider different value systems (IPBES-Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services et al. 2019), be founded on a human rights-based approach (UNSDG-HRWG United Nations Sustainable Development Group Human Rights Working Group 2003), enable free, prior and informed consent and the right to participatory decision-making by indigenous peoples and local communities (IPLCs; United Nations General Assembly 2007), as well as support

collaboration of all, providing implementation for a whole-of-government and whole-of-society approach.

Figure 22. doi

EXAMPLE CODE 17: The policy life-cycle: an offer (with links to PROV objects) and request (with links to PROV objects) evolve into a contractual agreement as the basis for a machine-actionable assessment, including a final human decision step, in preparation of and leading up to a loan event.

IPLCs have already been working for some time on transforming such principles into operational implementations. Members of the communities developed the CARE principles (Carroll et al. 2021) and have started to build systems of concrete flags that can be

associated with physical objects, e.g. collection specimens, and data, e.g. genomic data. One such system are the Local Contexts labels, including Traditional Knowledge and Biocultural labels (https://localcontexts.org/).

The approaches that have been developed by and for IPLCs can provide inspiration, since biodiversity data infrastructures and natural science collections are embedded in highly cooperative social contexts that involve a wide range of stakeholders and rights holders, which at least in part have distinct and sometimes divergent interests.

These social structures are characterised by dynamically changing and sometimes complex conditions arising from multi-layered legal, ethical and social contexts, multiple interacting agents in a variety of roles and sets of situation-specific restrictions and decision processes. Examples are annotations of taxon identifications, and updates to a physical object's associated collection management information by members of the biodiversity sciences and collections communities. In addition, the same physical object might be subject to contracts, policies, regulations and laws at all legislative levels, from civil contracts between two parties to local, national and international legislation, including bilateral and multilateral treaties.

Thus, information on conditions associated with physical specimens and biodiversity data is and increasingly will be multi-faceted, structured and interlinked, dynamic and context-dependent. The modular structure of the openDS specification will enable the integration of one or more classes that store such information. In next steps, the overall structure as well as internal details of these classes will need to be clarified and developed. Together the Biodiversity Permit/Contract Typology and the Typology of Legal/Contractual Terms for Biodiversity Specimens provide a comprehensive and information-rich starting point for structuring data and information on conditions within the openDS specification.

One higher-level split that suggested itself already during this work cycle differentiates between, on one hand, odrl:Set and odrl:Offer policies spanning open a context of "lawful handling" and, on the other hand, odrl:Agreement policies determining concrete "contract conditions" between uniquely identified parties for specific assets (specimens). The first category provides information on conditions derived from applicable laws and regulations. Specifically for existing collection objects, due to resource limitations, it is practically impossible for natural science collections to retrospectively provide sufficient information for all objects. In addition, for a large part of old collection objects, legal conditions are not readily available and cannot be reliably inferred automatically. Furthermore, many common species might not have any legislation associated with them. Entries to the second category storing contractual information might be more realistically provided, at least for contracts covering future activities and transactions, as e.g. loan transactions. Most transactions are associated with contracts that are concrete, well-delimited and state conditions explicitly. Here, the decision is whether all available contractual information should be reflected in the DiSSCo technical infrastructure and e.g. its ELViS module.

Some challenges that the working group encountered and discussed during the development of the Typology of Legal/Contractual Terms for Biodiversity Specimens were

associated with use cases in which conditions are linked to past (e.g. collecting from nature) versus future (e.g. loan, use) events, as well as use cases that require the identification of implicit permissions. Implicit conditions can be associated with specimens and/or situations, which are characterised by an absence of existing policies or regulations.

The SYNTHESYS+ working group focused exclusively on legal and contractual conditions associated with physical specimens to build a starting point for continuing development. Participants of the MOBILISE-SYNTHESYS+ workshop in the fall of 2021 pointed out that legal conditions apply to a wide range of both analogue and digital transactions. It is clear that, compared to physical specimens, digital objects, (meta)data and processes exist in entirely different legal backgrounds, which currently are very dynamic and rapidly evolving, cp. e.g. the ongoing discussions and negotiation process on "Digital Sequence Information" within the CBD, and the latest UNCTAD report on transnational data flows (UNCTAD United Nations Conference on Trade and Development 2021). Understanding, differentiating and connecting the interrelated realms of physical and digital objects will be an important objective of upcoming work. It will increasingly require transdisciplinary collaboration with legal experts.

Further work is also needed to investigate the capabilities and limitations of applying ODRL and its extensions for the design and construction of normative statements for data and services required by and commonly occurring within the biodiversity sector. Very likely existing rights expression languages (RELs), with ODRL being one ontology system, will need to be modified, expanded and adapted to the specifics of the biodiversity sector with its own set of widely varied legal, ethical and social contexts.

Kebede et al. 2021 pointed out and discussed several challenges that they encountered while exploring ODRL for use cases arising in connection with data sharing infrastructures in the health and logistics sector. The limitations considered by them concern, e.g., the delegation of compound rights, in which some rights are passed on, while others are restricted or revoked, specifically in situations of unequal power; the handling of newly created entities from existing assets; normative statements pertaining to outcomes (e.g. of "use") versus the process-based approach taken by ODRL; agent granularity and the diversity of their roles; a dynamic user-driven consent management; and transitions within the policy- and rule-lifecycles.

Kebede et al. 2021 suggest eFlint (Binsbergen et al. 2020) as an alternative or complementary approach to ODRL. They describe and highlight eFlint as an action-based policy language. In addition, they point to several more RELs and suggest a review of Hohfeld's "Fundamental legal conceptions" (Hohfeld 1917), which introduces and informs about the fundamentals of legal decision processes. This provides essential background for investigations and work towards a decision for a fully applicable and functional REL for defining and implementing normative statements and their conditions for data and services.

