

Software Citation Implementation WG

Neil Chue Hong

Martin Fenner

Daniel S. Katz

11th RDA Plenary Meeting, Berlin

Looking at how software source code as an artefact is present in research object management workflows e.g. discovery, deposit, preservation

RDA Software Source
Code IG

Software Metadata is a common link, as are particular use case scenarios

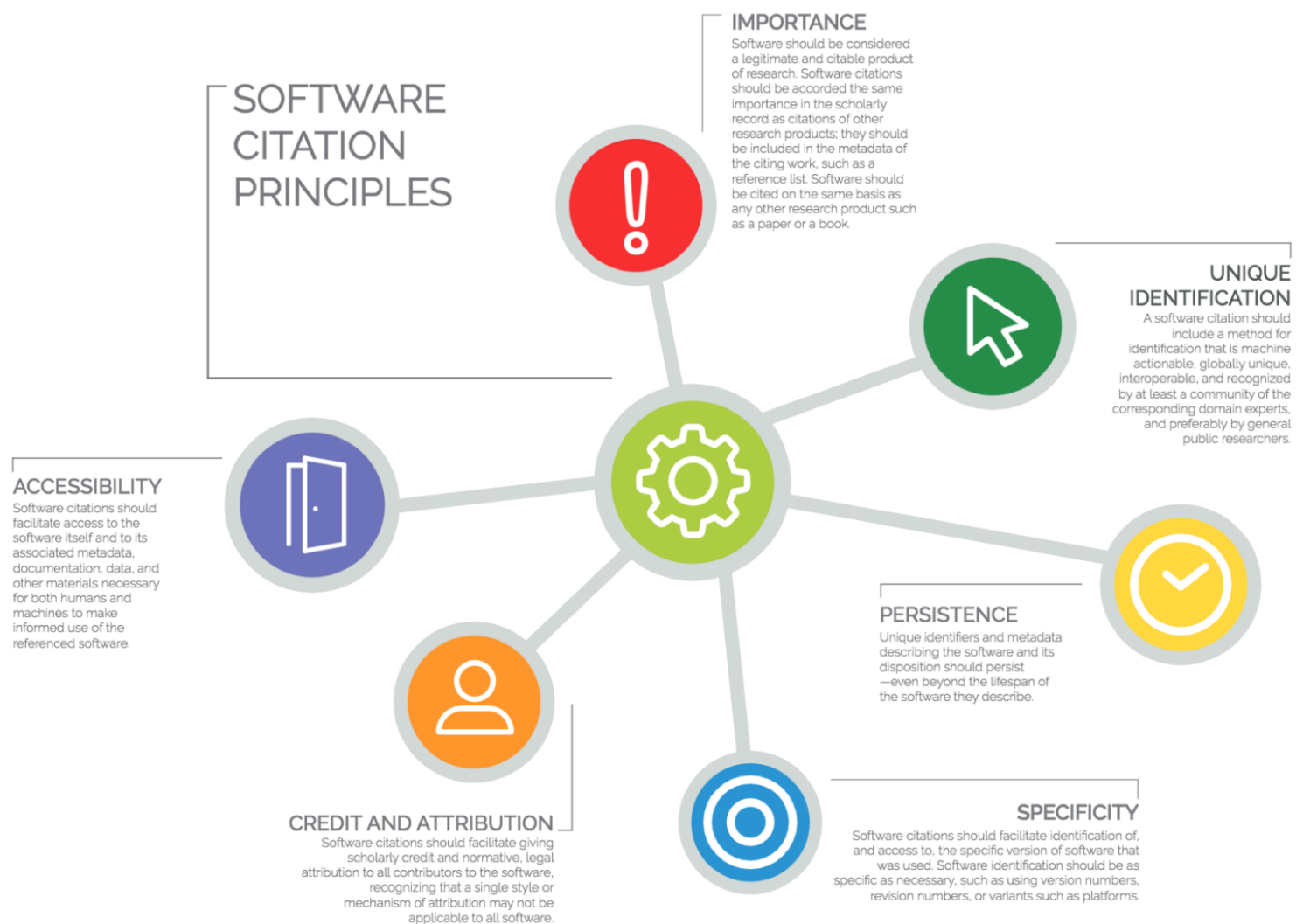
But metadata is somewhat different in the two parts, and some of the challenge is to see how much they overlap and can be joined in a possible superset

FORCE11
Software Citation
Implementation WG

Looking at how software (including source code) can be described and deposited to allow citation for different use cases including discovery, repeatability, reuse and credit

