OPEN ACCESS

CrossMark

Research Article

Assessing the FAIR Digital Object Framework for Global Biodiversity Research

Sharif Islam[‡,§], James Beach[|], Elizabeth R. Ellwood[¶,#], Jose Fortes[¤], Larry Lannom[«], Gil Nelson[#], Beth Plale[»]

‡ Naturalis Biodiversity Center, Leiden, Netherlands
§ DiSSCo, Leiden, Netherlands
| University of Kansas, Lawrence, United States of America
¶ Natural History Museum of Los Angeles County, Los Angeles, United States of America
# Florida Museum of Natural History, Gainesville, United States of America
¤ University of Florida, Gainesville, United States of America
« Corporation for National Research Initiatives (CNRI), Reston, Virginia, United States of America
» Indiana University Bloomington, Bloomington, IN, United States of America

Reviewed v 1

Abstract

In the first decades of the 21st century, there has been a global trend towards digitisation and the mobilisation of data from natural history museums and research institutions. The development of national and international aggregator systems, which focused on data standards, made it possible to access millions of museum specimen records. These records serve as an empirical foundation for research across various fields. In addition, community efforts have expanded the concept of natural history collection specimens to include physical preparations and digital resources, resulting in the Digital Extended Specimen (DES), which also includes derived and related data. Within this context, the paper proposes using the FAIR Digital Object (FDO) framework to accelerate the global vision of the DES, arguing that FDO-enabled infrastructures can reduce barriers to the discovery and access of specimens, help ensure credit back to contributors and increase the amount of research that incorporates biodiversity data.

Keywords

FAIR data, FDO, natural history museums, natural science collections, biodiversity specimens, interoperablity, persistent identifiers, FAIR implementation, extended specimen network
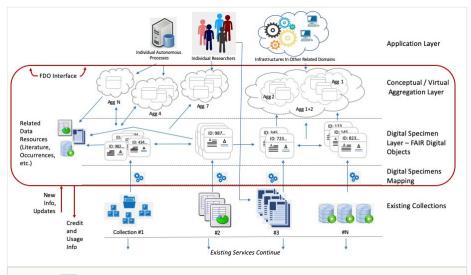
Introduction

The first decades of the 21st century have seen a massive increase in the digitisation and mobilisation of the data from hundreds of millions of specimens curated in thousands of natural history museums and research centres distributed around the world (Nelson and Ellis 2019). Driven by the need to make data from specimens more easily and widely accessible to scientists and to marshal those data for applied research addressing the global biodiversity crisis (Pimm et al. 2014), digitisation and the Internet have eliminated distances of space and time as barriers to data availability (Hedrick et al. 2020), which, for centuries, has separated researchers from remote specimens and their accompanying data. Within this data-driven ecosystem, the biodiversity community is invested in expanding the notion of the natural history collection specimen to include derivative preparations and metadata in addition to the original specimen (Webster 2017). In digital terms, this translates into an expansion from the digitised specimen record to an extended record which links derived and/or related data to the digital specimen record including CT scans, isotopes and even information discerned from artificial intelligence. Linked information about a specimen thus creates a rich extended digital object which we refer to as the Digital Extended Specimen (DES) (Hardisty et al. 2022). The Extended Specimen Network (Lendemer et al. 2020) is also used to describe such specimens and the foundation of such a network relies on a fragmented and global landscape of biological, geological and environmental data, which has emerged over time through various global and regional projects, datasets and databases (Bingham et al. 2017). These developments were driven by independent institutional advances in catalogue management (Nadim 2021), the adoption of accessible database technologies and the efforts to establish standards for describing specimens (Wieczorek et al. 2012, Groom et al. 2019). Considering the historical foundations and the utilisation of the Extended Specimen Network, the FDO framework can significantly contribute to widening access to fragmented specimen data and, consequently, bolster data stewardship and curation (Thomer et al. 2019). Moreover, when fully implemented, the DES can be instrumental in addressing the existential crises of unprecedented rates of species extinction, biodiversity loss and climate change (Corlett 2023).

