

# Qualitative Investigation of the Disappearance of Classes and Individuals with Newer Versions of Terminologies

Anette Ganske

January 2024

## 1 Investigation of the Changes of Terminologies in the TIB ESS collection

The histories of the 35 terminologies (Status as of 25.01.2024) in the ESS collection of the TIB TS (see Table 1) were investigated with several methods. For three terminologies (ABCD, SSN, and SOSA) we could not find earlier versions than the one published.

We only searched for deleted classes or individuals and did not analyse other changes e.g. of the description of the terms. We were aiming at a qualitative and not a quantitative analysis, i.e. we only investigated consecutive earlier versions of a respective terminology until we found deleted classes or individuals.

Many ontologies in the ESS which were suggested by SGN are stored in Bioportal<sup>1</sup>, as they were originally made for biological research. Bioportal stores for all terminologies earlier versions and provides for each version a file with the differences of it to the antecedent version. Therefore, we used these files for our investigation if applicable (see Table 1).

### 1.1 Tools

The following tools were used:

**Bubastis**<sup>2</sup>: EML-EBI tool for comparing two OWL ontologies, e.g. two different versions of an ontology. The tool writes the results of its comparisons (changed, new and deleted classes) to an xml file. Bioportal uses bubastis to show the changes between two successive versions of terminologies.

**Semantic Diff**<sup>3</sup>: analyses the terminology timeline on github and gitlab. It writes all added and removed terms to an xml report file.

---

<sup>1</sup><https://bioportal.bioontology.org/ontologies>

<sup>2</sup><http://www.ebi.ac.uk/efo/bubastis/>

<sup>3</sup><https://terminology.nfdi4ing.de/ts/sandbox/semanticdiff>

**Protégé**<sup>4</sup> (version 5.6.2): also provides a comparison tool for ontologies. We chose "make a resonable attempt to find refactor operations" and searched for "deprecate and replace patterns".

In addition, lists of changes provided for specific terminologies were analysed manually, see Tables 2 and 3.

## 1.2 Versioning

There is no standard way of versioning terminologies. For some terminologies semantic versioning<sup>5</sup> is used, i.e. they have version numbers x.y.z like those used for software. For other terminologies, the date of publication may be used to differentiate between different versions. Some terminologies published on github don't have version numbers, but the changes to these terminologies are stored in the timeline. The versions (if any) of the studied terminologies can be found in the third column of Tables 2 and 3.

## 1.3 Results

With these tools we were able to analyse 32 out of 35 terminologies: for three terminologies (ABCD, SSN and SOSA) we could not find earlier versions.

Where applicable, we either compared different versions of a terminology or analysed the timeline with the tracking of changes. In six of the 32 terminologies, we found classes or individuals that had been deleted. Additional checks showed that the corresponding URLs were no longer resolvable, see Tables 2 and 3. As Bubastis and Semantic Diff do not automatically check whether URLs or purls still resolve, there might be additional classes and individuals with broken links in the analysed terminologies.

Some of the changes were found in very early versions that may not have been used to annotate data. Also, FAIR terminology issues and practices have evolved over the years, so there is a high probability that fewer classes will be deleted for future versions. Nevertheless, there is still a risk that terms of terminologies won't be available after some time.

