A Yarnball of Frayed Strings - Communication, Impact & Cohesion of European Biodiversity Research Infrastructures

Erik Kusch[‡], Allan T Souza[§], Yi-Ming Gan[|], Carrie Andrew[¶]

‡ University of Oslo, Oslo, Norway
§ Institute for Atmospheric and Earth System Research INAR, Forest Sciences, Faculty of Agriculture and Forestry, University of Helsinki, Helsinki, Finland
| Royal Belgian Institute of Natural Sciences, Brussels, Belgium
¶ Natural History Museum, University of Oslo, Oslo, Norway

Corresponding author: Erik Kusch (erik.kusch@nhm.uio.no)

Abstract

European biodiversity research infrastructures (BioRIs) play a critical role in addressing the escalating biodiversity crisis by providing data, tools, and services necessary for scientific research and policy-making. However, despite their potential, these infrastructures are often fragmented in terms of communication and coordination, scientific fields most impacted as well as the interoperability and cohesion of their services. This fragmentation impedes their ability to collaborate effectively and facilitate comprehensive solutions to biodiversity challenges. Motivated by the need to enhance the coherence and impact of European BioRIs, this study investigates four major infrastructures: Distributed System of Scientific Collections (DiSSCo), Integrated European Long-Term Ecosystem Research (eLTER), Global Biodiversity Information Facility (GBIF), and LifeWatch ERIC. Through surveys of managerial staff and literature tracking, we assess the communication and collaboration efforts within and across these infrastructures, evaluate their scientific impact, and explore the cohesion of their services and data. Our results show that while internal communication is stronger within individual BioRIs, cross-infrastructure collaboration is limited. Notably, eLTER and LifeWatch exhibit higher levels of internal interaction compared to DiSSCo and GBIF. Our study also highlights significant challenges in public engagement and data standardization. We conclude with recommendations for improving interoperability and communication to strengthen the role of BioRIs in addressing biodiversity issues at both European and global scales.

Keywords

DiSSCo, eLTER, GBIF, LifeWatch, Biodiversity, BioDT, Digital Twin

Introduction

Research Infrastructures (RIs) are facilities that offer resources and services for research communities to conduct research and promote innovation. An RI can be represented by a single site or function across sites in a distributed fashion including major scientific equipment, collections, archives or scientific data, computing systems and communication networks (Fabre et al. 2021, Hallonsten 2020).

RIs provide important services to the scientific community, for instance, producing standardized methods, tools, data and platforms that facilitate the development of scientific research (Fabre et al. 2021). More importantly they can facilitate access to standardized data across different regions, countries and continents, fostering the development of data intensive research (Farshidi et al. 2023). Such infrastructures are particularly relevant to the study of biodiversity patterns and processes, as there are still many gaps in knowledge and geographic coverage in this field. In fact, our planet is currently facing an unprecedented crisis in biodiversity (Johnson et al. 2017). This crisis has several angles (e.g. climatic, pollution, deforestation, overexploitation) and affects different ecosystems and taxonomic groups in a multitude of ways (Dobson et al. 2021). RIs can support the scientific community to produce more complex and data richer outputs tackling these challenges (Farshidi et al. 2023).

Due to the important role that RIs are already playing in the advance of scientific knowledge, it is crucial that these infrastructures are well-connected and aware of the developments occurring in their respective networks and their partners networks. Well connected and informed RIs provide a more coherent platform for data sharing, reduce redundancy, and enhance the ability to build on existing knowledge (Farshidi et al. 2023). Towards this goal, it is important to not only focus on the technical interoperability but also robust communication among the people who operate and use the RIs. Effective collaboration, free and easy access to establish dialogue among RI staff is essential for aligning goals, sharing best practices, and ensuring that diverse RIs can work together efficiently and facilitate novel and relevant research impacts (Hodapp and Hanelt 2022).

An example of the benefits of such collaboration is the development of the Biodiversity Digital Twin (BioDT) project (Fig. 1). This project is facilitated through the services and resources offered through four distinct RIs that play pivotal roles in biodiversity research: Distributed System of Scientific Collections (DiSSCo), Integrated European Long-Term Ecosystem, critical zone and socio-ecological Research (eLTER), Global Biodiversity Information Facility (GBIF), and LifeWatch ERIC. Each of these distributed RIs (see Fig. 2 for an overview of their respective national representations) represents an international intiative varying in maturity, scope, and areas of expertise, but together they form a comprehensive network that enhances our ability to understand and address biodiversity challenges (Table 1).