This paper introduces and demonstrates the application of the FAIR Digital Objects (FDO) framework (Anders et al. 2023) to the DES through a thought exercise and practical use. Drawing on the work facilitated by the Data Foundation and Terminology Working Group in Research Data Alliance (RDA), as well as subsequent efforts within the FAIR Digital Objects Forum, we offer that FAIR Digital Objects, a technological abstraction and set of

services, is a viable approach to bringing greater FAIR-ness (Findability, Accessibility, Interoperability and Reusability) to the DES.

The foundational concept of the FDO framework is the digital object. While the term "*digital object*" can be applied to any digitised piece of information (Hui 2013), it has achieved recognition to connote a structured collection of data which can be acted upon programmatically, independent of any specific storage technologies (Kahn and Wilensky 2006, Goble et al. 2020, Harjes et al. 2020, Rozenberg 2021). Additionally, the FDO framework aligns with the Digital Object Architecture (DOA), an architectural model supporting interoperability amongst digital objects (Wittenburg et al. 2019, Wittenburg and Strawn 2019).

The FDO framework does not directly result in biodiversity data integration or immediately resolve interoperability challenges, but provides methods, mechanisms and objects within which those problems can be more readily and robustly solved (see Fig. 1 for a conceptual model of the framework applied to the DES). FDOs could be used to integrate and combine the ever-evolving informational bits making up a DES, regardless of the location of those bits. FDO could also (separately) be used as a protocol by which DES objects are retrieved. As the FDO framework is technology and implementation agnostic, we also envision one or more such protocols (for instance, to resolve digital object identifiers, locate the associated digital objects or run certain operations) to be supported to work with FDOs.

Figure 1. doi

A conceptual model of the FDO framework applied to the DES. This figure pinpoints the FDO interface with protocols that is needed for access to the objects. The figure also shows existing digital objects and Collections (as in museum collections) contributing to the digital extended specimen objects through create and update operations and receiving credit and usage information for their contributions.

The authors of this paper bring together their perspectives, based on involvement in museum collections, biodiversity aggregators, data fabric infrastructure research and biodiversity data usage. Starting in May 2020, we began a series of discussions to discern the guiding principles by which the FDO vision could accelerate the global realisation of the DES. We think that the FDO framework can reduce barriers to the discovery and access of biodiversity specimens and can facilitate the emergence of digital extended specimens. The FDO vision could increase the volume of research that can incorporate biodiversity data into its research questions while appropriating credit to the specimen sources.

This paper provides a conceptual model of the FDO framework applied to the DES, based on the consensus of the authors. In our subsequent work, we aim to detail the challenges of how specific existing DES data storage and management systems, such as relational databases and collections management systems, may or may not seamlessly fit into the FDO framework. This ongoing work will provide an opportunity to share experiences with concrete examples and empirical evidence.

Approach

The extended data of a physical specimen is held in diverse information sources, such as the Global Biodiversity Information Facility (GBIF), Geoscience Collections Access Service (GeoCASe), World Register of Marine Species (WoRMS), Barcode of Life Data System (BOLD) and Biodiversity Heritage Library (BHL) emerging from a variety of information models, data formats, application programming interfaces (API) and access controls. The FDO vision abstracts away the discord of multiple systems through a conceptual layer of digital objects that would be standardised everywhere across the Internet. In the simplest design, one FDO exists for each DES (in a 1:1 relation).

The implementation and application of the FDO framework (or any other framework) are closely linked to the schemas that define metadata and data. These schemas are built on rich data models supported by community culture, research-driven processes and agreements (Sansone et al. 2019). Before deploying the FDO framework, existing efforts concerning domain-specific data models, structures and ontologies must be fine-tuned and supported. The DiSSCo project (see below for more details) addresses this step through technical work packages that are focusing on open Digital Specimen (openDS) specification (Addink and Hardisty 2020). Additionally, the work on Minimum Information about a Digital Specimen (MIDS) is relevant, as it specifies and classifies the essential information elements that can be assigned to a specimen within a digitisation framework (Haston and Hardisty 2020). Further discussions have been driven by GBIF's new data model (Robertson et al. 2022), the Alliance for Biodiversity Knowledge and the BiCIKL project.