In conclusion, the investigation of the ODRL ontology, its applicability and limits, as well as explorations of additional existing REL standards and ontologies will need to be continued. Furthermore, a comprehensive structural framework for integrating information about

conditions into the openDS specification will need to be developed, considering (known) challenges and a wider range of use cases. For this, the interconnections between and potential inheritance of conditions associated with physical and digital entities will need to be developed.

The goal of all these developments is the integration of conditions into events and transactions that are representing the dynamic and evolving nature of physical and digital biodiversity records as captured in and by biodiversity data infrastructures.

10 Future development

Two lines of action seem to be obvious for future development. First, a process among a group of like-minded institutions to decide, on one hand, the extent to which they want to implement the Typology of Legal/Contractual Terms for Biodiversity Specimens and, on the other hand, to decide what kind of updates and amendments it might need. Updates and amendments could be based on the needs of their own staff or typical stakeholders respectively. Amendments could include fossil, geological and mineralogical objects. A pilot implementation carried out for a number of different CMS and different collection types could provide additional insight in the applicability of the proposed terminology and potential gaps. In order to have the possibility to further develop the vocabulary in the future, the version published here has been transferred to the GGBN Wiki platform: (https://wiki.ggbn.org/ggbn/Permits_and_Contracts_and_Terms_for_Biological_Specimens) . The wiki system allows collaborative curation and transparent versioning as well as the documentation of discussions on specific topics. In this way we can ensure that current findings and developments can be documented dynamically.

The second line of action would be at the level of international platforms such as GGBN or the upcoming DiSSCo. Based on the institutional processes, both the Biodiversity Permit/Contract Typology and the Typology of Legal/Contractual Terms for Biodiversity Specimens could be implemented in GGBN, and DiSSCo could extend its basic model and make the latter Typology machine-actionable. These steps facilitate the management of juristic terms by worldwide access to the information in the Typology. It should also be considered to transfer the standardisation process to the TDWG Biodiversity Information Standards organisation, either in the form of a task group or a more general working group.

Glossary of terms used in the Typologies

ABS Clearing House

Global information portal (https://absch.cbd.int) developed under the [Convention on Biological Diversity (CBD)](#) and further elaborated in the [Nagoya Protocol (NP)](#) to make information available on national contacts (especially National Focal Points and Competent National Authorities), national legislation and other matters relevant to [Access and Benefit Sharing](#) and the [Nagoya Protocol](#)

Access and Benefit-Sharing (ABS)

A system based on public international law that outlines the way in which genetic resources or (where applicable) traditional knowledge associated with such resources is accessed and how the benefits that result from the utilisation of such resources and associated traditional knowledge are shared with the countries and/or indigenous peoples and local communities providing them.

Access to genetic resources or associated traditional knowledge

The acquisition of genetic resources or associated traditional knowledge from the country that has sovereign right over those resources (providing country). Note that this term may be used differently by some countries or organisations. The EU Regulation defines access as 'the acquisition of genetic resources or of traditional knowledge associated with genetic resources in a Party to the Nagoya Protocol'. Both the Convention on Biological Diversity (CBD) and - in more detail - the Nagoya Protocol (NP) contain provisions for granting access to genetic resources.

Biodiversity Beyond National Jurisdiction (BBNJ)

An international agreement, currently at the time of writing this glossary in 2022 under negotiation at the United Nations, on the conservation and sustainable use of marine biological diversity of areas beyond national jurisdiction.

Bioprospecting

The search for potentially valuable genetic data and biochemical compounds in biodiversity for the purpose of developing economically valuable products for different applications (e.g. pharmaceutical, cosmetic, agricultural).

The UNDP defined bioprospection "Biodiversity prospecting or bioprospecting is the systematic search for biochemical and genetic information in nature in order to develop commercially-valuable products for pharmaceutical, agricultural, cosmetic and other applications. Bioprospecting activities must comply with the definition of utilisation of genetic resources of the Nagoya Protocol or as stated in the national law or policy. The Nagoya Protocol applies to the utilisation of genetic resources and their derivatives" (https://www1.undp.org/content/dam/sdfinance/doc/Bioprospecting%20_%20UNDP.pdf, https://de.scribd.com/document/505930765/Bioprospecting-UNDP)

Convention on Biological Diversity (CBD)

International agreement designed to promote three goals, the "conservation of biological diversity, the sustainable use of its components and the fair and equitable sharing of the benefits arising out of the utilization of genetic resources" (https://www.cbd.int/convention/text/). The agreement was adopted by the states that participated in the 1992 UN Conference on Environment and Development in Rio de Janeiro; it entered into force on 29 December 1993.

Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES)

International agreement designed to ensure that international trade in wild fauna and flora does not deteriorate the situation of endangered or strongly exploited species. The Convention on International Trade in Endangered Species of Wild Fauna and Flora (1973) resulted from a resolution by the World Conservation Union (IUCN) in 1963; it entered into force on 01 July 1975.

Environmental sample

A collected volume of water, soil, sediment, or any other material containing living or dead organisms, or genetic material such as DNA.

Genetic resources (GR)

Term identified in the Convention on Biological Diversity (CBD) and refers to all "genetic material of actual or potential value", thus encompasses "any material of plant, animal. microbial or other origin containing functional units of heredity" that is potentially valuable to humans. Genetic resources can be taken from the wild, domesticated or cultivated. They may be sourced from natural environments (in situ) or human-made collections (ex situ) (e.g. botanical gardens, gene banks, seed banks and microbial culture collections).

