

GBIC2 – A Roadmap for International Biodiversity Informatics - Update

Copenhagen, Denmark, 23-27 July 2018

The GBIC2 workshop¹, *A Roadmap for International Biodiversity Informatics*, will take place at the Biocentre of Copenhagen University (Ole Maaløes Vej 5, 2200 København N), 24-27 July 2018.

The workshop will explore models for a shared mechanism to support collaborative planning and implementing a fully interconnected global infrastructure for biodiversity data. The target is to enable all stakeholders jointly and transparently to develop and agree priorities on a decadal or longer timescale and then work in a coordinated fashion to secure funding and sustainable delivery of all agreed components.

The GBIF Governing Board has allocated funding to make this workshop possible. We are grateful also to the following sponsors that have contributed additional funds to support attendee travel: Atlas of Living Australia, DiSSCo, Field Museum, iDigBio, Museum of Comparative Zoology (Harvard), NLBIF, Pensoft, and United Nations Environment Programme. We also gratefully acknowledge the ability of many invitees to cover their own expenses and the readiness of all attendees to contribute their valuable time to this activity.

An international Steering Committee has been established to ensure the transparency and visibility of the activity:

- Brigitte Baptiste, Alexander von Humboldt Institute, Colombia
- Rob Guralnick, Florida Museum of Natural History, USA
- Eun-Shik Kim, Kookmin University, Korea
- Sandy Knapp, Natural History Museum, UK
- Isayvani Naicker, African Academy of Science, Kenya
- Laetitia Navarro, German Centre for Integrative Biodiversity Research (iDiv), Germany
- Edwin van Huis, Naturalis, Netherlands
- Donald Hobern, GBIF Secretariat, Denmark

[A candidate mechanism for coordinating infrastructure development](#)

The intended goal of the workshop is to develop a proposal for a new mechanism that allows all stakeholders around the world to collaborate openly in building a truly interconnected ecosystem of tools and infrastructures to publish, organize, analyse and use all available information on the world's biodiversity. Achieving this will involve facilitating alignments between investments and energies at all scales, across all regions and in many different domains. There are many challenges in doing this, but the likely benefits are many times greater, enabling or accelerating our understanding of biodiversity patterns, evolutionary history, taxonomy, ecology, conservation, biosecurity risks and food security.

In preparation for the workshop, we here propose a **candidate mechanism** for achieving such coordination. We plan to use the workshop to explore and to replace or revise this proposal, taking into account concerns or

¹ GBIC2 is planned as a workshop with a clear deliverable. The event is named the *Global Biodiversity Informatics Conference 2*, since it continues the work of the first GBIC (which, in fact, was also a workshop).

ideas raised by conference attendees. This candidate mechanism is intended as a simple starting-point for further discussion. All aspects can be reconsidered or changed, subject to the consensus of attendees.

1. Fundamental values

We believe that any effective coordination mechanism will have to satisfy several key criteria. It should be:

1. **Open:** The mechanism should support the principles of Open Science and Open Access by maximizing² free and open access to infrastructure and data.
2. **Transparent:** All processes for deciding and agreeing plans should be open for review and input by all interested stakeholders. The provenance and justification for any recommendations should be clear.
3. **Inclusive:** The mechanism should, to the fullest extent possible, enable full participation by relevant stakeholders, regardless of nationality or language, and with full consideration for issues of diversity and inclusion.
4. **Supportive:** The mechanism should not seek to duplicate the activities of any existing entity that already has a recognized mandate and capacity to coordinate or implement relevant activity.
5. **Sustainable:** The mechanism should facilitate long-term sustainable planning for all essential components of the shared infrastructure, regardless of current arrangements for developing or delivering these components.
6. **Cost-effective:** The costs and resources required to implement mechanism should be appropriate for the scale of the benefits achieved.

Questions for GBIC2:

- Are there other fundamental values which also need to be agreed?

2. Scope and modularity

The first GBIC workshop in 2012 developed a model which identified 20 key **component areas** of importance to the development and implementation of biodiversity informatics. These are presented in the [Global Biodiversity Informatics Outlook \(GBIO\)](#):



² There are legitimate reasons to manage some information and resources may in a more restrictive way (e.g. see the Atlas of Living Australia [report on sensitive data](#)). The focus should nevertheless be on maximizing access wherever feasible.