Here, we assess the capabilities of these four European biodiversity research infrastructures (BioRIs) in facilitating state-of-the-art scientific advancements. To this end, we investigate (1) communication & coordination efforts within and across the four BioRIs, (2) the impact on the scientific frontiers of biodiversity-related research, and (3) the cohesion and interopeability of their services and resources. Subsequently, we identify strenghts and shortcomings among the four BioRIs operating in Europe. Finally, we provide a set of recommendations for how continued usefulness of these infrastructures can be maintained and how coordination of their efforts may be aligned better in the future thus facilitating more cutting-edge research at European as well as Global scales.

Communication & Coordination Efforts of European BioRIs

Effective communication and coordination of RIs, particularly of distributed systems, has been identified as a major key performance indicator by funding bodies, research infrastructure managers, as well as end-users (Reid 2012, Kolar et al. 2019, Raes et al. 2020). However, communication and coordination are complex processes involving (1) diverse sets of stakeholders, (2) means of sharing information, as well as (3) distinct modes of directionality of flow of information. Analysing the state of the communication & coordination capabilities of European BioRIs thus necessitates investigation of all these three aspects simultaneously. Doing so enables subsequent recommendations of how to accomplish the key recommendation for clear channels of RI communications made by the European Commission (Commission 2010).

Stakeholders & Communication Participants

In assessing stakeholders and participants in a communication and coordination network comprised of several research infrastructures, there exist three distinct levels of communication and coordination efforts: (1) within a singular research infrastructure (i.e., information-flow between entities belonging to the same distributed research infrastructure), (2) between and across multiple research infrastructures (e.g., information-flow between an entity of GBIF and an entity of eLTER), and finally (3) between research infrastructures and the public (e.g., information-flow from within a research infrastructure to the public in response to an inquiry made by a member of the public). Due to the heterogenous nature of maturity level, target audiences, and services rendered by our four target RIs (eLTER, DiSSCo, GBIF, LifeWatch), generating an objective, quantitative, and standardised assessment of communication/coordination efficiency across these levels and audiences is nigh impossible to generate and thus does not exist.

To overcome this limitation of our current understanding of the communication & coordination capabilities of European BioRIs, we carried out a survey of research infrastructure managerial staff. Within this survey (see Suppl. material 1), we requested that managerial staff of national representations (i.e., nodes/sites) of each target BioRI self-report their communication and coordination with other nodes belonging to the four target BioRIs. In designing this survey, we made sure to follow long-established good-practices of survey design (Schaeffer and Dykema 2020) and solicitated rankings of collaboration intensity across a likert scale (ranging from 1 to 5) whose values correspond to:

1. Shared group calls or mailing lists
2. Sporadical personal/one-on-one interaction
3. Collaboration on shared projects
4. Collaborate on shared tasks
5. Shared offices and or staff

By providing these set interpretations of the likert scale intervals, we attempted to diminish the effect of subjectivity in responses given by BioRI managerial staff. Lastly, we also allowed a selection of "no interaction/collaboration" to denote instances of the European biodiversity BioRI network being entirely unconnected and without communiation pathways or coordination efforts. Leveraging the BioRI-self-reporting of communication/coordination intensity, we indexed communication and coordination intensity (1) within singular BioRIs and (2) across multiple BioRIs. Finally, the response rate to the survey itself is used as a proxy for (3) communication intensity with members of the public.

The responses to our survey identify a clear and statistically significant trend of communication and coordination being most pronounced within a singular BioRI and much weaker across BioRIs (Fig. 3). While this is entirely expected - for several reasons including alignment of topics, shared vocabularies, as well as shared means of communications within singular BioRIs - it is nevertheless concerning that most national BioRI representations hardly interact at all with individual national representations of a different BioRI or worse: know of their existence. While the four target BioRIs perform similarly poorly at communicating with entities of other BioRIs, there do exist pronounced differences of collaboration intensity within singular BioRIs. Particularly, eLTER and LifeWatch RIs collaborate much more intensely within their RIs than DiSSCo and GBIF do within theirs.

Assessing the communicativity of the European BioRIs with representatives of the public, we identify a regrettable and alarmingly low rate of responsiveness - both in rate-of-response (i.e., how many of the queried managerial staff responded to the survey issued) as well as time-to-response (i.e., time elapsed between issuance of inquiry and response rendered from the queried managerial staff). The response rates over time to our information-gatherin survey can be found in Fig. 4. Notice how only two the queried four BioRIs achieved response rates of more than 50% with none of the BioRIs surpassing this responsiveness threshold without the involvement of their board of directors or in-person meetings as was the case for GBIF whose lionshare of responses were solicitated during an in-person meeting of European and Central Asia nodes (GBIF ECA). These low rates of response and long times of response we find are exacerbated by the fact that our survey was made to fulfill a deliverable for the BioDT project - an EU horizon project which all four BioRIs are involved in - and thus one could argue that our information-gathering inquiry represents an inquiry made from a not-enitrely disconnected member of the public.