International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA)

Also referred to as the Plant Treaty or Seed Treaty. International agreement designed to promote the "conservation and sustainable use of plant genetic resources for food and agriculture and the fair and equitable sharing of the benefits arising out of their use, in harmony with the Convention on Biological Diversity, for sustainable agriculture and food security" (https://www.fao.org/3/i0510e/i0510e.pdf). It entered into force on 29 June 2004.

Nagoya Protocol (NP)

Short for "The Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization". An international agreement related to the Convention on Biological Diversity (CBD) , which primarily aims at sharing the benefits arising from the utilisation of genetic resources in a fair and equitable way (https://www.cbd.int/abs/text/). It entered into force on 12 October 2014.

Pandemic Influenza Preparedness framework (PIP)

International instrument that aims to "improve pandemic influenza preparedness and response" by improving and strengthening a system for the global sharing "of H5N1 and other influenza viruses with human pandemic potential and access to vaccines and sharing of other benefits" (https://apps.who.int/iris/rest/bitstreams/1351857/retrieve). It was negotiated by Member States of the World Health Organization (WHO) and entered into force on 24 May 2011.

Phytosanitary

Scientific and regulatory frameworks relating to plant health, including the control of plant pests or pathogens.

Reintroduction

The intentional release of a species from captivity (e.g. zoo, botanical garden, seed bank) in an area inside its indigenous range from which it has disappeared.

Translocation

The intentional movement of a species within its indigenous range to an area where it has disappeared.

Utilisation of genetic resources

To "conduct research and development on the genetic and/or biochemical composition of genetic resources, including through the application of biotechnology" ("biotechnology means any technological application that uses biological systems, living organisms, or derivatives thereof, to make or modify products or processes for specific use") as defined in the --> Nagoya Protocol (NP).

Acknowledgements

- the presenters, session moderators and participants at the MOBILISE and SYNTHESYS+ workshops, as well as the many colleagues who responded to our surveys and questionnaires.
- COST Action CA 17106 supported the joint MOBILISE-SYNTHESYS+ workshop on a loans and permits data standard for scientific collections.
- the European Union for funding our work through the SYNTHESYS+ Grant in the framework of the Horizon 2020 programme.

Funding program

Horizon 2020

Grant title

H2020-INFRAIA-2018-2020 | SYNTHESYS PLUS | Grant Agreement Number 823827

Conflicts of interest

The **majority of European authors** is affiliated with collection institutions that are members of the Consortium of European Taxonomic Facilities (CETAF). Furthermore, **EH** is a member of the association's executive committee. CETAF is engaged in promoting taxonomy and natural history collections at the European level through collaboration and advocacy.

Chris Lyal provides consultancy to the European Commission with regard to the implementation of the EU regulation on Access and Benefit Sharing. **Edmund K. Schiller** was occasionally consulted by the Austrian Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology in his capacity as the museum's ABS-representative. **Suzete R. Gomes** is a Health Public Researcher in Brazil and was occasionally consulted in her capacity as a taxonomist and as a zoological collection curator in Oswaldo Cruz Foundation (FIOCRUZ), Brazil, to provide scientific and technical information in relation to regulatory and rule-based constraints on access to biological material, exchange and use in this country. **Jutta Buschbom** provides consultancy in the context of the science-policy interface to the UN Convention on Biological Diversity (CBD). Her engagement is focused on making accessible facts and knowledge of the biodiversity sciences for biodiversity conservation. Towards this goal, she provides advocacy for taxonomic research, and the analog and digital infrastructures developed and maintained by natural science collections and biodiversity informatics. She is an independent scientist and consultant. She provides scientific, technical and advisory services in a business context.

References

- Addink W, Hardisty A (2020) 'openDS' – Progress on the New Standard for Digital Specimens. Biodiversity Information Science and Standards 4 https://doi.org/10.3897/biss.4.59338
- Austria, Federal Ministry for Arts Culture the Civil Service and Sport, Bundesdenkmalamt (Ed.) (2023) Richtlinien Archäologische Maßnahmen. [Guidance on Archaeological Measures]. 6. Bundesdenkmalamt, 120 pp. [In German]. URL: https://www.bda.gv.at/dam/jcr:311e0b74-72ea-4e28-97a1-779e4ceed928/230209_Publikation_Richtlinien_Arch%C3%A4ologie_A4_BF.pdf
- Belgian Science Policy Office (2023) Biosafety and Biosecurity facts [internet]. https://bccm.belspo.be/legal/biosafety-biosecurity. Accessed on: 2023-1-01.
- Bill & Melinda Gates Foundation (2022) ki - Knowledge integration Global Health. https://www.kiglobalhealth.org/. Accessed on: 2022-11-01.
- Binsbergen LTv, Liu L, Doesburg Rv, Engers Tv (2020) eFLINT: a domain-specific language for executable norm specifications. In: Association for Computing Machinery (Ed.) Proceedings of the 19th ACM SIGPLAN International Conference on Generative Programming: Concepts and Experiences. GPCE 2020 - 19th International Conference on Generative Programming: Concepts & Experiences, virtual, November 2020. GPCE 2020 - 19th International Conference on Generative Programming: Concepts & Experiences, New York, 124-136 pp. [ISBN 978-1-4503-8174-1]. https://doi.org/10.1145/3425898.3426958
- Carroll SR, Herczog E, Hudson M, Russell K, Stall S (2021) Operationalizing the CARE and FAIR Principles for Indigenous data futures. Scientific data 8 (1): 108. https://doi.org/10.1038/s41597-021-00892-0
- Center for Digital Antiquity (2023) About the Digital Archaeological Record [internet]. https://www.tdar.org/about/. Accessed on: 2023-2-27.
- Commonwealth Scientific and Industrial Research Organisation CSIRO (Ed.) (2023) Atlas of Living Australia [internet]. http://www.ala.org.au. Accessed on: 2023-1-01.
- Conference of the Parties to the Convention on Biological Diversity (2022) Kunming-Montreal Global Biodiversity Framework. In: Secretariat of the Convention on Biological Diversity (Ed.) Decision 15/4 adopted by the Conference of the Parties to the Convention on Biological Diversity. 15. Meeting of the Conference of the Parties to the Convention on Biological Diversity, Part II, Montreal, 7.-19. December 2022. Secretariat of the Convention on Biological Diversity, Montreal, 14 pp. URL: https://www.cbd.int/doc/decisions/cop-15/cop-15-dec-04-en.pdf
- Conference of the Parties to the Convention on International Trade in Endangered Species of Wild Fauna and Flora (2002) Permits and Certificates. In: CITES Secretariat (Ed.) Resolution Conf. 12.3, Rev. CoP19. 12. Meeting of the Conference of the Parties to the Convention on International Trade in Endangered Species of Wild Fauna and Flora, Santiago, Chile, 4.-15. November 2002. CITES Secretariat, Geneva, 27 pp. URL: https://cites.org/sites/default/files/documents/COP/19/resolution/E-Res-12-03-R19.pdf
- Convention on Biological Diversity with Annexes (1992) opened for signature 05 June 1992. United Nations Treaty Series 1760: 79. URL: https://www.cbd.int/doc/legal/cbd-en.pdf

