

Software Citations and Versioning

Force11 Software Citation Implementation WG

Martin Fenner, DataCite

Software is Versioned

- *Every single commit (with commit hash)*
- *Via version tag, optionally using semantic versioning (major.minor.patch)*
- *Via persistent identifier, e.g. DOI, corresponding to a specific commit hash*

Software Gets Cited

- *In the scholarly literature and other places*
- *Informally by name*
- *Using a URL pointing to a code repository*
- *Using a persistent identifier, e.g. DOI*
- *Using the commit hash*

Force11 Software Citation Principles

2. Credit and attribution: *Software citations should facilitate giving scholarly credit and normative, legal attribution to all contributors to the software, recognizing that a single style or mechanism of attribution may not be applicable to all software.*

6. Specificity: *Software citations should facilitate identification of, and access to, the specific version of software that was used. Software identification should be as specific as necessary, such as using version numbers, revision numbers, or variants such as platforms.*

Which Version to Cite

- For **specificity**: as granular as possible, ideally commit hash or specific version
- For **credit and attribution**: as broad as possible, e.g. using a “concept identifier” that describes any version of the software

How do we bring these very diverging approaches to software citation together?

How to Track all Citations for a specific Software?

- *We need more than one identifier for the software*
- *We need to link all citations to specific versions, and then in turn the versions that have been cited, forming a graph*
- *We need tools and services that allow users to query this graph*

DataCite has built a GraphQL API to query this version/citation Graph

```
{
  software(id: "https://doi.org/10.5281/zenodo.2799488") {
    id
    titles {
      title
    }
    version
    versionOfCount
    citationCount
    versions {
      nodes {
        id
        titles {
          title
        }
        version
        versionCount
        versionOfCount
        citationCount
      }
    }
    citations {
      nodes {
        id
        titles {
          title
        }
      }
    }
  }
}
```

- One software
- One “concept identifier” and one version
- Three citations
- Plus some confusion

<https://api.datacite.org/graphql>