Setting Sustainability Goals for Biodiversity Informatics Infrastructure

Giuditta Parolini ‡

‡ Museum für Naturkunde – Leibniz Institute for Evolution and Biodiversity Science, Berlin, Germany

Corresponding author: Giuditta Parolini (giuditta.parolini@mfn.berlin)

Abstract

Biodiversity data are of crucial importance in the implementation of sustainable development strategies (UN Sustainable Development Goals, EU Taxonomy for Sustainable Activities), because they contribute to deepening our understanding of marine and terrestrial ecosystems and our assessment of the impact of human activities on the planet. Yet, biodiversity informatics itself is lacking an overall sustainability plan that addresses "the three E's of environment, equity, and economics" (Brinkmann 2021) in relation to biodiversity data, their lifecycle, and related technologies. This is especially problematic given that data-driven biodiversity research has had a ten-fold increase (GBIF Secretariat 2022). This implies higher costs for data management, storage and transmission, a larger carbon footprint, and open issues about data access and reuse. SMART (Specific, Measurable, Assignable, Realistic, Time-Related) goals can help to set a sustainability strategy for biodiversity informatics because they bring together the ambitions of an organisation with the action plan necessary to realise these ambitions. When G.T. Doran proposed the use of SMART goals (Doran 1981), his aim was not to create a checklist, but a tool to get results and produce deliverables, which is precisely what is required for implementing an effective sustainability strategy. Three examples are provided here to explain the value of a SMART approach to set sustainability goals in biodiversity informatics.

Goal 1, Increase energy efficiency of computing hardware: A large part of the energy consumption in data centres is due to hardware operations. Several strategies can be pursued to increase energy efficiency in this area, such as activity optimisation for Graphic Processing Units (GPUs), improvements in the effectiveness of cooling systems, and replacement of Central Processing Units (CPUs) with more efficient GPUs. This sustainability goal is specific because it clearly identifies what needs to be achieved; it is measurable because pre- and post- energy consumption are known and can be easily compared with the expected improvement; it is assignable because data centre managers can be identified as the agents of change; it is realistic because there are multiple technological solutions to achieve the goal; and time-related because a roadmap for change stretching through months or years can be set up. Achieving this

goal is evidently beneficial for both the budget and the carbon footprint of biodiversity informatics, as higher energy efficiency is associated to lower running costs and also contributes to reduce greenhouse gases emissions.

Goal 2, Reducing data storage costs per terabyte: Data storage is becoming one of the most significant operating expenses in data-driven organisations due to the considerable increase in the amount of data produced. It is therefore imperative to reduce costs by selecting appropriate storage options (on-premises or cloud services), by adopting compressed formats whenever possible, and by avoiding data redundancy. This goal is SMART because it sets a target (the expected cost reduction can be quantified) and establishes measures (cost per terabyte) to assess its achievement. It can be assigned to data managers, who can choose among different technologies and service providers to lower costs, and can be scheduled over time according to budget needs. The achievement of this goal has clear economic benefits, but also increases the social sustainability of biodiversity informatics, because it makes data-driven research affordable even for organisations with limited financial means.

Goal 3, Decreasing API (Application Programming Interface) response time and bandwidth for data transmission: Data sharing is a key activity in biodiversity informatics and APIs have become the standard technology for achieving it. Improving API performance is therefore crucial to enhance the user experience for stakeholders interested in biodiversity data. In particular, API response time and bandwidth required for data transmission control whether users will be able to access immediately and download satisfactorily the data they are interested in. Goal 3 is SMART because API response time and bandwidth are known measures and can be improved using strategies such as caching and concurrency to handle data requests more rapidly and delivering data in compressed formats to reduce bandwidth. Software developers and project managers involved in API design and maintenance are the people in charge of this goal. They decide the best strategy and timeline to pursue changes. Improving API performance increases the social sustainability of biodiversity informatics because it promotes uncomplicated data access for all stakeholders interested in biodiversity data. It also lowers energy consumption for downloading data benefitting the environment.

Keywords

SMART goals, green IT, data storage, API, bandwidth

Presenting author

Giuditta Parolini

Presented at

TDWG 2023

Hosting institution

Museum für Naturkunde – Leibniz Institute for Evolution and Biodiversity Science, Berlin, Germany

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Brinkmann R (2021) Defining Sustainability. In: Brinkmann R (Ed.) The Palgrave Handbook of Global Sustainability. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-38948-2_1-1
- Doran GT (1981) There is a SMART way to write management's goals and objectives. Management Review 70: 35-36.
- GBIF Secretariat (2022) GBIF Science Review 2021. https://doi.org/10.35035/w3p0-8729