Networks of data are inherently dynamic and technologies evolve. A key first principle in the FDO framework is global referential integrity, a principle underpinned by the first FAIR principle: *F1. (Meta)data are assigned globally unique and persistent identifiers.* The topic of persistent identifiers (PIDs) has been well covered in the literature (see Meadows et al.

(2019), Juty et al. (2020), Hardisty et al. (2021)) and we will not revisit that discussion here, other than to emphasise that PIDs are the required starting point in constructing a global data space of FDOs. We cannot manage what we cannot reference.

To harmonise information management tasks and ensure consistency across all FDOs, the following key principles are applied:

- Every object has a globally unique, persistent and actionable identifier;

- Every object is typed i.e. classified against a specific definition of what the object represents and how it is represented;

- Every object has tightly associated metadata that describes it;

- Every object has a queryable set of operations that can be requested of it, as determined by its type;

- Every object can be addressed and accessed via a common protocol, for example, the Digital Object Interface Protocol (DOIP).

In the absence of implementation detail, the components and services minimally required to implement the FDO framework include:

- Persistent identifiers plus an identifier resolution system;

- A minimum set of metadata (known as PID Kernel records or FDO records);

- Defined digital object types accessible from a well-known set of type registries;

- Digital Object Repositories, aka "Object Servers", including repositories of metadata, aka "metadata registries";

- Mapping/brokering software and services to map existing data storage and management systems, such as relational databases and collections management systems to the FDO paradigm;

- An access protocol, such as DOIP, implemented by FDO repositories and applications.

In a fully built-out FDO framework for the DES, there will exist a fixed set of organisational functions. These responsibilities and control points can be distributed across multiple organisations or centralised in a few or exist in some mix of the two. These functions may consist of existing roles or new ones, with the same set of alternatives applying across consortia and standards bodies. Starting with an assumed set of existing records of some type, the top level organisational functions can be summarised as such:

- Establishing a standard set of types into which DESs are categorised and differentiated;

- Agreeing on PID regime(s) for identifying and resolving FDOs;

- Registering type information for new FDO types into a global type registry;

- Ensuring machine-actionable access to DES objects, including authorisation and authentication;

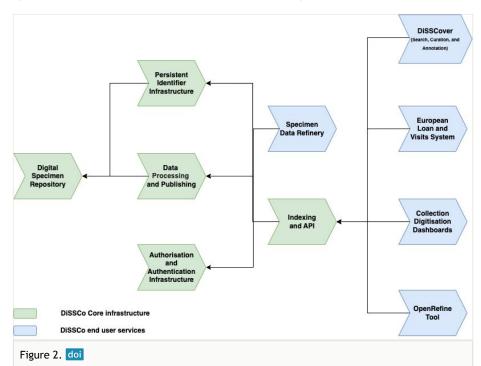- Providing backups or dark archives of FDOs for guaranteed persistence.

We acknowledge that the creation of specific FDOs accounts for much of the effort involved in implementing the FDO framework on top of existing practices. However, the prevalence of certain Collection Management Systems (CMSs), such as Specify and widely used standards, such as Darwin Core, will allow many of the early efforts to be reused and perhaps even built into subsequent CMS versions and other tooling. Standards will also play a role in the specific schemas used for DES, such as DiSSCo's proposed openDS. Obtaining agreement on one or a few such schemas will be essential for interoperability amongst different types of digital specimen objects.

