

# Software Deposit: How to describe a software deposit

Michael Jackson (ed.), The Software Sustainability Institute

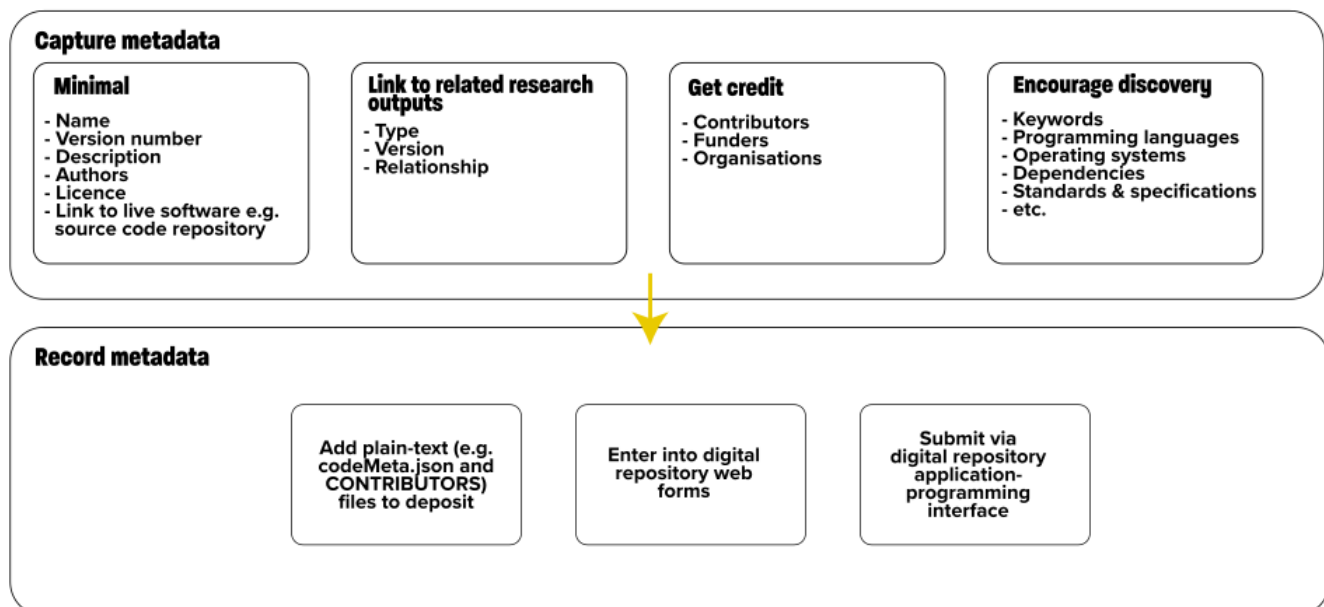
Version 1.0

doi:[10.5281/zenodo.1327321](https://doi.org/10.5281/zenodo.1327321)

07 August 2018

## Introduction

Papers have titles, authors, abstracts, subject categories and keywords, all of which are metadata, data which describes the paper. These metadata help others to categorise and index papers and find papers of interest to them. Metadata can serve a similar purpose for software deposits, providing a description of your software deposit. This metadata can help you get credit for your software, help digital repository managers categorise and index your software deposit and help other researchers find your software according to criteria of relevance to them. You will need this metadata when depositing your software into a digital repository. This guide summarises the metadata you can, and should, capture.



## How to describe a software deposit

## About this guide

This guide is one of a series of guides on software deposit, written by The Software Sustainability Institute<sup>[1](#)</sup>, funded by Jisc<sup>[2](#)</sup>. For an overview of the series, see Michael Jackson (ed.) (07 August 2018). Software Deposit: Guidance for Researchers (Version 1.0). Zenodo. doi:[10.5281/zenodo.1327310](https://doi.org/10.5281/zenodo.1327310). Online: <https://softwaresaved.github.io/software-deposit-guidance/SoftwareDepositGuidance.html>.

## Record a minimal set of metadata

You should always record the following metadata for your software deposit:

- Name: Name of your software.
- Version: Version number of your software.
- Description: Brief overview of your software, what it does (for example, what research problem it was written to solve) and what makes it novel (or different or distinct from similar software already available)
- Authors: Name, affiliation, current email address and ORCID [3](#) identifier. These provide a point of contact for others to get in touch if they have questions about your deposit and your research more generally, as well as allowing you and your fellow authors to get credit for your work. Make sure that any named people have given their consent for their names, affiliations, emails and ORCID identifiers to be included.
- Licence: name of your software licence. This allows others to know how they can, or cannot, use your software deposit.
- Link to live software: for ongoing projects links to your project web site and/or repository hosting service. This allows others to easily find your current software from your deposit and to engage with you and your project.

Though this represents a minimal set of metadata, you don't need to stop there. There is other metadata you can capture about your software deposit which can provide added richness to your description of your software deposit.

## Record metadata about related research outputs

Taken together, research software, data, facilities, experimental equipment, services and an overarching research question can be viewed as a research activity or experiment, worthy to be published. Conversely, a publication can be considered as a narrative that describes how the research objects are used together to address a research question.

Recording links to research objects related to your software (i.e. publications, data, other research software, facilities, experimental equipment, services) in your metadata allows others to see your software not in isolation but as an integral part of addressing a research question.

Consider not just research outputs directly produced by your or your project but those produced by others who have used your software and have cited it in their own papers or software, for example.

Provide the following metadata about each related research object:

- Type: paper, software, data, publication, etc.
- Version: this could be a persistent digital identifier (such as a DOI [4](#) or an ARK [5](#)), another form of unique identifier, a URL plus version number or timestamp.
- Relationship: nature and semantics of relationship with the other research object.

