

Wikidata Type Specimen Data Model Proposal - DRAFT

A data model for type specimens in Wikidata

Drafted and peer-reviewed by BHL's Wikimedia Working Group Members and Volunteers

Community Investigation

- Should we model type specimens as [instance of \(P31\) type specimen \(Q51255340\)](#) or [holotype \(Q1061403\)](#)?

On how to model holotype on type specimen, 3 options below:

Option 1

Statements

instance of	holotype	Ouratea sipaliwiniensis	edit
	of		
	0 references		+ add reference

Option 2

	Ouratea sipaliwiniensis	holotype	edit
	object has role		
	1 reference		+ add value

Option 3

Statements

instance of	 type specimen 
	▼ 1 reference
	reference URL http://n2t.net/ark:/65665/3d25cb92-b-ce95-4e84-8263-fc568e0fb767
	+ add reference
+ add value	

Coupled with this statement

subject has role	 holotype 
	of Okeanosaster hohonui
	▼ 1 reference
	reference URL http://n2t.net/ark:/65665/3d25cb92-b-ce95-4e84-8263-fc568e0fb767
+ add reference	
+ add value	

- **What is the difference between holotype and type specimen?**

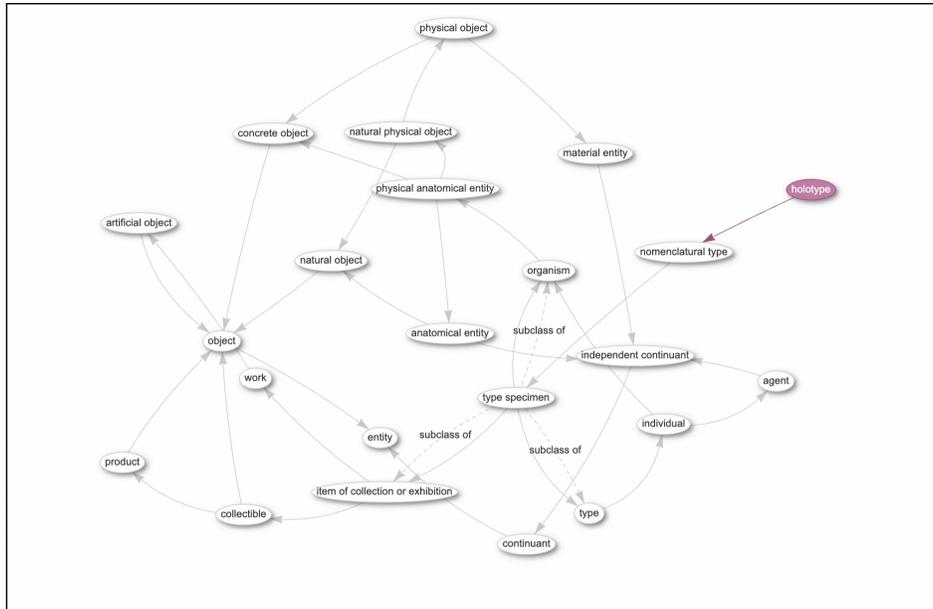
Holotype is a designation of Type Specimen e.g. holotype is a “subclass” of the main element “type specimen”

- Holotype – a single specimen expressly designated as the name-bearing “type” by the original author of the species.
- Syntype – one of several specimens in a series of equal rank used to describe the new species where the author has not designated a single holotype. Thus each specimen in the series is known as a syntype (from which neither a holotype nor a lectotype has been designated).
- Lectotype – a single specimen selected from a group of syntypes and designated as the name-bearing type some time after the original description was published.
- Paratype – representative specimen(s), other than the holotype, in the type series referred to in the original description.
- Paralectotype – the type specimens remaining after a lectotype is designated.
- Neotype – a substitute specimen selected in special circumstances to replace the holotype after the species was first described and the original holotype has

been lost or destroyed (e.g. fire, war etc.).

- Cotype – no longer used; formerly used for either syntype or paratype.
- Allotype – a designated specimen of opposite sex to that of the holotype.

