



HAL
open science

Integration of Biodiversity Linked Data and Web APIs using SPARQL Micro-Services

Franck Michel, Catherine Faron Zucker, Fabien Gandon

► To cite this version:

Franck Michel, Catherine Faron Zucker, Fabien Gandon. Integration of Biodiversity Linked Data and Web APIs using SPARQL Micro-Services. Biodiversity Information Standards (TDWG), Aug 2018, Dunedin, New Zealand. 10.3897/biss.2.25481 . hal-01856365

HAL Id: hal-01856365

<https://hal.science/hal-01856365v1>

Submitted on 10 Aug 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Conference Abstract

Integration of Biodiversity Linked Data and Web APIs using SPARQL Micro-Services

Franck Michel[‡], Catherine Faron-Zucker[‡], Fabien Gandon[‡][‡] Université Côte d'Azur, CNRS, Inria, I3S, Sophia-Antipolis, FranceCorresponding author: Franck Michel (franck.michel@cirs.fr)

Received: 03 Apr 2018 | Published: 22 May 2018

Citation: Michel F, Faron-Zucker C, Gandon F (2018) Integration of Biodiversity Linked Data and Web APIs using SPARQL Micro-Services. Biodiversity Information Science and Standards 2: e25481.

<https://doi.org/10.3897/biss.2.25481>

Abstract

Web APIs (Application Programming Interface) are a common means for Web portals and data producers to enable HTTP-based, machine-processable access to their data. They are a prominent source of information*1 pertaining to topics as diverse as scientific information, social networks, entertainment or finance. The methods of Linked Data (Heath and Bizer 2011) similarly aim to publish machine-readable data on the Web, while connecting related resources within and between datasets, thereby creating a large distributed knowledge graph. Today, the biodiversity community is increasingly adopting the Linked Data principles to publish data such as trait banks, museum collections and taxonomic registers (Parr et al. 2016, Baskauf et al. 2016). However, standard approaches are still missing to combine disparate representations coming from both Linked Data interfaces and the manifold Web APIs that were developed during the last two decades to expose legacy biodiversity databases on the Web.

The *SPARQL Micro-Service* architecture (Michel et al. 2018) tackles the goal of reconciling Linked Data interfaces and Web APIs. It proposes a lightweight method to query a Web API using SPARQL (Harris and Seaborne 2013), the Semantic Web standard to query knowledge graphs expressed in the Resource Description Framework (RDF). A SPARQL micro-service provides access to a small RDF graph, typically resource-centric, that it builds at run-time by transforming a fraction of the whole dataset served by the Web API into RDF triples. Furthermore, Web APIs traditionally rely on internal, proprietary resource

identifiers that are unsuited for use as Uniform Resource Identifiers (URIs). To address this concern, a SPARQL micro-service can assign a URI to a Web API resource, allowing an application to look up this URI and get a description of the resource in return (this process is referred to as *dereferencing*).

In this demo, we wish to showcase the value of SPARQL micro-services in the biodiversity domain. We first query TAXREF-LD, a Linked Data representation of the French taxonomic register of living beings (Michel et al. 2017), to retrieve information about a given taxon. Then, we demonstrate how we can enrich our knowledge about this taxon with various types of data retrieved on-the-fly from multiple Web APIs:

- trait data from the Encyclopedia of Life trait bank (Parr et al. 2016),
- articles or books from the [Biodiversity Heritage Library](#),
- audio recordings from the [Macaulay scientific media archive](#),
- photos from the [Flickr photography social network](#), and
- music tunes from [MusicBrainz](#).

Different visualizations are demonstrated, ranging from raw RDF triples to Web pages generated dynamically and integrating heterogeneous data, as suggested in Fig. 1. Depending on the audience's interests, we shall touch upon the alignment of Web APIs' proprietary vocabularies with well-adopted thesauri or ontologies, or more technical concerns *e.g.* related to the effort required to deploy a new SPARQL micro-service.