These component areas provide the starting-point for GBIC2. In particular, we propose that the resulting coordination mechanism should initially aim to deliver detailed implementation plans for each of these 20 component areas. In practice, it is likely that some may be combined as part of a single implementation plan, and it is possible that additional areas may be identified and added for further planning.

Each of the component areas already encompasses the activities of many organizations, agencies, projects and individual researchers. Delivering the GBIO vision for each component area may involve development of new or improved policies, data standards, processes, governance arrangements, software tools, IT infrastructure investments and research programmes. The coordination mechanism should establish the set of investments of all these kinds which will have the greatest effect in synergizing international efforts and delivering a functional global infrastructure.

Questions for GBIC2:

- Do you support using the GBIO component areas as a starting-point for planning infrastructure investments?
- Should the proposed mechanism seek to address all of the GBIO component areas, or are some of these already the responsibility of existing respected organizations? If so, how can we best collaborate with these organizations?
- Does the GBIO framework omit anything which you believe must be added?

3. Components of the mechanism

There are several discrete components which are all required if joint planning and implementation is to be achieved.

First, a process is required to identify, agree and justify a set of priority investments for each component area, including identifying the areas of overlap between component areas and defining each deliverable to the level of detail required to ensure interoperability between the work of different stakeholders. We here use the term **roadmap** to refer to an agreed set of priority investments, including necessary specifications, identification of dependencies and indicative costs. These roadmaps can serve as the basis for an international implementation strategy at a decadal scale.

Secondly, a process is required to coordinate efforts around the world to secure funding and resources for the priority investments identified in these roadmaps. Such a process should facilitate open communication between relevant stakeholders and maintain a shared view of how existing investments are contributing to design or delivery of each component. There should also be processes to enable projects which aim to deliver a shared component to secure letters of support and agree joint activities with other stakeholders around the world, based on the alignment of their activity with the roadmaps. We foresee this to be the most complex aspect of the mechanism in regard to developing a governance model which can support the needs and interests of all stakeholders.

Thirdly, a process is required to oversee the lifecycle of institutional investments in components of shared importance. Ideally, the global community needs to find a way to ensure that important catalogues, vocabularies and ontologies, data repositories, software infrastructures, and analytic services are maintained even when the current hosts are no longer able to maintain support.

Finally, a process is required to support a broader range of stakeholder projects and communities, by providing clear documentation for all agreed standards and best practices for aligning data activities with the collaborative infrastructure. This could extend to offering input into draft project proposals to help to guarantee such alignments.

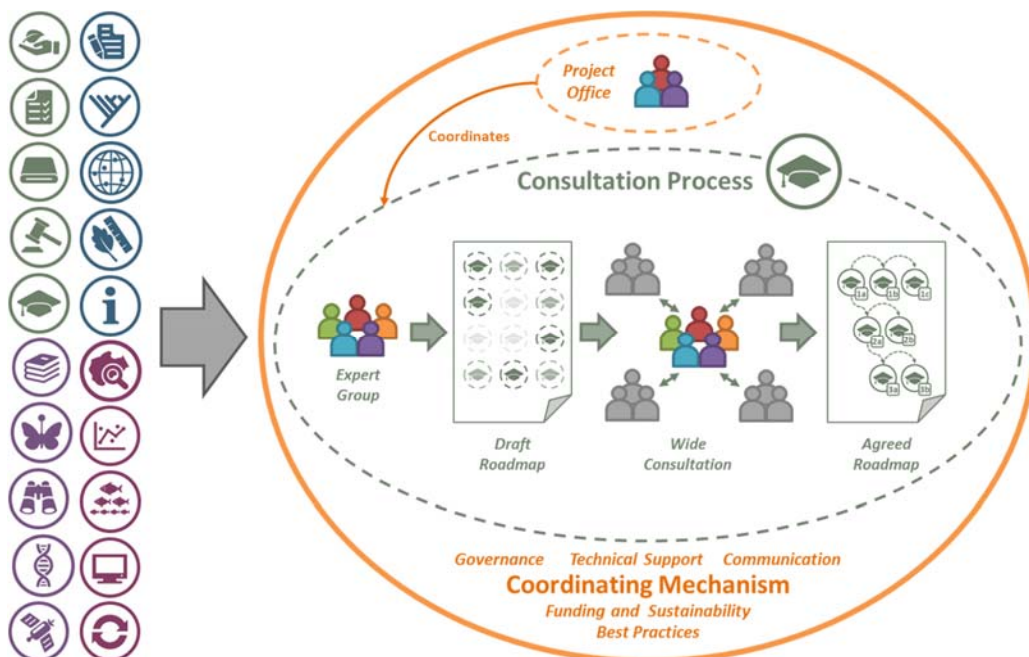
Questions for GBIC2:

