

1200 UTC, 4 March 2025

Headline:

Meet the Libroscope: a new vision for 'liberating' data from biodiversity publications

Text:

Some of the world's leading institutions, experts and scientific infrastructures relating to biodiversity information are uniting around a new 10-year roadmap to 'liberate' data presently trapped in research publications.

The initiative aims to enable the creation of a 'Libroscope' - a mechanism for unlocking and linking data from scientific literature to support understanding of biodiversity, as the microscope and telescope previously revolutionized science. The plan largely builds on existing technology and workflows, and does not rely on construction of a new technical infrastructure.

The proposals result from a symposium involving 51 experts from 10 countries held in August 2024 at the 7th-century monastery at Disentis in the Swiss Alps, supported financially by the <u>Arcadia Fund</u>. The symposium was a 10-year follow-up to the 2014 meeting at Meise Botanic Garden in Belgium, which led to the <u>Bouchout Declaration on open</u> <u>biodiversity knowledge management</u>. The Disentis meeting evaluated progress since then, and identified priorities for the decade ahead.

While acknowledging major advances in the sharing and use of open biodiversity data, the participants noted that accessing data within research publications is often very cumbersome, with databases disconnected from each other and from the source literature. Liberating and linking data from such publications - estimated to encompass more than 500 million total pages - would represent a compelling mission for the next decade.

Achieving this mission will support further research, and understanding of biodiversity vital to meeting global goals and targets such as the <u>Kunming Montreal Global Biodiversity</u> <u>Framework</u> (KMGBF), as well as assisting the compiling of knowledge assessments such as those carried out by the <u>Intergovernmental Science-Policy Platform on Biodiversity and</u> <u>Ecosystem Services (IPBES)</u>.

A roadmap for staged action over the next decade was agreed by the symposium participants, with the following vision: "By 2035, the power of biodiversity knowledge from research publications will be fully leveraged within an open science framework, including



unencumbered data discovery, access, and re-use across scientific disciplines and policy applications."

The <u>Disentis Roadmap</u>, further developed following the symposium, and now released publicly, has already been signed by 26 institutions and a further 46 individual experts on five continents - among them major natural history collections such as Meise Botanic Garden, Botanic Garden and Botanical Museum Berlin, the National Museum of Natural History in Paris, and Royal Botanic Gardens, Kew; infrastructures such as the Global Biodiversity Information Facility (GBIF), Biodiversity Heritage Library (BHL), Catalogue of Life, LifeWatch ERIC and the Swiss Institute of Bioinformatics (SIB); journal publishers such as Pensoft Publishers and the European Journal of Taxonomy; research institutions such as Chinese Academy of Sciences and the Senckenberg Society for Nature Research; and networks such as TDWG Biodiversity Information Standards and Consortium of European Taxonomic Facilities (CETAF). See the <u>full list of signatories here</u>.

The roadmap remains open for further signatures, ahead of the launch of an action plan at the <u>Living Data</u> conference in Bogotá, Colombia in October 2025. The original signatories hope that a much broader group of institutions and individuals, across global regions and disciplines, will join the initiative and help to shape implementation of its vision. Engagement of funders will also be critical to realize its objectives.

The specific goals of the roadmap are that by 2035:

- All major public biodiversity research funders and academic publishers will encourage and enable publication of data adhering to the FAIR principles (findable, accessible, interoperable and reusable);
- Biodiversity-focussed publications will be accessible in machine-actionable formats, with all non-copyrightable parts of articles flowing into public data repositories;
- Published research on biodiversity will be 'fully AI-ready', that is openly available for AI training and properly labelled for ingestion by machine-learning modelled, within appropriate ethical and legal frameworks;
- Dedicated funding from research and infrastructure grants will be reserved for ensuring access to biodiversity data and knowledge.

Donat Agosti of the Swiss organization Plazi, who convened the Disentis symposium, commented: "We finally have a chance to make a quantum leap in understanding and monitoring biodiversity, by leveraging the power of digital technologies, and combining modern genomic methods with the vast amount of research data published daily and currently stuck in the publication prison. The 'Libroscope' will help to explore the universe of



existing knowledge, accumulated over hundreds of years, and bring it to the forefront of developments in the digital age, helping nature and people across the globe."

A recent demonstration of the principles of the 'Libroscope' was the launch of data portals for the European Journal of Taxonomy (EJT) and the Biodiversity Data Journal, as part of the GBIF hosted portal programme. The new portals showcase the data contained within taxonomic literature published by the journals, making use of the workflow originally developed by Plazi and partners to extract re-usable data from articles traditionally locked in static PDF files. Once created, these data objects then flow into platforms such as <u>GBIF</u>, <u>Catalogue of Life</u>, <u>ChecklistBank</u> and the <u>BiodiversityPMC</u>, and are stored in the <u>Biodiversity</u> <u>Literature Repository</u> at <u>Zenodo</u> hosted by <u>CERN</u>. This process enables data on new species and the location of related specimens cited in the literature to be openly accessible in near-real time, and available for long-term access.