SPARQL response

```

prefix foaf: <http://xmlns.com/foaf/0.1/>
prefix schema: <http://schema.org/>

<http://taxref.mnh.fr/ld/taxon/56878/fo:fo>
schema:subjectOf <http://erebe-vn2.l3s.unice.fr/ld/flickr/photo/3117389893/> ;
schema:subjectOf <http://erebe-vn2.l3s.unice.fr/ld/flickr/photo/3117389151/> ;
...
foaf:depiction <https://farm6.staticflickr.com/5617/3117389893_e6d5c5a38a_2.jpg> ;
foaf:depiction <https://farm6.staticflickr.com/5567/3117389151_6108f9a5d5_2.jpg> ;
...
schema:contentUrl <https://download.ams.birds.cornell.edu/api/v1/asset/131336/audio/> ;
schema:contentUrl <https://download.ams.birds.cornell.edu/api/v1/asset/122866/audio/> ;
...
schema:subjectOf <https://musicbrainz.org/work/3ff62fec-7d58-44cb-0b5a-5f5876e12b56/> ;

```

Traits (Eol trait bank)

trait	value
clutch/broodsize	1-16
size	1-1000
body length (TV)	2,430-365 mm (adult)
gestation period	290-300 days
duration	9-27 months
body mass	30,369-59 g
onset of fertility	1,472-2 days
habitat	1/Urban 1,754 days (terrestrial)
egg at hatch	58-122 months
reproduction	55-122 months
parental care	No Parental Care No Parental Care

On the Common Dolphin, *Delphinus delphis* Linn. Proc. of the Zoological Society of London. (Biodiversity Heritage Library)

5. On the Common Dolphin, *Delphinus delphis*, Linn. By Wessale Haver Flower, F.R.S., F.Z.S.

It is somewhat remarkable that no well-known species of cetacean is so far from being entirely terrestrial as the Common Dolphin, *Delphinus delphis*, Linn. The last work which I have mentioned is that of Richardson's "Fauna and Ichthyology of the Strait of Gibraltar" in "Trans. Geogr. Soc. London," No. 4, p. 11, 1828. It is there stated that the Common Dolphin is a species of the genus *Delphinus*, Linn. This genus, however, is not entirely terrestrial, as the dolphins of the Atlantic coast, in the straits of the Channel, are sometimes seen on the shore, and are sometimes taken on the shore. It is also stated that the Common Dolphin is a species of the genus *Delphinus*, Linn. The last work which I have mentioned is that of Richardson's "Fauna and Ichthyology of the Strait of Gibraltar" in "Trans. Geogr. Soc. London," No. 4, p. 11, 1828. It is there stated that the Common Dolphin is a species of the genus *Delphinus*, Linn. The last work which I have mentioned is that of Richardson's "Fauna and Ichthyology of the Strait of Gibraltar" in "Trans. Geogr. Soc. London," No. 4, p. 11, 1828. It is there stated that the Common Dolphin is a species of the genus *Delphinus*, Linn.

MusicBrainz

Delphinus delphis

composer: Louis Vintury
part of: M&M (Ocean genus recordings)

Date	Title	Artist	Length
1970-01-01	Delphinus Delphis	Audite Vintury	4:51

Figure 1.

Visualization of multiple types of data retrieved from Web APIs about taxon *Delphinus delphis*.

Keywords

Web API, SPARQL, data integration, linked data, JSON-LD

Presenting author

Franck Michel is a research engineer at the University Côte d'Azur, CNRS, France. His research topics notably concern the integration and federation of heterogeneous data sources using Semantic Web technologies, and their publication in the Web of Data.

References

- Baskauf S, Wieczorek J, Deck J, Webb C (2016) Lessons Learned from Adapting the Darwin Core Vocabulary Standard for Use in RDF. *Semantic Web – Interoperability, Usability, Applicability* 7 (6): 617-627.
- Harris S, Seaborne A (2013) SPARQL 1.1 Query Language. W3C Recommendation.
- Heath T, Bizer C (2011) *Linked Data: Evolving the Web into a Global Data Space*. 1st. Morgan & Claypool
- Michel F, Gargominy O, Terceire S, Faron-Zucker C (2017) A Model to Represent Nomenclatural and Taxonomic Information as Linked Data. Application to the French Taxonomic Register, TAXREF. Proceedings of the 2nd International Workshop on Semantics for Biodiversity (S4BioDiv). CEUR, 1933
- Michel F, Faron-Zucker C, Gandon F (2018) SPARQL Micro-Services: Lightweight Integration of Web APIs and Linked Data. Proceedings of the Linked Data on the Web Workshop (LDOW2018).
- Parr C, Schulz K, Hammock J, Wilson N, Leary P, Rice J, Corrigan Jr R (2016) TraitBank: Practical semantics for organism attribute data. *Semantic Web* 7 (6): 577-588. <https://doi.org/10.3233/SW-150190>

Endnotes

*1 19,300+ Web APIs are registered on ProgrammableWeb.com as of March 2018.