---

<sup>4</sup><https://protege.stanford.edu/>

<sup>5</sup><https://semver.org/>

| Name of Terminology  | Abbreviation | Address   |
|--|--------------|---|
| ABCD Base Ontology   | ABCD         | <a href="https://github.com/tdwg/abcd/tree/master/ontology/abcd_concepts.owl">https://github.com/tdwg/abcd/tree/master/ontology/abcd_concepts.owl</a>                                   |
| The Biological Collections Ontology                                    | BCO          | <a href="https://biportal.bioontology.org/ontologies/BCO">https://biportal.bioontology.org/ontologies/BCO</a>   |
| Chemical Entities of Biological Interest                               | CHEBI        | <a href="https://biportal.bioontology.org/ontologies/CHEBI">https://biportal.bioontology.org/ontologies/CHEBI</a>   |
| Darwin Core Chronometric Age Extension vocabulary                      | CHRONO       | <a href="http://rs.tdwg.org/chrono.htm">http://rs.tdwg.org/chrono.htm</a>   |
| DFG Classification of Subject Areas Ontology                           | DFGFO        | <a href="https://raw.githubusercontent.com/tibonto/DFG-Fachsystematik-Ontology/main/dfgfo.ttl">https://raw.githubusercontent.com/tibonto/DFG-Fachsystematik-Ontology/main/dfgfo.ttl</a> |
| Data Privacy Vocabulary  | DPV          | <a href="https://www.w3.org/community/dpvcg/">https://www.w3.org/community/dpvcg/</a>   |
| Darwin Semantic Web  | DSW          | <a href="https://github.com/darwin-sw/dsw">https://github.com/darwin-sw/dsw</a>   |
| The Data Use Ontology  | DUO          | <a href="https://biportal.bioontology.org/ontologies/DUO">https://biportal.bioontology.org/ontologies/DUO</a>   |
| Darwin Core main vocabulary  | DWC          | <a href="http://rs.tdwg.org/dwc.htm">http://rs.tdwg.org/dwc.htm</a>   |
| The Ecosystems Ontology  | ECOS8        | <a href="https://biportal.bioontology.org/ontologies/ECOS8">https://biportal.bioontology.org/ontologies/ECOS8</a>   |
| Bioinformatics operations, data types, formats, identifiers and topics | EDAM         | <a href="https://biportal.bioontology.org/ontologies/EDAM">https://biportal.bioontology.org/ontologies/EDAM</a>   |
| Experimental Factor Ontology   | EFO          | <a href="https://biportal.bioontology.org/ontologies/EFO">https://biportal.bioontology.org/ontologies/EFO</a>   |
| The Environment Ontology   | ENVO         | <a href="https://biportal.bioontology.org/ontologies/ENVO">https://biportal.bioontology.org/ontologies/ENVO</a>   |
| Environmental Thesaurus  | EnvThes      | <a href="https://biportal.bioontology.org/ontologies/ENVTHER">https://biportal.bioontology.org/ontologies/ENVTHER</a>   |
| Ecological Trait-data Standard (ETS)                                   | ETS          | <a href="https://github.com/EcologicalTraitData/ETS">https://github.com/EcologicalTraitData/ETS</a>   |
| Friend of a Friend vocabulary  | FOAF         | <a href="http://xmlns.com/foaf/spec/">http://xmlns.com/foaf/spec/</a>   |
| General Multilingual Environmental Thesaurus                           | GEMET        | <a href="https://www.eionet.europa.eu/gemet/en/themes/">https://www.eionet.europa.eu/gemet/en/themes/</a>   |
| RDF/OWL vocabulary for representing spatial information                | GeoSPARQL    | <a href="https://opengeospatial.github.io/ogc-geosparql/geosparql11/index.html">https://opengeospatial.github.io/ogc-geosparql/geosparql11/index.html</a>                               |
| Gemeinsame Normdatei Ontology  | GND          | <a href="https://d-nb.info/standards/elementset/gnd">https://d-nb.info/standards/elementset/gnd</a>   |
| ISO 3166-1 Country Codes Ontology                                      | OM           | <a href="https://biportal.bioontology.org/ontologies/OM">https://biportal.bioontology.org/ontologies/OM</a>   |
| The Phenotype And Trait Ontology                                       | PATO         | <a href="https://biportal.bioontology.org/ontologies/PATO">https://biportal.bioontology.org/ontologies/PATO</a>   |
| Plant Ontology   | PO           | <a href="https://biportal.bioontology.org/ontologies/PO">https://biportal.bioontology.org/ontologies/PO</a>   |
| Provenance Ontology  | PROVO        | <a href="https://biportal.bioontology.org/ontologies/PROVO">https://biportal.bioontology.org/ontologies/PROVO</a>   |
| Plant Trait Ontology   | PTO          | <a href="https://biportal.bioontology.org/ontologies/PTO">https://biportal.bioontology.org/ontologies/PTO</a>   |
| Schema.org Vocabulary  | schemaorg    | <a href="https://schema.org/">https://schema.org/</a>   |
| Semantic Sensor Network Ontology                                       | SSN          | <a href="https://biportal.bioontology.org/ontologies/SSN">https://biportal.bioontology.org/ontologies/SSN</a>   |
| Semantic Web for Earth and Environment Technology Ontology             | SWEET        | <a href="https://biportal.bioontology.org/ontologies/SWEET">https://biportal.bioontology.org/ontologies/SWEET</a>   |
| Semanticscience Integrated Ontology                                    | SIO          | <a href="https://biportal.bioontology.org/ontologies/SIO">https://biportal.bioontology.org/ontologies/SIO</a>   |
| Sensor, Observation, Sample, and Actuator Ontology                     | SOSA         | <a href="https://www.w3.org/TR/vocab-ssn/">https://www.w3.org/TR/vocab-ssn/</a>   |
| Systems Biology Ontology   | SBO          | <a href="https://biportal.bioontology.org/ontologies/SBO">https://biportal.bioontology.org/ontologies/SBO</a>   |
| The Extensible Observation Ontology                                    | OBOE         | <a href="https://biportal.bioontology.org/ontologies/OBOE">https://biportal.bioontology.org/ontologies/OBOE</a>   |
| OWL-Time   | time         | <a href="https://biportal.bioontology.org/ontologies/TIME">https://biportal.bioontology.org/ontologies/TIME</a>   |
| Unified Astronomy Thesaurus (UAT)                                      | UAT          | <a href="https://github.com/astrothesaurus/UAT">https://github.com/astrothesaurus/UAT</a>   |
| Units of Measurement Ontology  | UO           | <a href="https://biportal.bioontology.org/ontologies/UO">https://biportal.bioontology.org/ontologies/UO</a>   |