"By repositioning scientific publications as an essential part of the research cycle, the Disentis Roadmap encourages publishers and the scientific community to move beyond open access towards FAIR access," said Laurence Bénichou, founder and liaison officer of the *European Journal of Taxonomy*. "Proactively ensuring data quality and dissemination is the core mission of the *European Journal of Taxonomy*. In this way, *EJT* enhances the immediate discoverability and usability of the taxonomic information it publishes, making it more valuable to the scientific community as a whole. Adherence to the Disentis vision marks a crucial step in the liberation and enrichment of knowledge about biodiversity."

The Chief Executive Officer of Meise Botanic Garden, Steven Dessein, who attended the Disentis Symposium, commented: "Meise Botanic Garden fully supports the Disentis Roadmap, which builds on the foundation laid by the Bouchout Declaration. Open biodiversity data is essential to tackling today's pressing environmental challenges, from biodiversity loss to climate change.

"By ensuring research publications become more accessible and interconnected, this roadmap represents a critical step toward harnessing biodiversity knowledge for science, policy, and conservation."

Christophe Déssimoz, Executive Director of the SIB Swiss Institute of Bioinformatics, another signatory of the Disentis Roadmap, added: "We have long championed the principles of open, structured, and interoperable data to advance life sciences. The Disentis Roadmap applies these same principles to biodiversity knowledge, ensuring that critical data is not just available, but truly actionable for research, policy, and conservation."



The director of the Botanic Garden and Botanical Museum of Berlin, Thomas Borsch, noted that more than any other branch of science, taxonomic research depended on the machineactionable availability of biodiversity data from the literature. "The 'Libroscope' postulated in the Disentis Roadmap will enable a new generation of research workflows through its interoperable approach," said Professor Borsch. "This will be very helpful to address pressing issues in biodiversity research and in particular to improve the use of quality information on organisms in national and global assessments."

The chief scientist of the national museum of natural history in Paris (MNHN) said: "We, like all similar museums and taxonomic institutions, are focussed on linking taxonomic and collection data with digital reproductions and molecular information to create the 'extended digital specimen.' However, the potential of taxonomic publications and text mining should not be underestimated either. On the contrary, it is a smart and accessible way to dig into scientific publications so as to retrieve, link and consolidate, research data of great relevance to many disciplines. This is why our institution fully supports the Disentis initiative."

Christos Arvanitidis, CEO of the Biodiversity and Ecosystem e-Science Infrastructure LifeWatch ERIC, commented: "LifeWatch ERIC is proud to be part of this initiative, as providing access and support to biodiversity and ecosystem data is fully aligned with our mission. The Disentis Roadmap opens up new opportunities for our research infrastructure to help make what science has provided us accessible and usable, and to improve the FAIRness of data for research and science-based policy."

"As a publisher of dozens of renowned academic journals in the field of biodiversity and systematics with experience in technology development, at Pensoft, we have always recognised the key role of academic publishers in scholarly communication. It's not only about publishing the latest research. Above all, it's about putting scientific work in the hands of those who need it: be it future researchers, policy-makers or their Al-powered assistants. Now that the Disentis roadmap is already a fact, we hope that many others will also join us on this ambitious journey to open up the knowledge we have today for those who will need it tomorrow," said Prof. Dr. Lyubomir Penev, founder and CEO at Pensoft, who also attended the Disentis symposium.

"We're excited to see the results from Disentis partners like Plazi, BHL, Pensoft and the European Journal of Taxonomy who are focussed on liberating data connected with scientific publications," said Tim Robertson, deputy director and head of informatics at the Global Biodiversity Information Facility (GBIF), who attended the Disentis meeting. "GBIF will continue to do our part to improve the standards, tools and services that help expand both the benefits and the impact of FAIR and open data on biodiversity science and policy."



Olaf Bánki, Executive Director of the Catalogue of Life, commented: "We call out to the scientific community, especially the younger generation, to join our effort in unlocking biodiversity data from literature. Actionable biodiversity and taxonomic data from digitized literature contributes to creating an index of all described organisms of all life on earth. We need such data to tackle and understand the current biodiversity crisis."

For further information, contact: Donat Agosti <u>agosti@plazi.org</u> Tim Hirsch <u>tim@timhirsch-consulting.com</u>

Social media:

@libroscope.bsky.social @libroscope@ecoevo.social https://www.linkedin.com/company/disentis-roadmap/