- Convention on International Trade in Endangered Species of Wild Fauna and Flora (1973) opened for signature on 03 March 1973. United Nations Treaty Series 993: 243. URL: https://cites.org/sites/default/files/eng/disc/CITES-Convention-EN.pdf, https://cites.org/sites/default/files/eng/app/2023/E-Appendices-2023-02-23.pdf
- COST Association AISBL (2018) Decision COST 017/18: Memorandum of Understanding for the implementation of the COST Action "Mobilising Data, Policies and Experts in Scientific Collections" (MOBILISE) CA17106 [internet] 15 pp. https://e-services.cost.eu/files/domain_files/CA/Action_CA17106/mou/CA17106-e.pdf. Accessed on: 2023-1-01.
- de Mestier A, Mulcahy D, Harris D, Korotkova N, Long S, Häffner E, Paton A, Schiller E, Leliaert F, Mackenzie-Dodds J, Fulcher T, Stahls G, von Rintelen T, Martín M, Lücking R, Williams C, Lyal C, Güntsch A, Aronsson H, Castelin M, Pielach A, Poczai P, Ruiz-León Y, Sanmartin Bastida I, Thines M, Droege G (2023) Policies Handbook on Using Molecular Collections. Research Ideas and Outcomes 9 https://doi.org/10.3897/rio.9.e102908
- DiSSCoTech (2020) What is a Digital Specimen? Technical posts about the design of the DiSSCo infrastructure [intnernet Blog], 31 Mar. 2020. https://dissco.tech/2020/03/31/what-is-a-digital-specimen/. Accessed on: 2022-11-01.
- Droege G, Barker K, Seberg O, Coddington J, Benson E, Berendsohn WG, Bunk B, Butler C, Cawsey EM, Deck J, Döring M, Flemons P, Gemeinholzer B, Güntsch A, Hollowell T, Kelbert P, Kostadinov I, Kottmann R, Lawlor RT, Lyal C, Mackenzie-Dodds J, Meyer C, Mulcahy D, Nussbeck SY, O'Tuama É, Orrell T, Petersen G, Robertson T, Söhngen C, Whitacre J, Wieczorek J, Yilmaz P, Zetzsche H, Zhang Y, Zhou X (2016) The Global Genome Biodiversity Network (GGBN) Data Standard specification. Database : the journal of biological databases and curation 2016 (baw125): 1-11. https://doi.org/10.1093/database/baw125
- Duke University Library Digital Repository Software Development and Integration Service (2023) About MorphoSource [internet]. https://ms1.morphosource.org/About/home. Accessed on: 2023-2-27.
- Eichert S, Richards N (Eds) (2022) THANADOS version 4.2 [internet]. https://thanados.net. Accessed on: 2022-11-01.
- European Commission (Ed.) (2022) Knowledge for Policy - Knowledge Center for Biodiversity [internet]. https://knowledge4policy.ec.europa.eu/biodiversity_en. Accessed on: 2023-1-31.
- European Organization for Nuclear Research, OpenAIRE (2013) Zenodo [internet]. https://zenodo.org/. Accessed on: 2023-2-27.
- Food and Agriculture Organization of the United Nations, International Plant Protection Convention Secretariat (Eds) (2023) Legislation: Phytosanitary Requirements/Restrictions/Prohibitions (1-713 of 713) [internet]. https://www.ippc.int/en/countries/all/legislation/. Accessed on: 2023-2-28.
- GBIF.org (2021) Community Consultation: Converging Digital Specimens and Extended Specimens [internet], 17 Feb. 2021. https://www.gbif.org/news/x7u6tmOhyfTE6rM3FQgIK/community-consultation-converging-digital-specimens-and-extended-specimens. Accessed on: 2022-11-01.
- German Museums Association (Ed.) (2021) Guidelines Care of Human Remains in Museums and Collections. Deutscher Museumsbund e.V., Berlin (Germany), 148 pp. [In

English]. URL: https://www.museumsbund.de/wp-content/uploads/2021/07/dmb-leitfaden-umgang-menschl-ueberr-en-web-20210625.pdf [ISBN 978 3 9822232 4 7]