Software Citation Principles: People and Process

- FORCE11 Software Citation group started July 2015
- WSSSPE3 Credit & Citation working group joined September 2015
- ~55 members (researchers, developers, publishers, repositories, librarians)
- Work on GitHub <https://github.com/force11/force11-scwg> & FORCE11 <https://www.force11.org/group/software-citation-working-group>
- Reviewed existing community practices & developed use cases
- Drafted software citation principles document
 - Started with data citation principles, updated based on software use cases and related work, updated based on working group discussions, community feedback and review of draft, workshop at FORCE2016 in April
 - Katz DS, Niemeyer KE, et al (2016) Software vs. data in the context of citation. PeerJ Preprints 4:e2630v1. DOI: 10.7287/peerj.preprints.2630v1
 - Discussion via GitHub issues, changes tracked
- Submitted, reviewed and modified (many times), now published
 - Smith AM, Katz DS, Niemeyer KE, FORCE11 Software Citation Working Group.(2016) Software Citation Principles. PeerJ Computer Science 2:e86. DOI: 10.7717/peerj-cs.86 and <https://www.force11.org/software-citation-principles>



... yet there is little support for its acknowledgement and citation

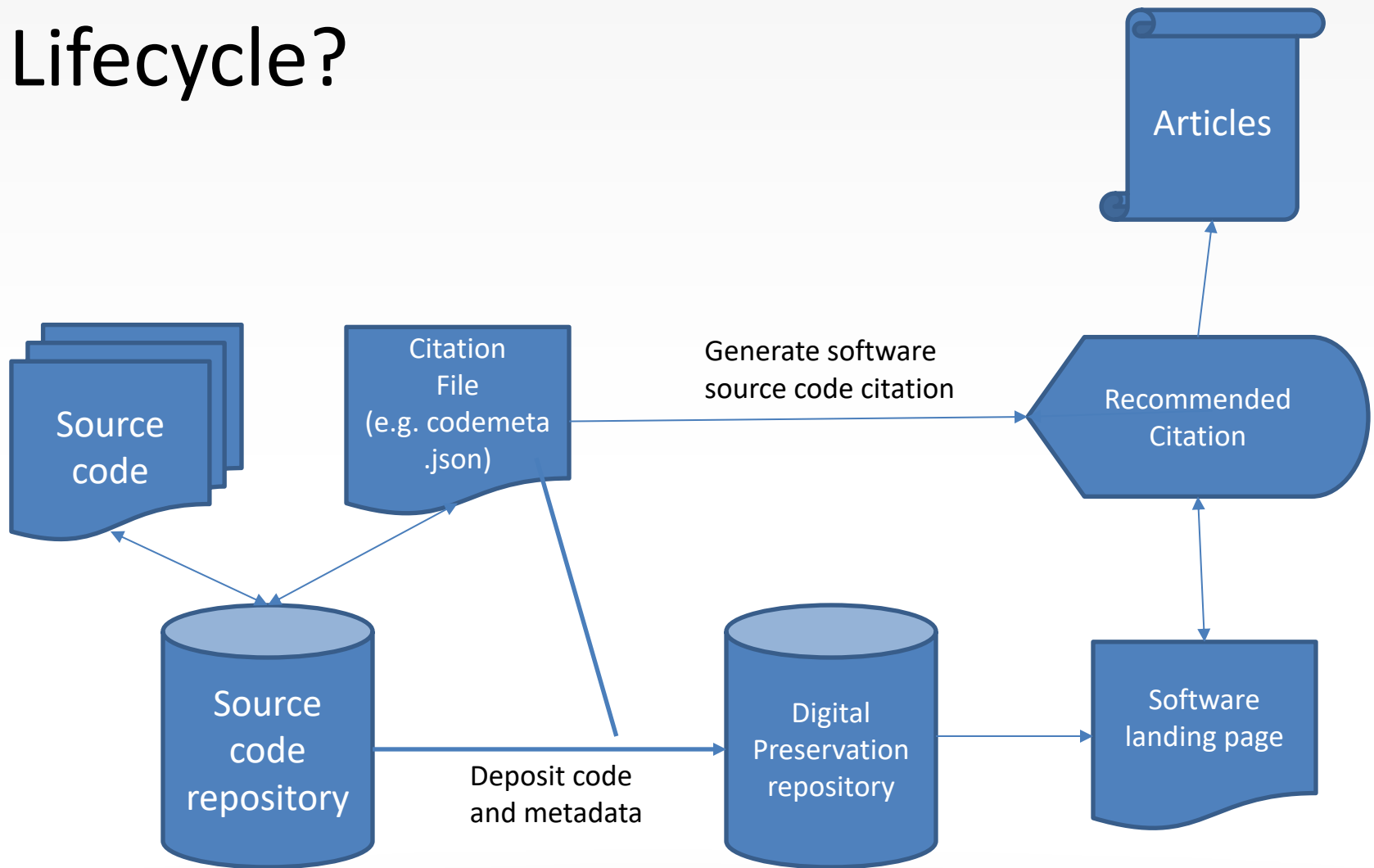
DataCite Schema Extensions

- DataCite Metadata Schema v4.1
 - <http://schema.datacite.org/meta/kernel-4.1/>
- Support for software citation
 - Facilitate software discovery, sharing and citation
 - Two new relation types
 - (HasVersion, IsVersionOf) – relate versions of the same software
e.g. relate an un-versioned code repository to one of its specific software versions
 - (IsRequiredBy, Requires) – software dependencies
 - Additions and modifications to documentation to clarify use for software citation
- See also:
 - <https://blog.datacite.org/metadata-schema-4-1/>

From Principle to Implementation

- FORCE11 Software Citation Implementation WG (Chairs: Katz, Fenner, Chue Hong)
 - <https://www.force11.org/group/software-citation-implementation-working-group>
 - <https://github.com/force11/force11-sciwg>
- Objectives
 - Endorse the principles
 - Develop sets of guidelines for implementing the principles
 - Help implement the principles
 - Test specific implementations of the principles
- Potential activities
 - Endorsement drive
 - Implementation hackathons / comparathons
 - Software Citation Support in Journal Article Tagging Suite (JATS)
 - Full implementation examples paper

Research Software Lifecycle?



Questions Arising – “social”

- Guidelines
 - How can we ensure consistent/interoperable citation across communities?
- Version mangling
 - How do we give guidance to production editors on e.g. on difference between version 2.1 and version 21 or version 2 1?
- Support in BiBTeX vs BiBLaTeX
 - The dreaded “misc” object

