Towards a Software Pillar for Open Science

from policy to implementation

Roberto Di Cosmo

Chair, Software Chapter, National Committee for Open Science Director, Software Heritage Inria and Université de Paris Cité

September 9th 2022



> Software Heritage

THE GREAT LIBRARY OF SOURCE CODE

Outline



Short Bio: Roberto Di Cosmo

Computer Science professor in Paris, now working at INRIA

- 30+ years of research (Theor. CS, Programming, Software Engineering, Erdos #: 3)
- 20+ years of Free and Open Source Software
- 10+ years building and directing structures for the common good



DemoLinux - first live GNU/Linux distro 1999 2007 Free Software Thematic Group 150 members 40 projects 200Me 2008 Mancoosi project www.mancoosi.org 2010 IRILL www.irill.org 2015 Software Heritage at INRIA 2018 National Committee for Open Science, France 2021 EOSC Task Force on Infrastructures for Software. **European Union**

Outline



R. Di Cosmo roberto@dicosmo.org

Open Science (Second National Plan for Open Science, France, 2021)

(CC-BY 4.0)

Unhindered dissemination of results, methods and products from scientific research. It draws on *the opportunity provided by recent digital progress* to develop *open access* to *publications* and – as much as possible – *data, source code* and *research methods*.

Towards a Software Pillar of Open Science

September 9th 2022

Open Science (Second National Plan for Open Science, France, 2021)

Unhindered dissemination of results, methods and products from scientific research. It draws on *the opportunity provided by recent digital progress* to develop *open access* to *publications* and – as much as possible – *data, source code* and *research methods*.

Jean-Eric Paquet (EU DGRI, on the objective of Open Science)

"Increase scientific quality, the pace of discovery and technological development, as well as societal trust in science."



Open Science (Second National Plan for Open Science, France, 2021)

Unhindered dissemination of results, methods and products from scientific research. It draws on *the opportunity provided by recent digital progress* to develop *open access* to *publications* and – as much as possible – *data, source code* and *research methods*.

Jean-Eric Paquet (EU DGRI, on the objective of Open Science)

"Increase scientific quality, the pace of discovery and technological development, as well as societal trust in science."

Mariya Gabriel (EU Commissionneer for Research)

The COVID-19 crisis has also shown that cooperation at international level in research and innovation is more important than ever, including through *open access to data and results*. *No nation, no country can tackle any of these global challenges alone.*

Open Science (Second National Plan for Open Science, France, 2021)

Unhindered dissemination of results, methods and products from scientific research. It draws on *the opportunity provided by recent digital progress* to develop *open access* to *publications* and – as much as possible – *data, source code* and *research methods*.

Jean-Eric Paquet (EU DGRI, on the objective of Open Science)

"Increase scientific quality, the pace of discovery and technological development, as well as societal trust in science."

Mariya Gabriel (EU Commissionneer for Research)

The COVID-19 crisis has also shown that cooperation at international level in research and innovation is more important than ever, including through *open access to data and results. No nation, no country can tackle any of these global challenges alone.*

Yuval Noah Harari (on COVID 19)

"The real antidote [to epidemic] is scientific knowledge and global cooperation."

French National plan for Open Science, 2018-2021



Launch on 4 July 2021 by Frédérique Vidal, Minister for Higher Education, Research and Innovation

- · First Commitment : Generalise open access to publications
- Second Commitment : Structure research data and make it available through open access
- Third Commitment : Be part of a sustainable european and international open science dynamic



Governance : The Open Science Committee





R. Di Cosmo roberto@dicosmo.org (CC-BY 4.0)

French National plan for Open Science, 2018-2021



Launch on 4 July 2021 by Frédérique Vidal, Minister for Higher Education, Research and Innovation

- · First Commitment : Generalise open access to publications
- Second Commitment : Structure research data and make it available through open access
- Third Commitment : Be part of a sustainable european and international open science dynamic



Main achievements

- Creation of the National Fund for Open Science (11.6 M€ of resources over 3 years), a
 dedicated funding instrument for open science policy in France
- · Two calls for projects for "open access publishing, open publication and its ecosystem"
- Support for the national open archive HAL: exceptional financial support, audit and ongoing renovation of the technical base, definition of a sustainable financing model and shared governance
- The French National Research Agency and other funding agencies request the deposit of publications in an open archive and the drafting of data management plans for the projects they fund
- ANR flash call for projects on open science: €2.3M, 25 projects supported to accelerate the
 maturation of disciplinary communities in the face of data management issues
- Creation of the position of chief data officer at the Ministry of Higher Education, Research and Innovation



