Software in bibliographies with biblatex-software

Roberto Di Cosmo
Director, Software Heritage

May 5th, 2020
Outline

1. Introduction
2. biblatex-software in a nutshell
3. Use and contribute
Inria (https://www.inria.fr)

- French national research institution for digital sciences (CS, applied Maths, …)
- 200 teams federating some 3500 researchers
- 1000+ open source software projects over 50 years

The software citation working group

Contributors  Pierre Alliez, Roberto Di Cosmo, Benjamin Guedj, Alain Girault, Mohand-Said Hacid, Arnaud Legrand, Xavier Leroy, Nicolas Rougier, Manuel Serrano

Outputs

- biblatex-software: software bibitems for BibLaTeX

biblatex-software
BibTeX/BibLaTeX reminder

BibTeX

Bibliographic format widely used in the scholarly world by LaTeX users to store and exchange bibliographic information.
At Inria it is used to generate bibliographies for the activity reports of 200 teams

BibLaTeX

Modern LaTeX package (and tool set) that produces rich bibliographies from BibTeX entries

State of software:

No support for software in BibTeX/BibLaTeX up to now:
- BibLaTeX has a `@software` bibitem, that is treated like `@misc`
1 Introduction

2 biblatex-software in a nutshell

3 Use and contribute
Four new bibliographic entries

@software Computer software.
  Required fields: author / editor, title, url, year

@softwareversion A specific version of a software.
  Required fields: author / editor, title, url, version, year

@softwaremodule A specific module of a larger software project.
  Required fields: author, subtitle, url, year

@codefragment A code fragment (e.g. a specific algorithm in a program or library).
  Required fields: url

Inheritance

softwareversion, softwaremodule and codefragment entries can inherit missing fields from entries designated via the crossref field.
### Four new software-specific fields

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>license</strong></td>
<td>list (literal). The license/s of the title in SPDX format.</td>
</tr>
<tr>
<td><strong>introducedin</strong></td>
<td>field (literal). If this is a software module or fragment, the version of the containing project where it has been first introduced.</td>
</tr>
<tr>
<td><strong>repository</strong></td>
<td>field (uri). The url of the code repository (e.g on GitHub, GitLab).</td>
</tr>
<tr>
<td><strong>swhid</strong></td>
<td>field (verbatim). The identifier of the digital object (a.k.a the software artifact itself). The intrinsic identifier of the item is an swh-id (swh:cnt for a content, swh:dir for a directory, swh:rev for a revision, swh:rel for a release, etc.). See <a href="https">the SWH-ID specification</a>.</td>
</tr>
</tbody>
</table>

### Two more new software-specific fields for HAL

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>hal_id</strong></td>
<td>field (verbatim). A digital identifier for the software record including its description and metadata on HAL.</td>
</tr>
<tr>
<td><strong>hal_version</strong></td>
<td>field (verbatim). The version of the HAL software record.</td>
</tr>
</tbody>
</table>
Software, Softwareversion and Softwaremodule

@software {cgal, 
title = {The Computational Geometry Algorithms Library}, 
author = {{The CGAL Project}}, 
editor = {{CGAL Editorial Board}}, 
year = 1996, 
url = {https://cgal.org/} 
}

@softwareversion{cgal:5-0-2, 
crossref = {cgal}, 
version = {{5.0.2}}, 
url = {https://docs.cgal.org/5.02}, 
year = 2020, 
swvid = {swh:1:rel:636541b6bfc7863908eae744610c3d91fa58855; origin=https://github.com/CGAL/cgal/} 
}

@softwaremodule{cgal:lp-gi-20a, 
crossref = {cgal:5-0-2}, 
author = {Menelaos Karavelas}, 
subtitle = {{2D} Voronoi Diagram Adaptor}, 
license = {GPL}, 
introducedin = {cgal:3-1}, 
url = {https://doc.cgal.org/5.0.2/Manual/packages.html#PkgVoronoiDiagram2}, 
}


Texts in blue are clickable links, try them!
Examples are worth a thousand words, cont’d

Software, Softwareversion and Codefragment

```latex
@software {parmap,
  title = {The Parmap library},
  author = {Di Cosmo, Roberto and Marco Danelutto},
  year = {2012},
  institution = {{University Paris Diderot} and {University of Pisa}},
  url = {https://rdicosmo.github.io/parmap/},
  license = {LGPL-2.0},
}

@softwareversion {parmap-0.9.8,
  version = {0.9.8},
  swhid = {swh:1:rev:0064fbd0ad69de205ea6ec6999f3d3895e9442c2;origin=https://gitorious.org/parmap/parmap.git},
  crossref = {parmap}
}

@codefragment {simplemapper,
  subtitle = {Core mapping routine},
  swhid = {swh:1:cnt:d5214ff9562a1fe78db51944506ba48c20de3379;origin=https://gitorious.org/parmap/parmap.git;anchor=swh:1:rev:0064fbd0ad69de205ea6ec6999f3d3895e9442c2;path=/parmap.ml;lines=101-143},
  crossref = {parmap-0.9.8}
}
```


Texts in blue are clickable links, try them!
Five new options to control typesetting of software related entries

swlabels=true|false  add a special label to software entries
license=true|false  show license information
halid=true|false  show the HAL identifier
swhid=true|false  show the SWHID identifier
vcs=true|false  show the repository URL

Adding other identifiers

- doi, hal_id, shwid, vcs, url are built-in
- other identifiers are easy to add using \texttt{eprint} (see full documentation)
Outline

1. Introduction
2. biblatex-software in a nutshell
3. Use and contribute
A style extension compatible with all existing BibLaTeX style

Three simple steps

1) copy the relevant files into your article’s directory

software.bbx software.dbx software-biblatex.sty
english-software.lbx

2) pass the datamodel=software option when loading biblatex

\usepackage[datamodel=software]{biblatex}

3) load software-biblatex

\usepackage{software-biblatex}

That’s it! (see the documentation for full details)
Availability and contributions

Distribution and documentation

biblatex-software is available on CTAN as
https://www.ctan.org/tex-archive/macros/latex/contrib/biblatex-contrib/biblatex-software
- documentation in software-biblatex.pdf
- demo example in sample-use-sty.pdf
- sample bibliography in biblio.bib

Development

contributions are welcome on
https://gitlab.inria.fr/gt-sw-citation/bibtex-sw-entry
BibLaTeX users can finally properly handle software in their bibliographies. *Let’s spread the word!*

### Articles using biblatex-software and Software Heritage

**R. Di Cosmo.** “Archiving and referencing source code with Software Heritage”. In: ICMS. Lecture Notes in Computer Science. to appear, preprint available. 2020. [URL](https://hal.archives-ouvertes.fr/hal-02526083).