Questions Arising – “technical”

- Software landing pages
 - Metadata plus links
 - How best to implement?
- Citation duplication
 - If archives / crawlers are creating citations, will duplication make it harder to assign credit?
- Compact identifiers
 - What role might they play in making software referencing easier?

Conference Guidelines

- Work with specific conferences to get the proceedings authors to implement software citation
- Encourage the software developers in those fields to make their software citable
- Create guidance to create guidelines
- E.g. ACAT (High Energy Physics)
 - https://indico.cern.ch/event/567550/papers/2696191/files/6046-Katz_Citation.pdf

Generating citations automatically

- Using Codemeta.json to automatically generate citations
 - Code Repo metadata -> codemeta.json -> DataCite DOI
 - APIs to incorporate into existing infrastructure
- Citation File Format
 - “Human and machine readable citations”
 - <https://citation-file-format.github.io/>

CaltechDATA software preservation

- GitHub software repositories automatically preserved, with DOI
- Extended to use codemeta.json file to incorporate full author list, keywords and license metadata in CaltechDATA landing page and associated with DOI
 - <https://www.library.caltech.edu/news/enhanced-software-preservation-now-available-caltechdata>

DataCite support for CodeMeta

- Support for codemeta.json file in new DOI registration web service
 - DOI for repo, not for specific version
- Provide GitHub URL to service to register DOI with metadata automatically populated
 - If there's a codemeta.json file in root of repo
 - Will also give error if required metadata missing
- Next: how to integrate into DataCite client's workflows

Astronomy Pilot

- Working on brokering of software citations within publisher (AAS), indexer (ADS), and software repository (Zenodo)
 - <https://github.com/AASJournals/Tutorials/blob/master/Repositories/CitingRepositories.md>
- Trying to implement this with ADS in the next couple of months
- Model is generic but being tested with astronomy
 - Extension to geosciences (AGU)
 - <https://danielskatzblog.wordpress.com/2018/02/08/better-scientific-software-fellowship/>

xSDKTrilinos v. 1.0

Full Project

RESOURCE

Publicly Accessible Repository <https://github.com/trilinos/xSDKTrilinos>

DOI: 10.11578/dc.20171025.1744

SAVE / SHARE

Export Metadata 

Abstract

The main purpose of the xSDKTrilinos package is to provide an interoperability layer that enables easy installation and combined usage of the IDEAS libraries, including PETSc, Hypre, and SuperLU. The package can be used for any project that uses or will use different libraries, such as Trilinos, PETSc, HYPRE and SuperLU, together. One particular example is AMANZI code from Los Alamos National Laboratories, that uses PETSc and Trilinos at the same time.