- Do you agree that all four components proposed here are necessary?
- Are there other unidentified components which must be included in a coordination mechanism?
- What stakeholders need to be engaged as partners in such a mechanism?
- How can the mechanism operate to offer stability and forward progress within biodiversity informatics, without limiting or excluding innovation or other project plans?

4. How the mechanism could work

Our proposed candidate mechanism would be as follows:

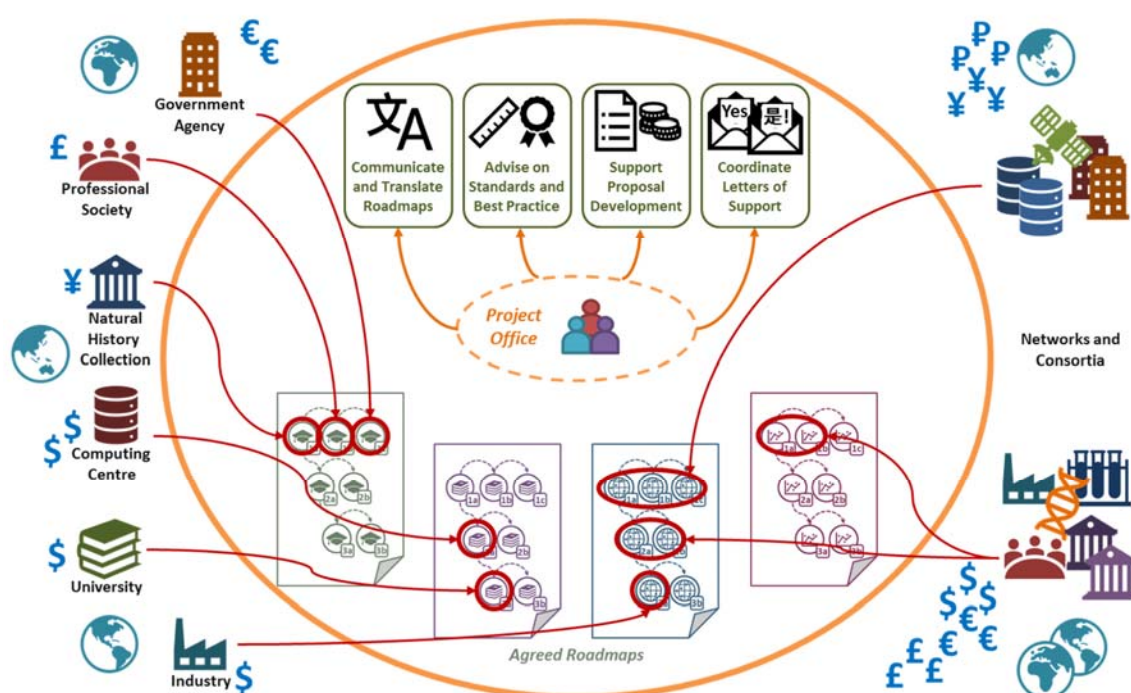
- A small **project office** should be established to support the coordination mechanism. This could initially be a small dedicated team of two or three people hosted by GBIF or some other entity. It could alternatively be a distributed team of individuals contributed or hosted by partner stakeholders around the world. Regardless, this project office would be tasked with operating the coordination mechanism in a neutral and open way on behalf of the entire community.



- The project office would coordinate a series of **international consultations** to establish agreed priority investments for each of the GBIO component areas. These consultations could be structured as 1) a call for international nomination and selection of a panel of relevant experts, 2) a face-to-face workshop for these experts to review the state of existing tools and infrastructure and

to propose a set of possible investments to advance the implementation of the component area, 3) an international, multilingual open consultation phase around the proposals from the expert group, and 4) development of a consensus view of the priority investments which the global community would most value, i.e. a **roadmap** for the component area.