Exemplar: DiSSCo

Distributed System of Scientific Collections (DiSSCo) is a research infrastructure in preparation for a portfolio of FAIR services along with capacity building and training to unify European natural science collections data. The FDO vision helps DiSSCo to apply the FAIR principles for natural science collections data use cases and services (De Smedt et al. 2020; Islam et al. 2020). DiSSCo is used here as an exemplar for the concepts and application of the FDO framework. As global efforts are increasing towards mass digitisation and extracting data at scale (Scott and Livermore 2021), it is important to understand the scope, context and different use cases of the various digital objects derived from and associated with the physical specimens. The conceptual design and implementation of DiSSCo utilises this scoping and use case exercises (Hardisty et al. 2020; Loo et al. 2023). In Fig. 1, we are envisioning DiSSCo in the Digital Specimen layer and for it to be part of the FDO interface.

The DiSSCo vision builds on the concept of the DES (which the project calls "Digital Specimen") -- a digital object acting as a digital surrogate on the internet for a specific physical specimen in a collection (Lannom et al. 2020). At the core of this implementation is the FDO foundational contribution of *abstraction*, which allows one to capture identification (via DOI or another type of persistent identifiers) and description of any entity (either a specimen, media object, machine agent or organisation) and further build services (Addink et al. 2023). The design decisions attempt to reach a balance between the flexibility of different specimens (for instance, ranging from marine specimen collections to botany to mineralogy) and provide structured descriptions that can be integrated into workflows, such as machine learning (Grieb et al. 2021, Davis 2023) or digital twinning applications (Schultes et al. 2022, Peters and Schindler 2023).

At the time of writing of this paper, the DiSSCo sandbox implementation is using open source components, such as Postgres, Kafka and Kubernetes, to create an agile, modular and scalable implementation (see Fig. 2) that can ingest data adhering to different standards (namely Darwin Core and ABCD(EFG)) and generate FDOs with a persistent identifier and structured attributes (Leeflang et al. 2022b). We illustrate the approach with an example Digital Specimen currently being considered within the DiSSCo project (see Fig. 3). We look specifically at the FDO records that provide minimal attributes for each digital object. In this implementation, these records are managed by a PID resolution system and are not stored in the respective biodiversity infrastructure.

Figure 2. doi

The envisioned DiSSCo data infrastructure and services. For more information, see https://www.dissco.eu/services.

A few notes about the Digital Specimen FDO record:

1. Digital Specimen identifies a unique digital specimen. The example FDO record contains the unique PID, the PID issuing organisation, the type of FDO ("Digital Specimen"), the location of the digital object, the licence as it applies to the metadata, PID lifecycle status (e.g. "active", "draft") and the organisation that hosts the specimen.

2. Two other FDO profiles (FDO profile describes the set of attributes in an FDO record) are also under consideration. Digital Media identifies a unique media object where the FDO record may differ from the Digital Specimen on a few specific elements that concern media objects like images, videos or

sound files and an Annotation FDO that includes the result of different annotation activities, such as comments and error correction both by human or machine enrichment processes (Leeflang et al. 2022a).

The structure and content of the DiSSCo FDO profile are evolving with ongoing discussions within DiSSCo and the FDO Forum. This dynamic development process ensures that the FDO profiles are continually refined and optimised to meet the diverse needs of the DiSSCo project and its stakeholders.

Figure 3. doi

DiSSCo Digital Specimen FDO record.

Discussion

We have introduced the FDO framework applied to the DES. Its adoption will have advantages and complications that will require a community effort. We address both here. The FDO framework's overriding advantage is that the existing global heterogeneity of collection information management and data repositories is pushed down a level of abstraction such that their design details need to be known only to those who are maintaining those systems. Additionally, the FDO framework simplifies the biodiversity community's networked data space in the following ways:

- Every digital object can be treated the same until it has to be treated differently to accomplish the specific purpose for which it was created. Consider the parallels with the Internet, in which each packet of data is routed from source to destination in the same way, regardless of the contents of the packets, which are examined and utilised only after they reach their destination.
- The mechanism for obtaining and interacting with DES digital objects is common across all sites, irrespective of the organisation, semantics, logic and design peculiarities of different information management and storage systems. This is analogous to the way that the Darwin Core simplified data exchange by obviating

community-level exposure to the internal structure of specimen data records in local information systems.