- Global Genome Biodiversity Network, GGBN Secretariat, GGBN Technical Management (Eds) (2023) The GGBN Data Portal [internet]. https://www.ggbn.org/ggbn_portal/search/index. Accessed on: 2023-2-28.
- Hardisty A (2020) Q101 What is a Digital Specimen (DS). Frequently asked questions (FAQ) about Digital Specimens [internet], 17 Nov. 2020. https://github.com/DiSSCo/openDS/blob/master/faqs/faqds.md. Accessed on: 2022-11-01.
- Hardisty A, Saarenmaa H, Casino A, Dillen M, Gödderz K, Groom Q, Hardy H, Koureas D, Nieva de la Hidalga A, Paul D, Runnel V, Vermeersch X, van Walsum M, Willemse L (2020) Conceptual design blueprint for the DiSSCo digitization infrastructure - DELIVERABLE D8.1. Research Ideas and Outcomes 6 (e54280): 1-164. https://doi.org/10.3897/rio.6.e54280
- Hardisty A, Buschbom J, Zimkus B (2021) 'Digital/Extended Specimen: 8. Meeting legal/regulatory, ethical and sensitive data obligations', Converging Digital Specimens and Extended Specimens - towards a global specification for data integration. [online forum comment], 17 May 2021. https://discourse.gbif.org/t/8-meeting-legal-regulatory-ethical-and-sensitive-data-obligations/2665. Accessed on: 2022-11-01.
- Hardisty A, Ellwood E, Nelson G, Zimkus B, Buschbom J, Addink W, Rabeler R, Bates J, Bentley A, Fortes JB, Hansen S, Macklin J, Mast A, Miller J, Monfils A, Paul D, Wallis E, Webster M (2022) Digital Extended Specimens: Enabling an Extensible Network of Biodiversity Data Records as Integrated Digital Objects on the Internet. BioScience 72 (10): 978-987. https://doi.org/10.1093/biosci/biac060
- Harrison PW, Ahamed A, Aslam R, Alako BTF, Burgin J, Buso N, Courtot M, Fan J, Gupta D, Haseeb M, Holt S, Ibrahim T, Ivanov E, Jayathilaka S, Balavenkataraman Kadhirvelu V, Kumar M, Lopez R, Kay S, Leinonen R, Liu X, O'Cathail C, Pakseresht A, Park Y, Pesant S, Rahman N, Rajan J, Sokolov A, Vijayaraja S, Waheed Z, Zyoud A, Burdett T, Cochrane G (2020) The European Nucleotide Archive in 2020. Nucleic Acids Research 49 https://doi.org/10.1093/nar/gkaa1028
- Haston E, Chapman C (2022) MIDS: The digitisation standard for Natural Science collections. Biodiversity Information Science and Standards 6 https://doi.org/10.3897/biss.6.94604
- Hohfeld WN (1917) Fundamental Legal Conceptions as Applied in Judicial Reasoning. The Yale Law Journal 26 (8): 710-770. https://doi.org/10.2307/786270
- India National Biodiversity Authority (2021) Internationally Recognized Certificate of Compliance ABSCH-IRCC-IN-255561-1 [internet]. https://absch.cbd.int/en/database/IRCC/ABSCH-IRCC-IN-255561. Accessed on: 2022-5-01.
- International Council of Museums (2017) ICOM Code of Ethics for Museums. International Council of Museums, Paris (France), 44 pp. [In English]. URL: https://icom.museum/wp-content/uploads/2018/07/ICOM-code-En-web.pdf [ISBN 978-92-9012-420-7]
- International Treaty on Plant Genetic Resources for Food and Agriculture (2001) opened for signature 03 November 2001. United Nations Treaty Series 2400: 303. URL: https://www.fao.org/3/a-i0510e.pdf
- IPBES-Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, Brondízio ES, Settele J, Díaz S, Ngo HT (2019) Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform

on Biodiversity and Ecosystem Services. 1. IPBES secretariat, Bonn, 1085 pp. [ISBN 978-3-947851-20-1] https://doi.org/10.5281/zenodo.6417333

- Islam S, Hardisty A, Addink W, Weiland C, Glöckler F (2020) Incorporating RDA Outputs in the Design of a European Research Infrastructure for Natural Science Collections. Data Science Journal 19 (50): 1-14. https://doi.org/10.5334/dsj-2020-050

- Islam S, Addink W (2023) DiSSCo/European Loans and Visits System ELViS: v1.0.0. Zenodo https://doi.org/10.5281/zenodo.8232436

- Kebede M, Sileno G, Van Engers T (2021) A Critical Reflection on ODRL. In: Rodriguez-Doncel V, Palmirani M, Araszkiewicz M, Casanovas P, Pagallo U, Sartor G (Eds) AI Approaches to the Complexity of Legal Systems XI-XII. AICOL International Workshops 2018 and 2020: AICOL-XI@JURIX 2018, AICOL-XII@JURIX 2020, XAILA@JURIX 2020, Groningen, Brno, 12 December 2018, 09 December 2020. Springer, Cham, Lecture Notes in Computer Science (Lecture Notes in Aritficial Intelligence), 13048, 48-61 pp. [ISBN 978-3-030-89811-3]. https://doi.org/10.1007/978-3-030-89811-3_4

- Kenya National Environment Management Authority (2019) Internationally Recognized Certificate of Compliance ABSCH-IRCC-KE-242932-1 [internet]. https://absch.cbd.int/en/database/IRCC/ABSCH-IRCC-KE-242932/1. Accessed on: 2022-5-01.

- Kenya National Environment Management Authority (2020a) Internationally Recognized Certificate of Compliance ABSCH-IRCC-KE-253313-1 [internet]. https://absch.cbd.int/en/database/IRCC/ABSCH-IRCC-KE-253313/1. Accessed on: 2022-5-01.