While a multitude of competing and complementary key performance indicators have been proposed to evaluate communication and collaboration performance of BioRIs (Kolar et al. 2019), the results of our survey at present already indicate a pronounced need for action to be taken. Our analyis presented here does not represent the entire communication/coordiation gammut, but nevertheless identifies particularly relevant communication levels which BioRIs perform differently across. This allows for the investigation of reasons behind differing levels of communicativity within BioRIs and thus serve to establish recommendations and guidelines for remedying shortcomings in one BioRI by learning from the processes enacted in another. Lastly, we have also identified commonalities of poor communication/coordination across differing BioRIs and with the public which need to be addressed at a cross-BioRI level.

Means of Communication

Sharing of information can be carried out via a multitude of means - each presenting distinct benefits and downsides for the purpose of communication with a specific stakeholder group. Throughout our process of querying BioRI managerial staff for self-reports of realised communication and coordination potential, we also engaged in exhaustive conversations and explorations of which means of communication the different BioRIs utiliise. What we find is that there exists no alignment across several or even just within single BioRIs as to which single means of communication is used for sharing of information. While spreading out information-flow across several means of communication itself is not necessarily detrimental to communication efforts, doing so without clear understanding of what information to communicate with which means is.

This absence of clear knowledge of which means of communication to engage with to obtain relevant information from BioRIs is exacerbated further by drawbacks of the most frequently used means of communication across our four target BioRIs. Firstly, mailing lists, while used almost ubiquitously, are private by design and nigh impossible to penetrate or stay up-to-date with particularly as the number of communication participants increases. Secondly, discussion boards make for much easier gaining of an overview of information pertaining to topics of interest but may not be desired to be made public by BioRI coordinators for data protection reasons and become confusing for the purpose of across-BioRI information sharing when different layouts and terminology are employed. Thirdly, social media platforms arguably combine the worst aspects of both mailing lists and discussion boards in that social media platforms usually make follow-up on past information impossible unless it was shared publicly in the first place and may preclude full access to relevant information through their compartmentalisation of communication into granular chat and group environments. Lastly, some national representations of research infrastructures rely heavily on information sharing via in-person communication such as workshops and conferences as well as what can best be described as "oral tradition" - a means of communication not befitting of sharing information outside a small group of people.

Ultimately, we find that, to realise more of the untapped communication and coordination potential within and across European BioRIs, a uniqfying means of communication is required which streamlines and simplifies information sharing through (1) persistence and traceability of information, (2) standardised layouts and terminology and (3) classification of information as either public or private with designated stakeholder audiences.

Directionality of Communications

There are two main types of communication engagement between research projects and stakeholders: One-way communication which is disconnected exchange of information in time and space such as publications, databases, newsletters, websites etc. and two-way communication which involves interaction between different parties connected in time and space, such as conferences, workshops, meetings etc regardless in-person or virtual (Jolibert and Wesselink 2012). While the former directionality is comparatively easier to achieve and likely sufficient for most interactions and use-cases of the public with regard to BioRIs, two-way communication is crucial to improving coordination of European biodiversity BioRIs.

Scientific Impact of European BioRIs

Research infrastructures have been designed to support research projects that - in the case of our target BioRIs - identify, quantify, and forecast biodiversity metrics, patterns, and processes across space and time. This is a non-trivial mission goal not just due to the magnitude of its impact once realised but also due to the variability of research topics that require supporting and the quantity of data and services required to do so. Likewise then, assessing or quantifying the impact BioRIs have on scientific advancements is a nuanced topic that justifies separate investigation such as done by Fabre et al. 2021 who propose a number of key metrics for quanitfication of scientific impact of RIs.

Due to the vastly different lifecycle stages of the BioRIs we are investigating (see Table 1), we are unable to compute a holistic set of scientific impact indicators for this study. Instead, we rely on BioRI-internal procedures of literature tracking to identify volume of publications in scientific journals using our target BioRIs. In addition, we also identify percentages of these publications which are peer-reviewd and open-access as these arguably represent the gold standard of research outcomes which European BioRIs ought to facilitate. Unfortunately, only two of the four BioRIs under investigation here track citations and so our explorative analysis is limited to eLTER and GBIF literature tracking (Fig. 5).

While literature using GBIF data and services far outnumbers those relevant to eLTER, we would like to point out that this may not necessarily be indicative of wider applicability of GBIF-mediated data or services but rather a product of GBIF's considerably longer lifespan, stronger global representation, and advanced data portals. Both GBIF and eLTER facilitate fair rates of peer-reviewed research, particularly in the case of eLTER. However, lower rates of the publications facilitated by these BioRIs are openly accessible (Fig. 5).