Table 1: List the long and short names of the investigated terminologies and their respective URLs, which we used for the comparisons.

| Terminology | Versions                           | Investigation Tool  | Results  |
|-------------|------------------------------------|---|--|
| ABCD        | Only one version available.        |   |  |
| BCO         | 2021-11-14 and 2015-04-10          | bubastis  | Only classes from other ontologies are deleted, own classes are deprecated.  |
| CHEBI       | 227 and 228                        | bubastis  | Deleted classes are moved within class tree, IRI still resolvable.   |
| CHRONO      | 2020-09-14 - 2021-02-21            | bubastis  | No deleted classes were found.   |
| DFGFO       | since 2005-10-05                   | Semantic Diff   | No deleted classes were found.   |
| DPV         | v1.0 and v0.82                     | bubastis  | <b>18 classes deleted. Older version might not have been published.</b>  |
| DSW         | 0.4 and 1.01                       | Semantic Diff   | <b>Deleted Classes, marked as deprecated, but purils do no longer resolve.</b>   |
| DUO         |                                    | bubastis  | Deleted classes were replaced by imported classes, original URIs are still resolvable.   |
| DWC         | 2008-11-19 and 2023-09-18          | Protégé   | No deleted classes were found.   |
| ECSO8       | 0.3 and 0.9                        | bubastis  | <b>Classes were deleted, respective URIs and also URIs of imported ontologies (ENVO and UO) don't resolve any longer.</b>  |
| EDAM        | 1.24 and 1.24                      | bubastis  | Classes are marked as deprecated, URIs still resolve.  |
| EFO         | 3.60 - 3.62                        | bubastis  | IRIs of deleted classes are still reachable (only a few checked). Comparison last two versions: All deleted classes are taken from MONDO ontology. Classes are there marked as deprecated. |
| ENVO        | 2021-05-14 and 2023-02-13          | bubastis  | <b>Many deleted classes, one not-resolving URI found.</b>  |
| EnvThes     | 5.03 - 5.05                        | bubastis  | No changes of classes between the last 3 archived versions   |
| ETS         | 0.9 and 0.10                       | protégé   | No deleted classes were found.   |
| FOAF        | 20100809 and 20140114              | <a href="http://xmlns.com/foaf/spec/#sec-changes">http://xmlns.com/foaf/spec/#sec-changes</a>               | No deleted classes were found.   |
| GEMET       | changes after 2004                 | <a href="https://www.eionet.europa.eu/gemet/en/changes/">https://www.eionet.europa.eu/gemet/en/changes/</a> | No deleted classes were found.   |
| GeoSPARQL   | 1.0 and 1.1                        | Protégé   | No deleted classes were found.   |
| GND         | 20230606 and 20230626              | bubastis  | Only deleted properties were found.  |
| OM          | 2.045 - 2.050                      | bubastis  | No deleted classes were found.   |
| PATO        | 2023-02-17 and 2023-05-18          | bubastis  | Only imported classes were deleted, respective URIs are still resolvable.  |
| PO          | 2019-06-05 - 2023-07-13            | bubastis  | No deleted classes were found.   |
| PROVO       | uploaded 07/23/2013 and 08/16/2013 | bubastis  | Classes were deleted, all URIs still resolve.  |
| PTO (TO)    | 9.2018-7.2023                      | bubastis  | No deleted classes were found.   |

Table 2: List of investigated terminologies, their respective URLs, the compared versions, the used tool and the result of the search for deleted classes

| Terminology | Versions                    | Investigation Tool  | Results  |
|-------------|-----------------------------|---|--|
| Schemaorg   | 0.91 - 24.0                 | <a href="https://schema.org/docs/releases.html">https://schema.org/docs/releases.html</a> | Deprecated classes are still reachable.  |
| SSN         | Only one version available. |   |  |
| SWEET       | 3.4.0 - 3.6.0               | protégé   | Deleted classes are marked as deprecated and are still reachable.                        |
| SIO         | 1.55 - 1.59                 | bubastis  | No deleted classes were found.   |
| SOSA        | Only one version available. |   |  |
| SBO         | 03:02:2018 and 28:08:2021   | bubastis  | URIs of deleted classes were changed from purl to internal IDs and do no longer resolve. |
| OBOE        | 1.1 and 1.2                 | bubastis  | No deleted classes were found.   |
| Time        | 09/18/2017 and 04/17/2020   | bubastis  | <b>The URI of one deleted class is not resolving.</b>                                    |
| UAT         | 4.2 and 5.0                 | protégé   | <b>2 deleted individuals where URIs don't resolve.</b>                                   |
| UO          | 06/10/2022 and 08/09/2023   | bubastis  | No deleted classes were found.   |

Table 3: Continuation of Table 2: List of investigated terminologies, their respective URIs, the compared versions, the used tool and the result of the search for deleted classes