Developers:

[Heroux, Michael](#) ^[1]; [Klinvex, Alicia](#) ^[1]; [Demeshko, Irina](#) ^[1] Show Author Affiliations

Release Date:

2016-04-14

Code Availability:

Open Source, Publicly Available Repository

Software Type

Scientific

Licenses:

BSD 3-clause "New" or "Revised" License

Citation Formats

MLA

APA

Chicago

Bibtex

Heroux, Michael A., Klinvex, Alicia M., and Demeshko, Irina. *xSDKTrilinos v. 1.0*. Computer Software. <https://github.com/trilinos/xSDKTrilinos>. USDOE. 14 Apr. 2016. Web. doi:10.11578/dc.20171025.1744.

October 12, 2017

Software

Open Access

Code repository that supports the research presented in the paper "Enabling adaptability in web forms based on user characteristics detection through A/B testing and Machine Learning".

Juan Cruz-Benito; Andrea Vázquez Ingelmo; José Carlos Sánchez-Prieto

Jupyter notebooks developed to support the research presented in the paper "Enabling adaptability in web forms based on user characteristics detection through A/B testing and Machine Learning" presented to the IEEE Access Journal in 2017 by Juan Cruz-Benito, Andrea Vázquez-Ingelmo, José Carlos Sánchez-Prieto, Francisco José García-Peñalvo, Roberto Therón and Martín Martín-González

Preview

paper-ieeeAccess-2017-1.0.zip

juan-cb-paper-ieeeAccess-2017-867b5c8

- README.md
- machinelearning-results.ipynb
- reinforcement-results.ipynb

Share



Cite as

Juan Cruz-Benito, Andrea Vázquez Ingelmo, & José Carlos Sánchez-Prieto. (2017, October 12). Code repository that supports the research presented in the paper "Enabling adaptability in web forms based on user characteristics detection through A/B testing and Machine Learning". (Version 1.0). Zenodo. <http://doi.org/10.5281/zenodo.1009618>

Start typing a citation style...

Available in

GitHub

Publication date:

October 12, 2017

DOI:

DOI 10.5281/zenodo.1009618

Keyword(s):

machine-learning scikit-learn pandas
jupyter-notebook jupyter python research paper
adaptability

Related identifiers:

Supplement to:

<https://github.com/juan-cb/paper-ieeeAccess-2017/tree/1.0>

Communities:

GRIAL Research Group - University of Salamanca

License (for files):

[Other \(Open\)](#)

Versions

Version 1.0 10.5281/zenodo.1009618 Oct 12, 2017

Cite all versions? You can cite all versions by using the DOI [10.5281/zenodo.1009617](https://doi.org/10.5281/zenodo.1009617). This DOI represents all versions, and will always resolve to the latest one. [Read more.](#)

Improving deposit workflows

- Jisc Research Data Shared Service Software deposit pilot
 - Preservation of software in institutional digital repositories
 - Software metadata -> SWORD deposit protocol
 - Code repository -> digital repository -> preservation archive
 - What information gets lost?

Other WIP / use links

- Hackathon at FORCE2017
 - Ideas: <https://github.com/force11/force11-sciwg/blob/master/hackathons/FORCE2017-hackathon.md>
 - Report: <https://www.software.ac.uk/blog/2017-11-09-hacking-future-software-citation>
- Implementations
 - Recommended citations: <http://citeas.org/>
 - CodeMeta software metadata exchange: <https://codemeta.github.io/>
 - Software Catalogs, Archives and Landing pages
 - <http://www.swmath.org/>
 - <https://zenodo.org/>
 - <https://www.nitrc.org/>
 - Extracting and archiving software metadata
 - Software Heritage: <https://www.softwareheritage.org/>
 - Microwler: <https://github.com/helgeho/Microwler>
 - Automated deposit of software and metadata into archives
 - Zenodo: <https://guides.github.com/activities/citable-code/>
 - Figshare: <https://support.figshare.com/support/solutions/articles/6000150264-how-to-connect-figshare-with-your-github-account>
 - SWORD: work starting by SSI as part of Jisc RDSS programme
 - Journal / Publisher Guidelines
 - AAS: <https://github.com/AASJournals/Tutorials/blob/master/Repositories/CitingRepositories.md>

Get Involved

- Join FORCE11 Software Citation Implementation WG
 - And join a task force
 - <https://www.force11.org/group/software-citation-implementation-working-group>
 - <https://github.com/force11/force11-sciwg>
- Contribute your experience / implementations
- Join us today
 - Discussions over lunch
 - Software Citation “Do-a-thon” hosted by FIZ Karlsruhe this afternoon: 13:30 – 17:00
 - Meet at Registration Desk at 13:15
 - Abteilung Mathematik und Informatik, Franklinstr. 11
 - <https://www.fiz-karlsruhe.de/en/fiz-karlsruhe/niederlassungen/buero-berlin.html>