APIs: A Common Interface for the Global Biodiversity Informatics Community

Ben Norton ‡

‡ North Carolina Museum of Natural Sciences, Raleigh, North Carolina, United States of America

Corresponding author: Ben Norton (michaelnorton.ben@gmail.com)

Abstract

Web APIs (Application Programming Interfaces) facilitate the exchange of resources (data) between two functionally independent entities across a common programmatic interface. In more general terms, Web APIs can connect almost anything to the world wide web. Unlike traditional software, APIs are not compiled, installed, or run. Instead, data are read (or *consumed* in API speak) through a web-based transaction, where a client makes a request and a server responds. Web APIs can be loosely grouped into two categories within the scope of biodiversity informatics, based on purpose. First, Product APIs deliver data products to end-users. Examples include the Global Biodiversity Information Facility (GBIF) and iNaturalist APIs. Designed and built to solve specific problems, web-based Service APIs are the second type and the focus of this presentation (referred to as Service APIs). Their primary function is to provide on-demand support to existing programmatic processes. Examples of this type include Elasticsearch Suggester API and geolocation, a service that delivers geographic locations from spatial input (latitude and longitude coordinates) (Pejic et al. 2010).

Many challenges lie ahead for biodiversity informatics and the sharing of global biodiversity data (e.g., Blair et al. 2020). Service-driven, standardized web-based Service APIs that adhere to best practices within the scope of biodiversity informatics can provide the transformational change needed to address many of these issues. This presentation will highlight several critical areas of interest in the biodiversity data community, describing how Service APIs can address each individually. The main topics include:

1. standardized vocabularies,
2. interoperability of heterogeneous data sources and
3. data quality assessment and remediation.

Fundamentally, the value of any innovative technical solution can be measured by the extent of community adoption. In the context of Service APIs, adoption takes two primary forms:

1. financial and temporal investment in the construction of clients that utilize Service APIs and
2. willingness of the community to integrate Service APIs into their own systems and workflows.

To achieve this, Service APIs must be simple, easy to use, pragmatic, and designed with all major stakeholder groups in mind, including users, providers, aggregators, and architects (Anderson et al. 2020Anderson et al. 2020; this study). Unfortunately, many innovative and promising technical solutions have fallen short not because of an inability to solve problems (Verner et al. 2008), rather, they were difficult to use, built in isolation, and/or designed without effective communication with stakeholders. Fortunately, projects such as Darwin Core (Wieczorek et al. 2012), the Integrated Publishing Toolkit (Robertson et al. 2014), and Megadetector (Microsoft 2021) provide the blueprint for successful community adoption of a technological solution within the biodiversity community. The final section of this presentation will examine the often overlooked non-technical aspects of this technical endeavor. Within this context, specifically how following these models can broaden community engagement and bridge the knowledge gap between the major stakeholders, resulting in the successful implementation of Service APIs.

Keywords

Application Programming Interfaces, standardization

Presenting author

Ben Norton

Presented at

TDWG 2021

Conflicts of interest

References

- Anderson RP, Araújo MB, Guisan A, Lobo JM, Martínez-Meyer E, Peterson AT, Soberón JM (2020) Optimizing biodiversity informatics to improve information flow, data quality,

and utility for science and society. Frontiers of Biogeography 12 (3). https://doi.org/10.21425/F5FBG47839
- Blair J, Gwiazdowski R, Borrelli A, Hotchkiss M, Park C, Perrett G, Hanner R (2020) Towards a catalogue of biodiversity databases: An ontological case study. Biodiversity Data Journal 8 https://doi.org/10.3897/bdj.8.e32765
- Microsoft (2021) AI for Earth camera trap image processing API. Github repository of its Megadetector recognizer. URL: https://github.com/Microsoft/CameraTraps
- Pejic B, Pejic A, Covic Z (2010) Uses of W3C's Geolocation API. 2010 11th International Symposium on Computational Intelligence and Informatics (CINTI) https://doi.org/10.1109/cinti.2010.5672223
- 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
- Verner J, Sampson J, Cerpa N (2008) What factors lead to software project failure? 2008 Second International Conference on Research Challenges in Information Science https://doi.org/10.1109/rcis.2008.4632095
- 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