- Kenya National Environment Management Authority (2020b) Internationally Recognized Certificate of Compliance ABSCH-IRCC-KE-253235-1 [internet]. https://absch.cbd.int/en/database/IRCC/ABSCH-IRCC-KE-253235/1. Accessed on: 2022-5-01.

- Krug C, Obura D, Gonzales A, Shannon L, Londono-Murcia M, Radulovici A, Millette K, Rankovic A, Archer E, Armah FA, Bax N, Chaudhari K, Kutchhi S, Kutchhi A, Costello M, Davalos L, Roque F, Declerck F, Dee LE, Essl F, Ferrier S, Genovesi P, Guariguata M, Hashimoto S, Speranza CI, Isbel F, Kok M, Lavery S, Leclere D, Loyola R, Lwasa S, McGeoch M, Mori AS, Nicholson E, Ochoa-Quintero JM, Öllerer K, Polasky S, Rondinini C, Schroer S, Selomane O, Shen X, Strassburg B, Sumaila R, Tittensor DP, Turak E, Urbina Gonzalez LF, Vallejos M, Vazquez-Dominguez E, Verburg PH, Visconti P, Woodley S, Xu J (2022) Transformative actions on all drivers of biodiversity loss are urgently required to achieve the global goals by 2050. In: Secretariat of the Convention on Biological Diversity, bioDISCOVERY program of Future Earth, Secretariat of the Group on Earth Observations Biodiversity Observation Network-GEO BON (Eds) Expert input to the post-2020 Global Biodiversity Framework [internet]. Secretariat of the Convention on Biological Diversity, Montreal, 183 pp. URL: https://www.cbd.int/doc/c/5735/c241/efeeac8d7685af2f38d75e4e/sbstta-24-inf-31-en.pdf

- Lebo T, McGuinness D, Sahoo S (2013) PROV-O: The PROV Ontology [internet]. https://www.w3.org/TR/2013/REC-prov-o-20130430/. Accessed on: 2023-1-31.

- Lendemer J, Thiers B, Monfils AK, Zaspel J, Ellwood ER, Bentley A, LeVan K, Bates J, Jennings D, Contreras D, Lagomarsino L, Mabee P, Ford LS, Guralnick R, Gropp RE, Revelez M, Cobb N, Seltmann K, Aime MC (2020a) The Extended Specimen Network: A Strategy to Enhance US Biodiversity Collections, Promote Research and Education. BioScience 70 (1): 23-30. https://doi.org/10.1093/biosci/biz140

- Lendemer J, Thiers B, Monfils AK, Zaspel J, Ellwood ER, Bentley A, LeVan K, Bates J, Jennings D, Contreras D, Lagomarsino L, Mabee P, Ford LS, Guralnick R, Gropp RE, Revelez M, Cobb N, Seltmann K, Aime MC (2020b) Corrigendum: The Extended

Specimen Network: A Strategy to Enhance US Biodiversity Collections, Promote Research and Education. BioScience 70 (2): 195. https://doi.org/10.1093/biosci/biz165

- Local Contexts (2023) Grounding Indigenous Rights [internet]. Local Contexts, Inc., New York. Accessed on 2023-1-31. URL: https://localcontexts.org/

- Mulligan C, Boyer D, Turner T, Delson E, Leonard W (2022) Data sharing in biological anthropology. American Journal of Biological Anthropology 178 (Suppl.74): 26-53. https://doi.org/10.1002/ajpa.24499

- Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity (2010) opened for signature on 29 October 2010. United Nations Treaty Series 3008: 3. URL: https://www.cbd.int/abs/doc/protocol/nagoya-protocol-en.pdf

- Peru Servicio Nacional Forestal y de Fauna Silvestre (2020a) Internationally Recognized Certificate of Compliance ABSCH-IRCC-PE-249103-1 [internet]. https://absch.cbd.int/en/database/IRCC/ABSCH-IRCC-PE-249103/1. Accessed on: 2022-5-01.

- Peru Servicio Nacional Forestal y de Fauna Silvestre (2020b) Internationally Recognized Certificate of Compliance ABSCH-IRCC-PE-249106-1 [internet]. https://absch.cbd.int/en/database/IRCC/ABSCH-IRCC-PE-249106/1. Accessed on: 2022-5-01.

- Peru Servicio Nacional Forestal y de Fauna Silvestre (2021) Internationally Recognized Certificate of Compliance ABSCH-IRCC-PE-254865-1 [internet]. https://absch.cbd.int/en/database/IRCC/ABSCH-IRCC-PE-254865/1. Accessed on: 2022-5-01.

- Sanderson R, Ciccarese P, Young B (Eds) (2016) Web Annotation Ontology [internet]. https://www.w3.org/ns/oa. Accessed on: 2023-1-31.

- Schindel D, Cook J (2018) The next generation of natural history collections. PLOS Biology 16 (7): 1-8. https://doi.org/10.1371/journal.pbio.2006125

- Schultes E, Wittenburg P (2019) FAIR Principles and Digital Objects: Accelerating Convergence on a Data Infrastructure. In: Manolopoulos Y, Stupnikov S (Eds) Communications in Computer and Information Science, 1003. XX International Conference Data Analytics and Management in Data Intensive Domains DAMID/RCDL, Moscow, 9-12 October 2018. Springer, Cham, 3-16 pp. https://doi.org/10.1007/978-3-030-23584-0_1

- Sharp C (2016) Overview of the Digital Object Architecture (DOA). An Internet Society Information Paper [internet], 26 Oct. 2016. https://www.internetsociety.org/resources/doc/2016/overview-of-the-digital-object-architecture-doa/. Accessed on: 2023-1-01.