Quantification of scientific impact purely by numbers of publications - even when taking into account rates of open-access and peer-review - may obfuscate further relevant nuance with respect to how well BioRI data and services align with relevant fields of research. To this end, GBIF has already implemented a literature classification scheme with which the impact of GBIF infrastructure of scientific advancement in distinct fields can be quantified. Investigating these classifications, we find that peer-reviewed, open-access scientific journal-based publications faciliated by GBIF target mostly research topics such as Ecology, Conservation, Climate Change and Species Distributions. However, we also find wider applicability of GBIF data and services (Fig. 6). We expect this wide applicability of data and services to be mirrored, if not greater (owing to its application beyond biodiversity data purveying). for eLTER if such a classification system were implemented.

Cohesion & Interoperability of Services and Resources

Biodiversity data standards originated as a computational offshoot of the data digitization process. Without data standards, there are no uniform understandings of, or notations for, the compilation and integration of biodiversity data. In addition, machine readability of data, which is important for research infrastructures as well as for AI, depends upon data standards for interoperability. Thus, there are many forms of data standards that are critical for biodiversity data and BioRIs. Andrew et al. 2024 provide a more comprehensive guide to understanding biodiversity data standards. For our purposes, the important point is that data standards bring information into a standardized format for storage, and that they support computational communication.

Biodiversity data standards were earliest applied to museum data (GRAHAM et al. 2004), and soon after integrated into research data infrastructures like GBIF (Wieczorek et al. 2012, Robertson et al. 2014, Pooter D et al. 2017). With more forms of data represented in a greater number of infrastructures, the practicalities of data standardization are changing (Schneider et al. 2019). One impediment is that standards have typically been created on an as-needed basis. Hence, the protocols and methods for their development vary, and the terms they include are never comprehensive of all forms of biodiversity data (Andrew et al. 2024).

The four BioRIs that we focus on all contain different forms of biodiversity data, and this reduces the amount of data standards that they share (Andrew et al. 2024). This, in turn, limits the potential for data interoperability between them. Thus, in addition to our survey results demonstrating low communication between personnel of the different RIs (Fig. 3), the likelihood for data interoperability, through shared data standards, is also likely limited. GBIF and DiSSCo share the most data standards, because both contain museum data. Each apply the Access to Biological Collection Data (ABCD) and Darwin Core (DwC) data standards (Holetschek et al. 2012, Wieczorek et al. 2012, Walls et al. 2014). They also implement other data standards that are unique to each: Collection Descriptions (CD) and the Minimum Information about a Digital Specimen (MIDS) for DiSSCo, and the Humboldt Core Extension (https://eco.tdwg.org; Guralnick et al. 2017) for GBIF. In contrast, eLTER data come from site-level field data (Wohner et al. 2022), and are more heterogeneous. They cover datasets in the atmospheric, ecological, geological, hydrological & social sciences; it would be exceptionally challenging to standardize data across all of those topics. Instead, eLTER has created standards observations that emphasize the standardization of measurements and protocols, rather than of variables between datasets. eLTER does not, then, overlap with any other RI in their data standards. They do support, but not require, the use of existing established vocabularies, such as the Ecological Thesaurus (EnvThes; Schentz et al. 2013). The final BioRI under investigation here, LifeWatch, lacks any specific data standards because they do not directly aggregate data. Instead, LifeWatch mostly provides support for data logistics. LifeWatch does provide resources for applying data standards, such as EcoPortal (https://ecoportal.lifewatch.eu), which contains around 30 data standards and related domains, including the eLTER-supported EnvThes. Despite the differences in data, they all ultimately support biodiversity data. Hence, greater interoperability between datasets should conceivably increase communication between the RIs, in addition to providing greater likelihood of supporting the use of their combined data in research projects such as the Biodiversity Digital Twin (Westerlaken 2024, Trantas et al. 2023).

Discussion

Our investigation of European Biodiversity Research Infrastructures (BioRIs) highlights considerable fragmentation of within and across four distinct European BioRIs (i.e., eLTER, DiSSCo, GBIF, LifeWatch). Such fragmentation presents significant challenges that hinder efficient collaboration, resource sharing, and ultimately impede scientific progress through misallocation of efforts and funds. These challenges are particularly evident in several key areas: geographical representation, communication and coordination, scientific areas impacted, and data and service cohesion.

Despite one or more BioRIs being present in almost every European country (Fig. 2A), currently, BioRIs are unevenly distributed individually across Europe (Fig. 2B). This geographical fragmentation of individual BioRIs leads to unequal opportunities for researchers in different regions. To ensure that research infrastructures are accessible to all and avoid biases in research outcomes, there is a need for a more homogeneous representation of BioRIs across Europe. This would ensure that regions currently underrepresented are provided with the same opportunities for scientific collaboration and resource access as more established regions. A concerted effort to expand and balance BioRI representation could contribute to a more equitable research environment. This, however, is a costly undertaking which is unlikely to be realised without a major influx of monetary resources. Ultimately and in lieu of a sudden change to ongoing international efforts to support and grow BioRIs, overcoming geopgraphical fragmentation will likely be remedied better through increased communication and capacity sharing efforts across existing BioRI networks.

