Generating CodeMeta Metadata for R Packages

Carl Boettiger[1]

1 University of California, Berkeley

Summary

The CodeMeta project defines a JSON-LD format ("JSON-Ld 1.1" 2017) for describing software metadata, based largely on `schema.org` terms. This metadata format is being adopted by many leading archives for scientific software, including DataCite, Zenodo, and DataONE to address many of the needs identified in the NIH report on the need for a "Software Discovery Index" ("Software Discovery Index Meeting Report," n.d.). Many common software metadata formats have been mapped into CodeMeta by means of a crosswalk table ("CodeMeta Project Crosswalk" 2017), also implemented in this package. The `codemetar` package provides utilities to generate and validate these `codemeta.json` files automatically for R packages by parsing the DESCRIPTION file and other common locations for R metadata. The package also includes utilities and examples for parsing and working with existing codemeta files, and includes several vignettes which illustrate both the basic usage of the package as well as some more advanced applications.

References

"CodeMeta Project Crosswalk." 2017. https://codemeta.github.io/. https://codemeta.github.io/.

"JSON-Ld 1.1." 2017. https://json-ld.org/spec/latest/json-ld/.

"Software Discovery Index Meeting Report." n.d. http://www.softwarediscoveryindex.org/. http://www.softwarediscoveryindex.org/.