- Objects are self-describing in that they carry their type and access control information from one location to another, independently of whatever current system is making them available. That means constraints on their use or modification or attribution information move with the data to ensure that the publisher's intent is respected. Moving digital objects from one system to another does not in itself change access permissions or any other security details of the object.

The FDO framework can represent a dynamic representation of a specimen that, over time, accretes links and pointers to new scholarly treatments and research analyses. Seeded by museum collection data, DES digital objects will link to many types of related scientific and societal information not currently modelled or processed in Collections Management Systems (CMS) rooted in cataloguing and curation.

Our extensive discussions also ranged to address questions about the disruption that the technology might cause. Irrespective of the FDO framework, value-adding changes to digital objects derived from collection catalogue records will occur beyond the scope of CMS and curatorial practice. However, that is happening today as distributed researchers and automated methods update and annotate copies of specimen records in aggregation databases with no synchronisation with source CMSs. A global FDO framework could make it easier for annotated records to be linked back to the source CMS.

A global network of Digital Extended Specimens could attenuate the traditional status of a museum as the "Source of Authority" for information about their holdings. Specimen records in institutional CMS will continue to be the coin of the realm for collection curation and asset management, but, ultimately, many of those records may no longer be the authoritative, complete or up-to-date sources of information for the specimens they proxy.

Although it would be futile to predict the exact consequences of this technological change, the emergence of an FDO-enabled DES network will motivate researchers and their museum stakeholders to re-evaluate their roles and level of participation in the re-architected data community. Natural history museums represent a massive, distributed archive of Earth's biological and geological diversity. Once specimen data are pulsing through the pipelines and services of a networked DES, will there be incentives for collections' institutions to curate and update records in a re-partitioned specimen data space, organised along biogeographic or taxonomic themes and let the traditional collection cataloguing, curation and primary publishing paradigm give way? We are excited about exploring these questions with global stakeholders to find a sustainable implementation model.

Conclusion

The application of the FDO framework to the DES cannot solve all challenges in scientific practices and data standardisation. However, similar to the earlier technological

interoperability of the Internet, it can provide a needed layer to overcome the obstacles to data stewardship and standardisation of cross-disciplinary data. While our focus in this paper centres on the challenges related to specimens and extended specimens, it is worth noting that the concept of FDOs can be extended to encompass the broader field of biodiversity research.

The FDO framework is agnostic to the social and political challenges involved in global integration for facilitating easier research access. However, given the pressing issues of biodiversity and climate change, we are compelled to explore how data records can be transformed into FAIR Digital Objects, ensuring broader access and seamless interoperability. Implementing FAIR principles and adopting the FDO framework are keys to meeting the demands of the present times and maximising the potential for impactful research and collaboration across diverse scientific domains.

Acknowledgements

Conflicts of interest

A reviewer of this paper is affiliated with the DiSSCo Project.

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
- Addink W, Islam S, Dillen M, Güntsch A, Theocharides S (2023) Deliverable D7.1 Architecture Design for a pan-European PID system for Digital Specimens. ARPHA Preprints https://doi.org/10.3897/arphapreprints.e107168
- Anders I, Blanchi C, Broder D, Hellström M (2023) FAIR Digital Object Technical Overview. FAIR Digital Objects Forum. https://doi.org/10.5281/zenodo.7824714
- Bingham H, Doudin M, Weatherdon L, Despot-Belmonte K, Wetzel F, Groom Q, Lewis E, Regan E, Appeltans W, Güntsch A, Mergen P, et al. (2017) The biodiversity informatics landscape: elements, connections and opportunities. Research Ideas and Outcomes 3: 14059. https://doi.org/10.3897/rio.3.e14059
- Corlett R (2023) Achieving zero extinction for land plants. Trends in Plant Science 28 (8): 913-923. https://doi.org/10.1016/j.tplants.2023.03.019.
- Davis C (2023) The herbarium of the future. Trends in Ecology & Evolution 38 (5): 412-423. https://doi.org/10.1016/j.tree.2022.11.015