- Smith BW, Rodríguez-Doncel V, Esteves B (Eds) (2023) ODRL Implementation Best Practices [internet]. https://w3c.github.io/odrl/bp/. Accessed on: 2023-2-03.

- SPNHC_Wiki_contributors (2022a) Permitting [internet]. https://spnhc.biowikifarm.net/w/index.php?title=Permitting&oldid=4140. Accessed on: 2022-10-14.

- SPNHC_Wiki_contributors (2022b) Main Page [internet]. https://spnhc.biowikifarm.net/w/index.php?title=Main_Page&oldid=4215. Accessed on: 2022-10-14.

- Stony Brook Primate Locomotion Laboratory (2022) Welcome to the Stony Brook Primate Locomotion Laboratory [internet]. https://primatelocomotion.org/. Accessed on: 2022-11-01.

- The Ancient Human Occupation of Britain Project (2016) The Ancient Human Occupation of Britain (AHOB) [internet]. https://www.ahobproject.org/. Accessed on: 2022-11-01.

- Thessen A, Woodburn M, Koureas D, Paul D, Conlon M, Shorthouse D, Ramdeen S (2019) Proper Attribution for Curation and Maintenance of Research Collections:

Metadata Recommendations of the RDA/TDWG Working Group. Data Science Journal 18 (54): 1-11. https://doi.org/10.5334/dsj-2019-054

- UNCTAD United Nations Conference on Trade and Development (2021) Digital Economy Report 2021 - Cross-border data flows and development: For whom the data flow. United Nations Publications, New York, 236 pp. [ISBN 978-92-1-113022-5] https://doi.org/10.18356/9789210058254

- United Kingdom, His Majesty's Government, Department for Culture Media and Sports (2005) Guidance for the Care of Human Remains in Museums [internet], 11 Oct. 2005. UK Department for Digital, Culture, Media & Sports, London, 36 pp. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/906601/GuidanceHumanRemains11Oct.pdf

- United Nations General Assembly (2007) Resolution 61/295: United Nations Declaration on the Rights of Indigenous Peoples (13 September 2007). A/RES/61/295. URL: http://www.un.org/development/desa/indigenouspeoples/wp-content/uploads/sites/19/2018/11/UNDRIP_E_web.pdf

- United Nations General Assembly (2015) Resolution 70/1: Transforming our world: the 2030 Agenda for Sustainable Development (21 October 2015). A/RES/70/1. URL: https://documents-dds-ny.un.org/doc/UNDOC/GEN/N15/291/89/PDF/N1529189.pdf?OpenElement

- UNSDG-HRWG United Nations Sustainable Development Group Human Rights Working Group (2003) The human rights based approach to development cooperation towards a common understanding among UN agencies. In: UNSDG United Nations Sustainable Development Group (Ed.) UNSDG Resources Library [internet]. Interagency Workshop on a Human Rights based Approach in the context of UN reform, 3-5 May 2003. URL: https://unsdg.un.org/sites/default/files/6959-The_Human_Rights_Based_Approach_to_Development_Cooperation_Towards_a_Common_Understanding_among_UN.pdf

- USDA APHIS United States of America Department of Agriculture Animal and Plant Health Inspection Service (2023) Manuals [internet]. https://www.aphis.usda.gov/aphis/ourfocus/planthealth/complete-list-of-electronic-manuals. Accessed on: 2023-2-28.

- Voisin J, Condemi S, Frayer D, Wolpoff M (2021) Human fossil teeth [internet], 16 Sept. 2021. http://anthropologicaldata.free.fr/. Accessed on: 2022-11-01.

- Watanabe M (2015) The Nagoya Protocol on Access and Benefit Sharing: International treaty poses challenges for biological collections. BioScience 65 (6): 543-550. https://doi.org/10.1093/biosci/biv056

- Webster MS (2017) The Extendedd Specimen. In: Webster MS (Ed.) The Extended Specimen: Emerging Frontiers in Collections-based Ornithological Research. 1. CRC Press, Boca Raton, 1-10 pp. [ISBN 9781315120454]. https://doi.org/10.1201/9781315120454

- Wetterstrand KA (2021) DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP) [internet]. www.genome.gov/sequencingcostsdata. Accessed on: 2023-1-05.

- Wikipedia contributors (2022a) Best Practice [internet]. https://en.wikipedia.org/w/index.php?title=Best_practice&oldid=1117358591. Accessed on: 2022-12-01.

- Wikipedia contributors (2022b) Internationalized Resource Identifier [internet]. https://en.wikipedia.org/w/index.php?

title=Internationalized_Resource_Identifier&oldid=1074712086. Accessed on: 2023-1-31.