## Record metadata to help others discover your software

Both the minimal metadata and metadata about research outputs could help others discover your software if searching for software according to various criteria (e.g., if looking for software written by you or your co-authors, if looking for software with a specific licence, or if following links to related research outputs).

Recording the following metadata could help others discover your software according to richer criteria:

- Keywords: both generic and subject/domain/community-specific. This can include recording, as keywords, the names of algorithms implemented within your software, to allow for the discovery of more generic functionality which could be used elsewhere. For example, another researcher may be

interested in the implementation of a specific algorithm within the software, rather than the software itself.

- Any other metadata that describes your software, e.g., programming language(s), operating system(s), dependencies.

## Record metadata for credit

Recording yourself and your fellow authors as part of the minimal metadata can help you get credit for your research software, as this, coupled with your software's name and a unique persistent digital identifier (such as a DOI or an ARK) provides a minimal set of information which uniquely identifies you and your fellow authors, your software and where it can be obtained.

Additional metadata to further facilitate credit for those who contributed to your software includes:

- Contributors: Name, affiliation, current email address, ORCID identifier and an indication of how they contributed. You should include people who may not have designed or developed the software but provided input or contribution in other ways (for example, testers, documentation authors etc). Make sure that any named people have given their consent for their names, emails, affiliations and ORCID identifiers to be included.
- Funders: Name and ORCID identifier (if available).
- Organisations: Name and ORCID identifier (if available).

## How to record your metadata

When you submit a deposit to a digital repository you will be required to submit metadata. The metadata asked for will be specific to the digital repository. Digital repositories differ in both what metadata they can record and what they consider to be mandatory and optional metadata. At the least, the repository will collect a minimum set of metadata to allow it to generate a persistent digital identifier for your software deposit. Consult the documentation for your intended digital repository for information on its requirements.

Some digital repositories are limited in the metadata they record. They may not be able to record much of the metadata that is recommended above. Repository hosting services can also be limited in terms of the metadata they provide, which can limit the effectiveness of software deposit tools such as the figshare-GitHub integration [6](#) or Zenodo-GitHub integration [7](#).

To ensure none of your metadata is lost, it is recommended that you record all your metadata within plain-text files in your deposit. While the format of these plain-text files are up to you, we have two specific recommendations.

## Add a codeMeta.json file to your deposit

To ensure none of your metadata is lost, it is recommended that you record all your metadata within a codeMeta.json file, with your metadata expressed in the CodeMeta schema [8](#). CodeMeta is a structured format for software-related metadata and is both human- and machine-readable.

## Add a CONTRIBUTORS file to your deposit

While CodeMeta schema allows for authors and contributors to be specified it does not allow for the nature of their contribution to be specified (for example designer, developer, tester, documentation author etc). Provide a plain-text CONTRIBUTORS file which lists people who you deem to have made a significant contribution to your software (including yourself). This can include, for example, columns with each contributor's name, ORCID identifier and affiliation and brief summary of what their contribution was.

You should also state your funders, providing information on any grants that supported the software's development.

Make sure that any named people have given their consent for their names, ORCID identifiers and affiliations to be included.

For legacy software, you might want to include statement along the lines of "We have made all efforts to recall contributors. Please get in touch if you contributed to this software and we have forgotten you, so that this can be rectified in future deposits."

## Find out more

Related Software deposit guides:

- Michael Jackson (ed.) (07 August 2018). Software Deposit: What to deposit (Version 1.0). Zenodo. doi:[10.5281/zenodo.1327325](https://doi.org/10.5281/zenodo.1327325). Online: <https://softwaresaved.github.io/software-deposit-guidance/WhatToDeposit.html>.
- Michael Jackson (ed.) (07 August 2018). Software Deposit: How to choose a software licence (Version 1.0). Zenodo. doi:[10.5281/zenodo.1327316](https://doi.org/10.5281/zenodo.1327316). Online: <https://softwaresaved.github.io/software-deposit-guidance/HowToChooseSoftwareLicence.html>.

CodeMeta and codeMeta.json files:

- The CodeMeta Project, <https://codemeta.github.io/>. A machine-readable approach to documenting software metadata. Specific CodeMeta terms are listed at <https://codemeta.github.io/terms/>.
- schema.org, <http://schema.org/>. Metadata terms, many of which are used within CodeMeta.

## Cite this guide

Please cite as: Michael Jackson (ed.) (07 August 2018). Software Deposit: How to describe a software deposit (Version 1.0). Zenodo. doi:[10.5281/zenodo.1327321](https://doi.org/10.5281/zenodo.1327321). Online: <https://softwaresaved.github.io/software-deposit-guidance/HowToDescribeSoftwareDeposit.html>.



This work is published under a Creative Commons Attribution 4.0 International License (CC BY 4.0), <https://creativecommons.org/licenses/by/4.0/>.

- 
1. The Software Sustainability Institute, <https://www.software.ac.uk>.↵
  2. Jisc, <https://www.jisc.ac.uk>.↵
  3. ORCID, <https://orcid.org/>.↵
  4. Digital Object Identifier (DOI), <https://www.doi.org/>.↵
  5. Archival Resource Key (ARK). J. Kunze and R. Rogers (2008) The ARK Identifier Scheme, California Digital Library and US National Library of Medicine, May 2008. <https://escholarship.org/uc/item/9p9863nc>.↵
  6. How to connect figshare with your GitHub account, figshare knowledge, <https://knowledge.figshare.com/articles/item/how-to-connect-figshare-with-your-github-account-1>.↵
  7. Making your code citable, GitHub Guides, <https://guides.github.com/activities/citable-code/>.↵
  8. The CodeMeta Project, <https://codemeta.github.io/>.↵