- De Smedt K, Koureas D, Wittenburg P (2020) FAIR Digital Objects for Science: From Data Pieces to Actionable Knowledge Units. Publications 8 (2). https://doi.org/10.3390/publications8020021
- Goble C, Cohen-Boulakia S, Soiland-Reyes S, Garijo D, Gil Y, Crusoe M, Peters K, Schober D (2020) FAIR Computational Workflows. Data Intelligence 2: 108-121. https://doi.org/10.1162/dint_a_00033
- Grieb J, Weiland C, Hardisty A, Addink W, Islam S, Younis S, Schmidt M (2021) Machine Learning as a Service for DiSSCo's Digital Specimen Architecture. Biodiversity Information Science and Standards 5: 75634. https://doi.org/10.3897/biss.5.75634
- Groom Q, Dillen M, Hardy H, Phillips S, Willemse L, Wu Z (2019) Improved standardization of transcribed digital specimen data. Database https://doi.org/10.1093/database/baz129
- 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 https://doi.org/10.3897/rio.6.e54280
- Hardisty AR, Addink W, Glöckler F, Güntsch A, Islam S, Weiland C (2021) A choice of persistent identifier schemes for the Distributed System of Scientific Collections (DiSSCo. Research Ideas and Outcomes 7: 67379. https://doi.org/10.3897/rio.7.e67379
- Hardisty AR, Ellwood ER, Nelson G, Zimkus B, Buschbom J, Addink W, Rabeler RK, Bates J, Bentley A, Fortes JAB, Hansen S, Macklin JA, Mast AR, Miller JT, Monfils AK, Paul DL, Wallis E, Webster M, et al. (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
- Harjes J, Link A, Weibulat T, Triebel D, Rambold G (2020) FAIR digital objects in environmental and life sciences should comprise workflow operation design data and method information for repeatability of study setups and reproducibility of results. Database 2020 https://doi.org/10.1093/database/baaa059
- Haston E, Hardisty A (2020) An Introduction to the Minimum Information about a Digital Specimen (MIDS) Digitisation Standard. Biodiversity Information Science and Standards 4 https://doi.org/10.3897/biss.4.59214
- Hedrick BP, Heberling JM, Meineke EK, Turner KG, Grassa CJ, Park DS, Kennedy J, Clarke JA, Cook JA, Blackburn DC, Edwards SV (2020) Digitization and the future of natural history collections. BioScience 70 (3): 243-251. https://doi.org/10.1093/biosci/biz163
- Hui Y (2013) What is a digital object? In: Halpin H, Monnin A (Eds) Philosophical Engineering: Toward a Philosophy of the Web. https://doi.org/10.1002/9781118700143.ch4
- 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 https://doi.org/10.5334/dsj-2020-050
- Juty N, Wimalaratne SM, Soiland-Reyes S, Kunze J, Goble CA, Clark T (2020) Unique, persistent, resolvable: Identifiers as the foundation of FAIR. Data Intelligence 2 https://doi.org/10.1162/dint_a_00025.