- The project office would coordinate and oversee the translation of the roadmaps into relevant languages, probably focusing on the six UN languages, plus Portuguese, but open to extension according to needs and available funds, and publicise these throughout the biodiversity informatics and biodiversity research communities and to relevant funding agencies. A key role would be to build understanding with funders of the rationale behind the roadmaps and how such investments would interlock for form a greater whole.



- The project office would support efforts by any institution, organization, agency or network to implement any priority investment identified in a roadmap. Such support could include guidance on the requirements for each investment, coordinating interest and efforts between international stakeholders, securing international letters of support, etc.
- The project office should maintain an overview of existing investments and remaining gaps and of the operational lifecycle for time-bound investments. The project office should monitor interest from the community in revisiting and revising roadmap documents after some time, as the landscape and technologies change.

Questions for GBIC2:

- Is a project office the most appropriate way to support a coordination mechanism?

- If so, should the project office be located with an existing organization, or should it be a new entity, or should it be a virtual team comprising individuals from different organizations?
- How can the independence and transparency of such a project office be guaranteed?
- How many people/how much funding is necessary?
- What languages should be supported? At which stages in the process should materials be translated?
- Is it feasible for a project office to support international parallel efforts to secure funding for key components? How is competition to be handled?

The plan for the GBIC2 workshop

The mechanism outlined above is the starting-point for our discussions. We aim to end the conference with an improved model for such a mechanism which reflects the insights of the workshop attendees. Following the conference, we will work with the steering committee to develop a paper outlining the vision for the mechanism and will prepare translations. Our current plan is then to circulate the paper for wider input and improvement from other stakeholders not represented in the workshop. Early in 2019, we would hope to be sufficiently clear for the mechanism to become a reality.

The expected structure of the workshop will be as follows:

Tuesday 24 July afternoon

- Plenary presentations to open the workshop and provide perspectives to guide the working sessions:
 - Simon Hodson, CODATA, on initiatives around open science and research infrastructure
 - Robert Hanisch, NIST, on experiences in other scientific domains in coordinating international planning
 - Maria Uhle, NSF/Belmont Forum, on funders' perspectives in regard to best practice infrastructure investments
 - Donald Hobern, GBIF, on the GBIO model and the proposed candidate mechanism

Wednesday 25 July

- Plenary exploration of the candidate mechanism
- Parallel working sessions to identify issues with the candidate mechanism and to identify possible improvements
- Plenary session to incorporate thinking from the parallel sessions and to develop or refine a version 2 candidate mechanism

Thursday 26 July

- Four parallel working sessions, with attendees trialing the version 2 candidate mechanism by attempting to identify and agree a set of desired priority investments for each of four GBIO component areas:
 - **Biodiversity knowledge network (Culture):** Developing models, social networks, reputation and recognition systems to empower and reward professionals and knowledgeable amateurs to contribute to the curation and improvement of digital biodiversity information. This component area involves mostly sociological aspects of integration and change, bringing together skills from a different set of contributors. See [here](#).

- **Published materials** (Data): Addressing the challenges of offering digital access to centuries of information gathered and stored in non-digital formats, including transformation of this information into structured data for synthesis with existing digital resources. This component area involves a diverse stakeholders with opportunities to increase alignment and reuse of informatics tools, including optical character recognition, gazetteers and vocabularies, and automated translation services. See [here](#).
- **Integrated occurrence data** (Evidence): Aligning activities around the world to aggregate spatial evidence of species occurrence, removing duplication of effort, and ensuring delivery of the most comprehensive integrated data resource possible. This component area offers immediate opportunities for alignment since several key infrastructures are already starting discussions on how to achieve this. See [here](#).
- **Trends and predictions** (Understanding): Delivering time-sensitive analyses of available evidence and developing robust models for past and future scenarios. This component area presents different challenges around credibility, governance and integration of multiple sources of biotic and abiotic information. See [here](#).
- Plenary session to incorporate thinking from the parallel sessions and to develop or refine a version 3 candidate mechanism

Friday 27 July morning

- Plenary discussion of results from previous days and how to proceed, in particular to define the version of a collaboration mechanism supported by the attendees and to plan how to write this up, expand consultation and proceed with implementing the mechanism.