Such a need for increased communcation and capacity sharing efforts is already supported by our findings. We find that BioRIs across Europe are fragmented with regards to ongoing communication and coordination across each other and perform differingly at communicating and sharing efforts within themselves. Differing means of communication most of which are geared toward one-way flow of information are likely at the core of this insular nature of BioRI communication networks. To overcome this limitation, which we find to be central to many issues faced by BioRIs as well as their user-bases, we recommend the creation and adoption of a standardized communication framework. By implementing such a system, BioRIs can strengthen internal collaboration, promote new partnerships, and make information more accessible to the broader scientific community and the public. This would also help mitigate information silos and enhance stakeholder engagement.

Some efforts are already underway to address these communication challenges. For example, initiatives like the EU-funded GLOBIS-B project have demonstrated the potential for successful collaboration across geographically and thematically diverse RIs. GLOBIS-B has focused on mobilizing data and generating knowledge across multiple domains, organizing workshops that unite scientific communities and service providers under common objectives (Kissling et al. 2015). Another example is the ENVRIplus project, which has worked to harmonize observation methodologies and access policies across multiple European Environmental and Earth System RIs. This effort has fostered collaboration through staff exchange programs and the generation of common solutions for shared infrastructures (ENVRIplus 2024).

These examples highlight the importance of robust communication strategies. Previous research has shown that strong, user-friendly communication platforms can significantly enhance collaboration among RIs (D’Ippolito and Rüling 2019, Balli et al. 2022). Developing such tools, along with creating spaces for regular inter-RI meetings and collaborative online environments, would foster a sense of community and shared purpose. This can also lead to more efficient problem-solving and innovation, as researchers and infrastructure providers can exchange ideas and experiences more freely. Open dialogue can help RIs align their development efforts and streamline their approaches toward achieving common goals. Achieving the ambitious demands of cutting-edge research to adress biodiversity issues of the 21st century will depend on the development of seamless communication policies and tools that enable different infrastructures to integrate their data and services efficiently (Westerlaken 2024, Trantas et al. 2023).

While the diversity of research areas impacted by different BioRIs is a valuable asset poised to respond the research demands of the 21st century, it underscores the importance of data cohesion and compatibility. One ongoing concern is to what extent the lack of communication between BioRIs is due to systemic differences in their objectives and the types of data they handle. A better understanding of these underlying differences is needed to assess whether they contribute to fragmentation or if they could be harmonized to enhance interoperability. We suggest that this deeper understanding can only be created itteratively through efforts like literature tracking services already implemented in some of the BioRIs we have investigated. Subsequently, addressing these issues could pave the way for more comprehensive and coherent data integration, benefitting the entire biodiversity research landscape.

In conclusion, tackling fragmentation within European BioRIs requires a multi-faceted approach. Ensuring broader geographical representation, fostering collaboration through standardized communication platforms, and promoting data cohesion will be key steps toward a more integrated and efficient research infrastructure network. These efforts will ultimately support more innovative and impactful biodiversity research, ensuring that Europe remains at the forefront of global efforts to understand and protect biodiversity.

Data Availability

All data and code supporting this manuscript is available at https://github.com/ErikKusch/BioDT-Research-Infrastructure.

Acknowledgements

This study has received funding from the European Union's Horizon Europe Research and
Innovation Programme under grant agreement No 101057437 (BioDT project, https://doi.org/10.3030/101057437). Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them. This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 871128 (eLTER PLUS).

Funding program

European Union Horizon Europe Research and Innovation Programme

Grant title

Biodiversity Digital Twin for Advanced Modelling, Simulation and Prediction Capabilities

Ethics and security

This study relies on responses from RI managerial staff to a survey. Responses to this survey have been anonymised and only those responses declaring consent to be included in a publication following anonymisation provedures have been included in this study.

Author contributions

Erik Kusch lead this study and coordinated writing efforts. Allan Souza and Erik Kusch conceptualised this manuscript. Erik Kusch carried out analyses and produced figures. Erik Kusch conceptualised and subsequently carried out the RI manager survey with feedback from Allan Souza and Carrie Andrew. Allan Souza and Yi-Ming Gan provided literature tracking data. All other contributed to writing and gave final approval for publication.