- Kahn R, Wilensky R (2006) A framework for distributed digital object services. International Journal on Digital Libraries 6 (2): 115-123. https://doi.org/10.1007/s00799-005-0128-x
- Lannom L, Koureas D, Hardisty A (2020) FAIR Data and Services in Biodiversity Science and Geoscience. Science Data Bank https://doi.org/10.57760/sciencedb.j00104.00072
- Leeflang S, Addink W, Theocharides S (2022a) Human and Machine Working Together towards High Quality Specimen Data: Annotation and Curation of the Digital Specimen. Biodiversity Information Science and Standards 6: 90987. https://doi.org/10.3897/biss.6.90987
- Leeflang S, Weiland C, Grieb J, Dillen M, Islam S, Fichtmueller D, Addink W, Haston E (2022b) DiSSCo Prepare Deliverable D6.2 Implementation and construction plan of the DiSSCo core architecture. DiSSCo Prepare https://doi.org/10.34960/50b9-kj05
- 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, et al. (2020) 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
- Loo T, Benvenuti M, Cecchi L, Innocenti G, Biaggini M, Bellucci L, Moggi Cecchi V, Di Vincenzo F, Rossi de Gasperis S (2023) DiSSCo Prepare Deliverable D9.6 Compilation of Construction Masterplan. DiSSCo Prepare https://doi.org/10.34960/cy1m-b238
- Meadows A, Haak L, Brown J (2019) "Persistent Identifiers: The Building Blocks of the Research Information Infrastructure&rdquo. Insights 32 (1): 9. https://doi.org/10.1629/uksg.457
- Nadim T (2021) The datafication of nature: data formations and new scales in natural history. Journal of the Royal Anthropological Institute 27 (S1): 62-75. https://doi.org/10.1111/1467-9655.13480
- Nelson G, Ellis S (2019) The history and impact of digitization and digital data mobilization on biodiversity research. Philosophical Transactions of the Royal Society B 374 (1763): 20170391. https://doi.org/10.1098/rstb.2017.0391
- Peters D, Schindler S (2023) FAIR for digital twins. CEAS Space Journal https://doi.org/10.1007/s12567-023-00506-y
- Pimm SL, Jenkins CN, Abell R, Brooks TM, Gittleman JL, Joppa LN, Raven PH, Roberts CM, Sexton JO (2014) The biodiversity of species and their rates of extinction, distribution, and protection. science 344 (6187): 1246752. https://doi.org/10.1126/science.1246752
- Robertson T, Wieczorek J, Raymond M (2022) Diversifying the GBIF Data Model. Biodiversity Information Science and Standards (6.94420). https://doi.org/10.3897/biss.6.94420
- Rozenberg S (2021) Digital records as relational objects–Yuk Hui's concept of digital objects applied to archival science. Archival Science 21 (2): 193-218. https://doi.org/10.1007/s10502-021-09357-0
- Sansone S-, McQuilton P, Rocca-Serra P, Gonzalez-Beltran A, Izzo M, Lister A, Thurston M (2019) FAIRsharing as a community approach to standards, repositories and policies. Nature Biotechnology 37 (4): 358-367. https://doi.org/10.1038/s41587-019-0080-8

- Schultes E, Roos M, Bonino da Silva Santos LO, Guizzardi G, Bouwman J, Hankemeier T, Baak A, Mons B (2022) FAIR Digital Twins for Data-Intensive Research. Frontiers in big data 5: 883341. https://doi.org/10.3389/fdata.2022.883341
- Scott B, Livermore L (2021) Extracting Data at Scale: Machine learning at the Natural History Museum. Biodiversity Information Science and Standards 5 https://doi.org/10.3897/biss.5.74031
- Thomer A, Weber N, Twidale M (2019) Supporting the long-term curation and migration of natural history museum collections databases. Proceedings of the Association for Information Science and Technology 55 (1): 504-513. https://doi.org/10.1002/pra2.2018.14505501055
- Webster MS, et al. (2017) The extended specimen: emerging frontiers in collections-based ornithological research. CRC Press
- Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, Robertson T, Vieglais D (2012) Darwin Core: an evolving community-developed biodiversity data standard. PloS one 7 (1): 29715. https://doi.org/10.1371/journal.pone.0029715
- Wittenburg P, Strawn G (2019) Commenting on "Digital Object" Aspects. https://doi.org/10.23728/b2share.2317b12321764f669c92ebbcf7518164
- Wittenburg P, Strawn G, Mons B, Boninho L, Schultes E (2019) Digital objects as drivers towards convergence in data infrastructures. Technical paper. doi 10: 2. https://doi.org/10.23728/b2share.b605d85809ca45679b110719b6c6cb11