- Wikipedia contributors (2023a) GitHub [internet]. https://en.wikipedia.org/w/index.php?title=GitHub&oldid=1137678221. Accessed on: 2023-2-27.
- Wikipedia contributors (2023b) JSON [internet]. https://en.wikipedia.org/w/index.php?title=JSON&oldid=1136730806. Accessed on: 2023-1-31.
- Wikipedia contributors (2023c) JSON-LD [internet]. https://en.wikipedia.org/w/index.php?title=JSON-LD&oldid=1132644378. Accessed on: 2023-1-31.
- Wikipedia contributors (2023d) URL [internet]. en.wikipedia.org/w/index.php?title=URL&oldid=1136437341. Accessed on: 2023-1-31.
- Wilkinson M, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J, da Silva Santos LB, Bourne P, Bouwman J, Brookes A, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo C, Finkers R, Gonzalez-Beltran A, Gray AG, Groth P, Goble C, Grethe J, Heringa J, 't Hoen PC, Hooft R, Kuhn T, Kok R, Kok J, Lusher S, Martone M, Mons A, Packer A, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S, Schultes E, Sengstag T, Slater T, Strawn G, Swertz M, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B (2016) The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3 (1): 1-9. https://doi.org/10.1038/sdata.2016.18
- Wittenburg P, Strawn G, Mons B, Boninho L, Schultes E (2019) Digital Objects as Drivers towards Convergence in Data Infrastructures. EUDAT B2SHARE, 24 pp. https://doi.org/10.23728/b2share.b605d85809ca45679b110719b6c6cb11
- Zimkus B, Ford L, Morris P (2021) The Need for Permit Management within Biodiversity Collection Management Systems to Digitally Track Legal Compliance Documentation and Increase Transparency About Origins and Uses. Collection Forum 35 (1): 1-20. https://doi.org/10.14351/0831-4985-35.1.1
- Zimkus BM, Ford LS (2014) Best practices for genetic resources associated with natural history collections: Recommendations for practical implementation. Collection Forum 28 (1-2): 77-112. https://doi.org/10.14351/0831-0005-28.1.77

Supplementary materials

Suppl. material 1: The Biodiversity Permit/Contract Typology - 38 Document Types
doi

Authors: Edmund Schiller, Karin Wiltschke, Jutta Buschbom, Eva Häffner, Frederik Leliaert, Breda Zimkus, John Dickie, Suzete Gomes, Chris Lyal, Daniel Mulcahy, Alan Paton, Gabi Droege
Data type: standardised vocabulary
Download file (134.83 kb)

Suppl. material 2: The Typology of Legal/Contractual Terms for Biodiversity Specimens - 87 terms `doi`

Authors: Edmund Schiller, Karin Wiltschke, Jutta Buschbom, Eva Häffner, Frederik Leliaert, Breda Zimkus, John Dickie, Suzete Gomes, Chris Lyal, Daniel Mulcahy, Alan Paton, Gabi Droege
Data type: standardised vocabulary
Download file (114.64 kb)

Suppl. material 3: Links to Australian legislation related to biodiversity `doi`

Authors: Edmund Schiller
Data type: links to webpages
Brief description: Not covered in this publication is Australian legislation related to biodiversity, the links in this table for websites on Australian legislation showcase the need of specialist (legal) knowledge to inform on topics that might be relevant for natural history collections
Download file (10.95 kb)

Suppl. material 4: Legal characteristics of every Term for Biodiversity Specimens `doi`

Authors: Edmund Schiller, Karin Wiltschke, Jutta Buschbom, Eva Häffner, Frederik Leliaert, Breda Zimkus, John Dickie, Suzete Gomes, Chris Lyal, Daniel Mulcahy, Alan Paton, Gabi Droege
Data type: standardised vocabulary
Brief description: Every legal/contractual term from the Typology of Legal/Contractual Terms for Biodiversity Specimens with its individual characteristics (generally described in chapter 7.2.2), and their potential lawful source(s)
Download file (245.92 kb)

Suppl. material 5: GGBN vocabulary for permits, version 1 `doi`

Authors: Global Genome Biodiversity Network, Gabi Droege, Katie Barker
Data type: standardised vocabulary
Brief description: This copy of the GGBN vocabulary for permits was downloaded on October 18th, 2022 from https://wiki.ggbn.org/ggbn/GGBN_Data_Standard_v1#GGBN_Permit_Vocabulary
Download file (242.31 kb)

Suppl. material 6: GGBN vocabulary for loans, version 1 `doi`

Authors: Global Genome Biodiversity Network, Gabi Droege, Katie Barker
Data type: standardised vocabulary
Brief description: This copy of the GGBN vocabulary for loans was downloaded on October 18th, 2022 from https://wiki.ggbn.org/ggbn/GGBN_Data_Standard_v1#GGBN_Loan_Vocabulary
Download file (148.78 kb)

Suppl. material 7: SPNHC Categories of Legal/Compliance Documentation `doi`

Authors: Society for the Preservation of Natural History Collections
Data type: standardised vocabulary
Brief description: This copy of the SPNHC Categories of Legal/Compliance Documentation was downloaded on October 18th, 2022 from https://spnhc.biowikifarm.net/wiki/Permitting#Categories_of_Legal.2FCompliance_Documentation
Download file (206.82 kb)

Suppl. material 8: Permit Names US & EU `doi`

Authors: Edmund Schiller
Data type: text
Brief description: Examples for permits standardised for EU-wide use, or permits issued by federal authorities of the USA
Download file (75.12 kb)

Suppl. material 9: Summary of JSON-LD code for the loan use case `doi`

Authors: Jutta Buschbom
Data type: JSON-LD code / text
Brief description: This supplementary file provides a summary of the example code written for a use case representing an ELViS loan transaction. An interactive visualisation can be generated at https://json-ld.org/playground/ by first removing all comments and then copy-pasting the example code into the "JSON-LD Input" field of the website. Once validated, the tab "Visualized" below the input area produces a visual representation that can be explored by clicking on the graphical elements.
Download file (100.14 kb)

Suppl. material 10: Questionnaire on using molecular biological collections `doi`

Authors: SYNTHESYS+ project task groups for tasks NA3.2 and NA3.3
Data type: questionnaire
Brief description: For this file/pdf the questionnaire/survey template was copied from https://www.surveymonkey.com/r/HQDGKLN on Sept. 6th, 2022, field functions were removed, and the initially stacked answering options are displayed here in a row for reducing the number of pages.
Download file (837.34 kb)

Endnotes

[*1] this term is explained in the Glossary