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Andrew C, Islam S, Weiland C, Endresen D (2024) Biodiversity data standards for the organization and dissemination of complex research projects and digital twins: a guide. arXiv preprint. URL: https://arxiv.org/abs/2405.19857
- Balli E, Angioni B, Cerrato S, Filosa S, Goffi C, Postpichl L, Silva R, Stanculescu D, Cocco M (2022) Communication strategy and plans for research infrastructures: the EPOS case. Annals of Geophysics 65 (3). https://doi.org/10.4401/ag-8845
- Commission E (2010) Cost control and management issues of global research infrastructures.
- D'Ippolito B, Rüling C (2019) Research collaboration in Large Scale Research Infrastructures: Collaboration types and policy implications. Research Policy 48 (5): 1282-1296. https://doi.org/10.1016/j.respol.2019.01.011
- Dobson A, Rowe Z, Berger J, Wholey P, Caro T (2021) Biodiversity loss due to more than climate change. Science 374 (6568): 699-700. https://doi.org/10.1126/science.abm6216
- ENVRIplus (2024) ENVRIplus. https://www.envriplus.eu/. Accessed on: 2024-8-27.
- Fabre R, Egret D, Schöpfel J, Azeroual O (2021) Evaluating the scientific impact of research infrastructures: Therole of current research information systems. Quantitative Science Studies 2 (1): 42-64. https://doi.org/10.1162/qss_a_00111
- Farshidi S, Liao X, Li N, Goldfarb D, Magagna B, Stocker M, Jeffery K, Thijsse P, Pichot C, Petzold A, Zhao Z (2023) Knowledge sharing and discovery across heterogeneous

research infrastructures. Open Research Europe 1 https://doi.org/10.12688/openreseurope.13677.3

- GRAHAM C, FERRIER S, HUETTMAN F, MORITZ C, PETERSON A (2004) New developments in museum-based informatics and applications in biodiversity analysis. Trends in Ecology & Evolution 19 (9): 497-503. https://doi.org/10.1016/j.tree.2004.07.006
- Guralnick R, Walls R, Jetz W (2017) Humboldt Core – toward a standardized capture of biological inventories for biodiversity monitoring, modeling and assessment. Ecography 41 (5): 713-725. https://doi.org/10.1111/ecog.02942
- Hallonsten O (2020) Research Infrastructures in Europe: The Hype and the Field. European Review 28 (4): 617-635. https://doi.org/10.1017/s1062798720000095
- Hodapp D, Hanelt A (2022) Interoperability in the era of digital innovation: An information systems research agenda. Journal of Information Technology 37 (4): 407-427. https://doi.org/10.1177/02683962211064304
- Holetschek J, Dröge G, Güntsch A, Berendsohn WG (2012) The ABCD of primary biodiversity data access. Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology 146 (4): 771-779. https://doi.org/10.1080/11263504.2012.740085
- Johnson C, Balmford A, Brook B, Buettel J, Galetti M, Guangchun L, Wilmshurst J (2017) Biodiversity losses and conservation responses in the Anthropocene. Science 356 (6335): 270-275. https://doi.org/10.1126/science.aam9317
- Jolibert C, Wesselink A (2012) Research impacts and impact on research in biodiversity conservation: The influence of stakeholder engagement. Environmental Science & Policy 22: 100-111. https://doi.org/10.1016/j.envsci.2012.06.012
- Kissling WD, Hardisty A, García EA, Santamaria M, De Leo F, Pesole G, Freyhof J, Manset D, Wissel S, Konijn J, Los W (2015) Towards global interoperability for supporting biodiversity research on essential biodiversity variables (EBVs). Biodiversity 16: 99-107. https://doi.org/10.1080/14888386.2015.1068709
- Kolar J, Cugmas M, Ferligoj A (2019) Towards Key Performance Indicators of Research Infrastructures. arXiv https://doi.org/10.48550/arxiv.1910.00304
- Pooter D, W A, N B, S B, K D, M E, E F, A G, P G, M L, M L (2017) Toward a new data standard for combined marine biological and environmental datasets-expanding OBIS beyond species occurrences. Biodiversity data journal 5 URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5345125/
- Raes N, Casino A, Goodson H, Islam S, Koureas D, Schiller E, Schulman L, Tilley L, Robertson T (2020) White paper on the alignment and interoperability between the Distributed System of Scientific Collections (DiSSCo) and EU infrastructures - The case of the European Environment Agency (EEA). Research Ideas and Outcomes 6 https://doi.org/10.3897/rio.6.e62361
- Reid A (2012) Best Practice In the Management of research infrastructures. https://doi.org/10.13140/RG.2.2.15046.47685
- Robertson T, Döring M, Guralnick R, Bloom D, Wieczorek J, Braak K, Otegui J, Russell L, Desmet P (2014) The GBIF Integrated Publishing Toolkit: Facilitating the Efficient Publishing of Biodiversity Data on the Internet. PLoS ONE 9 (8). https://doi.org/10.1371/journal.pone.0102623
- Romano P, Giugno R, Pulvirenti A (2011) Tools and collaborative environments for bioinformatics research. Briefings in Bioinformatics 12 (6): 549-561. https://doi.org/10.1093/bib/bbr055