From Richard, holotype tree view showing that it is a subclass of type specimen



- **What properties are core for type specimen?**

- **From Rod on Telegram:** Let's start by looking at the common properties already in use: <https://w.wiki/9W97>

Property	Description	Count
instance of	that class of which this subject is a particular example and member; different from P279 (subclass of); for example: K2 is an instance of mountain; volcano is a subclass of mountain (and an instance of volcanic landform)	1351
collection	art, museum, archival, or bibliographic collection the subject is part of	1134
subject has role	role/generic identity of the item ("subject"), also in the context of a statement. For the role of the value of the statement	1046

	("object"), use P3831 ("object has role"). For acting roles, use P453 ("character role"). For persons, use P39.	
inventory number	identifier for a physical object or a set of physical objects in a collection	1033
location of discovery	where the item was located when discovered	888
location	location of the object, structure or event. In the case of an administrative entity as containing item use P131. For statistical entities use P8138. In the case of a geographic entity use P706. Use P7153 for locations associated with the object	867
image	image of relevant illustration of the subject; if available, also use more specific properties (sample: coat of arms image, locator map, flag image, signature image, logo image, collage image)	784
URL	location of a resource	700
time period	time period (historic period or era, sports season, theatre season, legislative period etc.) in which the subject occurred	130
quantity	number of instances of this subject in the universe of the subject (the actual number of instances in Wikidata may be lower or higher)	104
sex or gender	sex or gender identity of human or animal. For human: male, female, non-binary, intersex, transgender female, transgender male, agender, etc. For animal: male organism, female organism. Groups of same gender use subclass of (P279)	101
described by source	work where this item is described	98
type locality (geology)	the locality where a particular rock type, stratigraphic unit or mineral species is defined from (can coincide but not the same as p189)	73
discoverer or inventor	subject who discovered, first described, invented, or developed this discovery or invention or scientific hypothesis or theory	67
student	notable student(s) of the subject individual	30
time of discovery or invention	date or point in time when the item was discovered or invented	16

- **Where to put identifiers and which ids are crucial to have?**
 - From Rod/Sandra (P217 - inventory number; or combo P195 (collection) +P217 (inventory number) qualifier)

- **Who else needs to be a part of this discussion from biodiversity wiki world?**
 - [WikiProject taxonomy](#)
 - [WikiProject biodiversity](#)

- **What are some good examples that we can look at?**
 - **From Siobhan:** Example of a holotype specimen
<https://www.wikidata.org/wiki/Q136152297>
 - **From Siobhan:** Example of a **species taxon** Wikidata item with more than basic information -
<https://www.wikidata.org/wiki/Q21244160> This is an item I've added more detailed data to. But there are still other properties that could be added depending on the taxon - for example [invasive to](#), [taxon range](#), [taxon range map image](#), [named after](#), [different from described by source](#), [described at URL](#) etc. Likely there are multiple other properties that could also be used so best practice documentation is again needed!

- **Siobhan - Best practices to add references to taxonomic names?**

Rod Page's guidelines for adding references to taxonomic names may be useful
https://www.wikidata.org/wiki/Wikidata_talk:WikiProject_Taxonomy/Archive/2021/10

- **JJ - is there already an agreed upon best practice for modelling taxon?**
 - See [darwin element set for taxon](#);
 - And [GBIF Species API outputs](#); what other relevant fields here?
 - Useful resource [link](#)

- **Best practices to associate a type to a taxon name?**
 - From Rod: maybe P2868 - subject has role?

- **Siobhan - How to link taxon names to synonyms in Wikidata?**
 - - **How to link holotype to isotypes?**

Overview

Give a more detailed introduction to the data model.

Explain why the model was created and who the intended audience is.

Conceptual Model

Describe the conceptual model of the data, including entities, attributes, relationships, and constraints.

Use diagrams such as entity-relationship diagrams (ERDs) to visually represent the model.

Terminology

Define key terms and concepts used in the data model.

Provide explanations and examples to clarify any potentially confusing terminology.

Data Dictionary

Include a data dictionary that lists all entities, attributes, and relationships in the model.

For each item in the data dictionary, provide a description, data type, and any relevant constraints.

Data Quality and Governance

Address data quality considerations and how they are managed within the model.

Discuss data governance practices and policies that ensure the integrity and security of the data.

Examples and Use Cases

Provide examples and use cases to illustrate how the semantic data model is applied in practice.

Show sample queries or transactions to demonstrate how data is retrieved and manipulated.

Implementation Guidance

Offer guidance on implementing the semantic data model, including best practices and recommended tools or technologies.

Discuss considerations for scaling the model and adapting it to evolving requirements.

References and Resources

Provide references to any external sources or standards that influenced the design of the data model.

Include links to additional resources for further reading or exploration.

Appendices

Include any supplementary information, such as additional diagrams, technical specifications, or data samples.