

Software Ontologies and Metadata Schemes References

Morane Gruenpeter

September 2017

A list for collecting and organizing software ontologies and metadata schemes describing software components and software source code.

1 Preliminary references

1. The term ontology [1]
2. Encyclopedia of Database Systems [2]
3. Software citation principles [3]

2 dedicated software ontologies

1. DOAP- Description Of A Project[4]
2. SEON - Software Evolution ONtologies [5]
3. ADMS.SW- Asset Description Metadata Schema for Software
https://joinup.ec.europa.eu/asset/adms_foss/description
4. CodeMeta- the rosetta stone of software ontologies [6]

3 scholarly ecosystem

1. OntoSoft for scientific software[7]
2. DataCite <https://schema.datacite.org/>
3. TOTEM - the Trustworthy Online Technical Environment Metadata Registry <http://www.keep-totem.co.uk/>

4 linked data

1. Wikidata [8]
 - Software- <https://www.wikidata.org/wiki/Q7397>
2. schema.org :
 - *SoftwareApplication* - <http://schema.org/SoftwareApplication>
 - *SoftwareSourceCode*- <http://schema.org/SoftwareSourceCode>

5 digital preservation schemes

1. Dublin Core (DCMI)- <http://dublincore.org/documents/2012/06/14/dcmi-terms/?v=terms#dcmitype-Software>
2. PRONOM -
<https://www.wikidata.org/wiki/User:YULdigitalpreservation/Software>
3. Marc - <http://www.loc.gov/marc/>

6 package management vocabularies

Found mostly in the source code itself as manifest files. The name of the file is defined bellow by the context or development ecosystem in which it is used:

Table 1: manifest files for package management

context	filename	in CodeMeta
java- Maven	pom.xml	yes
Octave	DESCRIPTION	yes
R package	DESCRIPTION	yes
ruby gems	.gemspec or Rakefile	yes
Javascript npm	package.json	yes
Perl CPAN::META	META.json, META.yml, .sDpec	yes
Dart	pubspec.yaml	no
Debian package	debian/upstream/metadata	yes
puppet	metadata.json	no
PyPI	setup.py	yes
Scientific software	CITATION	no
CodeMeta	CODE, code.json, codemeta.json	yes
Java gradle	gradle.properties	no
Jekyll	_config.yml -no	no
clojure	project.clj or build.boot	no
haskell	project.cabal	no
scala	build.sbt	no
Ocaml	opam	no

A list of the most popular terms used in the package management domain:

```
package_management_popular_terms = {
  "developmentStatus": [],
  "version": [],
  "operatingSystem": [],
  "description": [],
  "keywords": [],
  "issueTracker": [],
  "name": [],
  "author": [],
  "relatedLink": [],
  "url": [],
  "type": [],
  "license": [],
  "maintainer": [],
  "email": [],
```

```
        "softwareRequirements": [],
        "identifier": [],
        "codeRepository": []
    }
```

References

- [1] Wikipedia. Ontology (information science) — wikipedia, the free encyclopedia, 2017. [Online; accessed 26-May-2017].
- [2] Ling Liu and M. Tamer Özsu, editors. *Encyclopedia of Database Systems*. Springer US, 2009.
- [3] Arfon M Smith, Daniel S Katz, and Kyle E Niemeyer. Software citation principles. *PeerJ Computer Science*, 2:e86, 2016.
- [4] Edd Wilder-James. *DOAP: Description of a project*, 2007.
- [5] Michael Würsch, Giacomo Ghezzi, Matthias Hert, Gerald Reif, and Harald C. Gall. Seon: a pyramid of ontologies for software evolution and its applications. *Computing*, 94(11):857–885, 2012.
- [6] Matthew B Jones, Carl Boettiger, Abby Cabunoc Mayes, Arfon Smith, Peter Slaughter, Kyle Niemeyer, Yolanda Gil, Martin Fenner, Krzysztof Nowak, Mark Hahnel, Luke Coy, Alice Allen, Mercè Crosas, Ashley Sands, Neil Chue Hong, Patricia Cruse, Dan Katz, and Carole Goble. Codemeta: an exchange schema for software metadata. knb data repository, 2016.
- [7] Yolanda Gil, Varun Ratnakar, and Daniel Garijo. Ontosoft: Capturing scientific software metadata. In *Proceedings of the Eighth ACM International Conference on Knowledge Capture*, Palisades, NY, 2015.
- [8] Wikidata community. Wikidata, 2016.