- Schaeffer NC, Dykema J (2020) Advances in the Science of Asking Questions. Annual Review of Sociology 46 (1): 37-60. https://doi.org/10.1146/annurev-soc-121919-054544
- Schentz H, Peterseil J, Bertrand N (2013) EnvThes–interlinked thesaurus for long term ecological research, monitoring, and experiments. *Proceedings of EnviroInfo*.
- Schneider F, Fichtmueller D, Gossner M, Güntsch A, Jochum M, König-Ries B, Le Provost G, Manning P, Ostrowski A, Penone C, Simons N (2019) Towards an ecological trait-data standard. Methods in Ecology and Evolution 10 (12): 2006-2019. https://doi.org/10.1111/2041-210x.13288
- Trantas A, Plug R, Pileggi P, Lazovik E (2023) Digital twin challenges in biodiversity modelling. Ecological Informatics 78 https://doi.org/10.1016/j.ecoinf.2023.102357
- Walls R, Deck J, Guralnick R, Baskauf S, Beaman R, Blum S, Bowers S, Buttigieg PL, Davies N, Endresen D, Gandolfo MA, Hanner R, Janning A, Krishtalka L, Matsunaga A, Midford P, Morrison N, Tuama É, Schildhauer M, Smith B, Stucky B, Thomer A, Wieczorek J, Whitacre J, Wooley J (2014) Semantics in Support of Biodiversity Knowledge Discovery: An Introduction to the Biological Collections Ontology and Related Ontologies. PLoS ONE 9 (3). https://doi.org/10.1371/journal.pone.0089606
- Westerlaken M (2024) Digital twins and the digital logics of biodiversity. Social Studies of Science https://doi.org/10.1177/03063127241236809
- 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). https://doi.org/10.1371/journal.pone.0029715
- Wohner C, Peterseil J, Klug H (2022) Designing and implementing a data model for describing environmental monitoring and research sites. Ecological Informatics 70 https://doi.org/10.1016/j.ecoinf.2022.101708

Figure 1.

Conceptual representation of the Biodiversity Digital Twin (BioDT) with the integration of the four research infrastructures (DiSSCo, eLTER, GBIF and LifeWatch) involved in the development of the digital twin

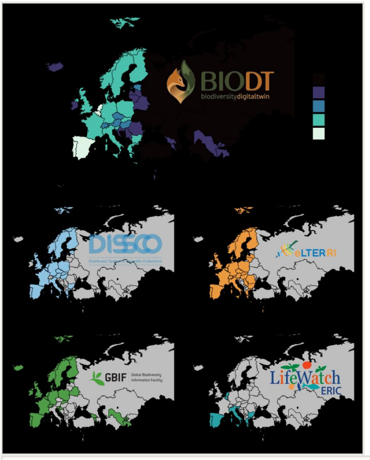

Figure 2.

Distribution & Representation of European Biodiversity Research Infrastructures. The network of european biodiversity research infrastructures is fragmented (A) with each individual research infrastructure differing in its geographical coverage (B).

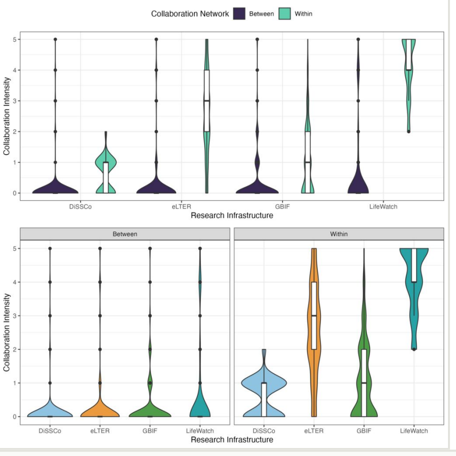

Figure 3.

Self-reported collaboration intensity within and across BioRIs. Upper panel: Collaboration profiles of all of the four target BioRIs show a clear bias for within-BioRI collaboration with little to no collaboration across BioRIs being reported. Lower panel: While BioRIs perform similarly poorly at collaborating with other BioRIs (lower-lefthand panel), some BioRIs collaborate much more intensely within themselves than do others (lower-righthand panel).

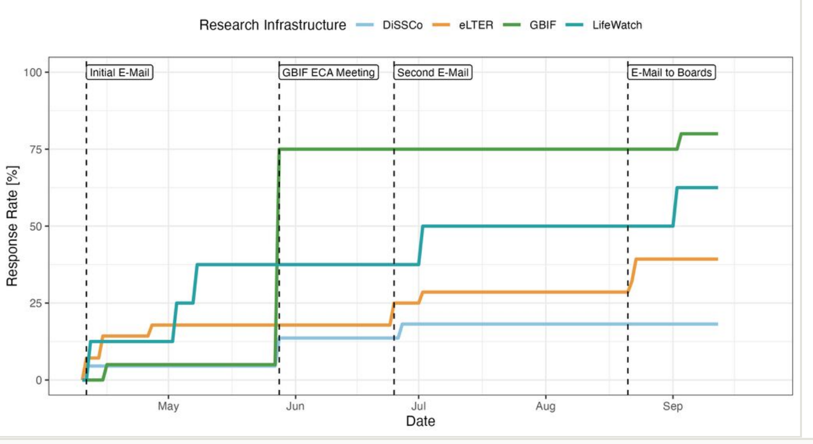

Figure 4.

Survey responses of queried BioRIs. Plotting rate of response (% of queried national representations of each BioRI who responded to the survey inquiry) over time since first issuance of the inquiry reveals poor responsiveness across all target BioRIs with some BioRIs performing particularly poorly at communicating with the public. In addition, despite all BioRIs represented here being distributed systems, the value of in-person meetings (e.g. GBIF ECA) in solicitating information from within BioRIs remains invaluable.

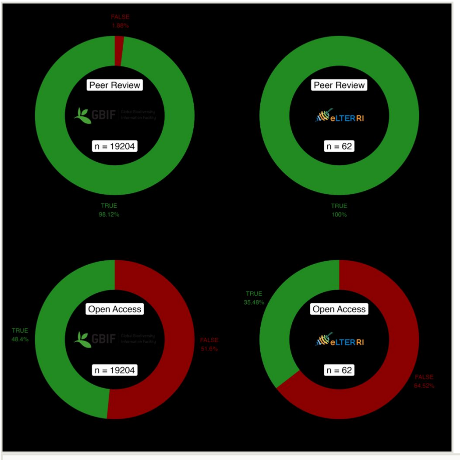

Figure 5.

BioRI-Facilitated Literature Tracking. BioRI-driven literature tracking reveals that GBIF data and services have facilitated a vast number of publications whereas the fledging eLTER RI data and services have less so. Nevertheless, eLTER-facilitated publications are overwhelmingly peer-reviewed as compared to GBIF-facilitated counterparts. Publications facilitated by either, however, are openly accessible at lower rates. Only eLTER and GBIF currently track publications their efforts have facilitated.

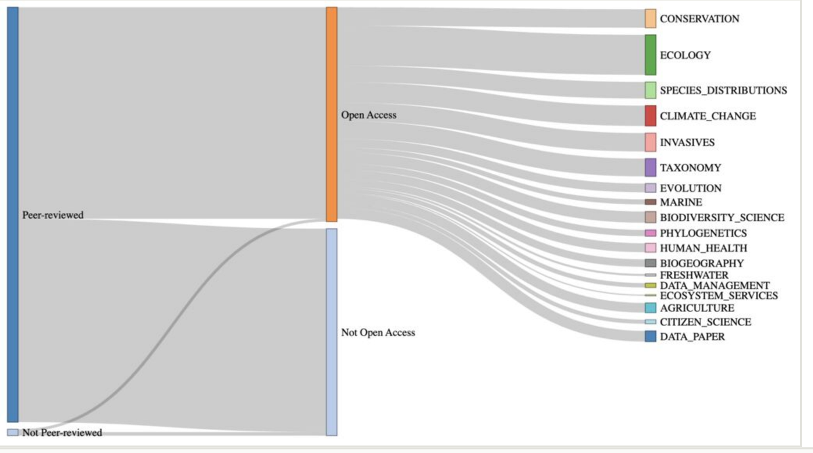

Figure 6.

GBIF-Facilitated Literature Topics. Classification of publications facilitated by an RI (i.e. GBIF) can be used to determine which topics an RI's chief scientific impact falls on. In the case of GBIF, the topics contained in publications which are facilitated by GBIF, peer-reviewd and oppen access (the gold standard of research outcomes), scientific topics are broad and far ranging. Only GBIF currently tracks topics of publications its efforts have facilitated.

Table 1.

Key characteristics of the biodiversity research infrastructures under investigation.

RI	Started year	Current number of member countries	Coverage	Main aim	Spheres of knowdledge
DiSSCo	2018	23	European	To digitize and mobilize data from natural history museums	Biosphere
eLTER	2020	26	European	To facilitate long-term ecological observations and research across various European ecosystems	Atmosphere, Geosphere, Biosphere and Sociospher
GBIF	1999	45	Global	To provide free and open access to biodiversity data on a global scale	Biosphere
LifeWatch	2006	8	European	To provide advanced e-Science research facilities for biodiversity and ecosystem studies	Biosphere

Supplementary material

Suppl. material 1: Research Infrastructure Communication Survey

Authors: Erik Kusch, Allan T. Souza
Data type: Survey Form
Brief description: A PDF of the google form used to solicit self-reported metrics of communication & collaboration within and between RIs.
Download file (114.12 kb)