Governance : The Open Science Committee





Main achievements

- Creation of the open science barometer: measurement of the rate of French scientific publications in open access
- Progressive deployment of open science strategies within research performing organisations and universities, creation of a network of open science referents
- Training actions targeted at doctoral students: Passport to open science, practical guide for Phd students
- Contribution to the structuring and governance of the EOSC: structuring of EOSC France, French presence on the board and other EOSC governance bodies
- Support to international open science infrastructures: SCoSS labelled projects (DOAB, PKP, OpenCitations), RDA, Software Heritage

R. Di Cosmo roberto@dicosmo.org (CC-BY 4.0)

.

Outline



6 Focus on ARDC and infrastructures

🕖 Demo time!

8 Actions

Software powers modern research



[...] software [...] essential in their fields.

Top 100 papers (Nature, 2014)

Sometimes, if you dont have the software, you dont have the data

Christine Borgman, Paris, 2018





[...] software [...] essential in their fields.

Top 100 papers (Nature, 2014)

Sometimes, if you dont have the software, you dont have the data Christine Borgman, Paris, 2018





[...] software [...] essential in their fields.

Top 100 papers (Nature, 2014)

Sometimes, if you dont have the software, you dont have the data Christine Borgman, Paris, 2018



The links in the picture are important





access to the source code is essential!



Nota Bene

software may be a tool, a research outcome and a research object

access to the *source code* is essential!

Preserving (the history of) source code is necessary for reproducibility



Harold Abelson, Structure and Interpretation of Computer Programs (1st ed.)

"Programs must be written for people to read, and only incidentally for machines to execute."

Apollo 11 source code (excerpt)

P63SP0T3	CA EXTEND RAND	BIT6 CHAN33	#	IS THE LR ANTE	ENNA IN POSITION 1 YET
	EXTEND BZF	P63SP0T4	#	BRANCH IF ANTE	ENNA ALREADY IN POSITION 1
	CAF	CODE500	#	ASTRONAUT:	PLEASE CRANK THE
	TC	BANKCALL	#		SILLY THING AROUND
	CADR	GOPERF1			
	TCF	GOTOPOOH	#	TERMINATE	
	TCF	P63SP0T3	#	PROCEED	SEE IF HE'S LYING
P63SP0T4	TC CADR	BANKCALL SETPOS1	#	ENTER	INITIALIZE LANDING RADAR
	TC CADR	POSTJUMP BURNBABY	#	OFF TO SEE THE	E WIZARD

1985

Harold Abelson, Structure and Interpretation of Computer Programs (1st ed.)

"Programs must be written for people to read, and only incidentally for machines to execute."

Apollo 11 source code (excerpt)

P63SP0T3	CA EXTEND RAND	BIT6 CHAN33	#	IS THE LR ANTE	ENNA IN POSITION 1 YET
	BZF	P63SP0T4	#	BRANCH IF ANTE	ENNA ALREADY IN POSITION 1
	CAF	CODE500	#	ASTRONAUT:	PLEASE CRANK THE
	TC	BANKCALL	#		SILLY THING AROUND
	CADR	GOPERF1			
	TCF	GOTOPOOH	#	TERMINATE	
	TCF	P63SP0T3	#	PROCEED	SEE IF HE'S LYING
P63SP0T4	тс	BANKCALL	#	ENTER	INITIALIZE LANDING RADAR
	CADR	SETP0S1			
	TC CADR	POSTJUMP BURNBABY	#	OFF TO SEE THE	E WIZARD

Quake III source code (excerpt)

```
float 0_rsqrt( float number )
{
    long i;
    float x2, y;
    const float threehalfs = 1.5F;
```

```
x2 = number * 0.5F;
y = number;
i = * ( long * ) &y; // evil floating point bit level hacking
i = 0x5f3759df - ( i >> 1); // what the fuck?
y = % ( float * ) &i;
y = y * ( threehalfs - ( x2 * y * y ) ); // Ist iteration
// y = y * ( threehalfs - ( x2 * y * y ) ); // Ist iteration, this
can be removed
```

1985

return y;

Harold Abelson, Structure and Interpretation of Computer Programs (1st ed.)

"Programs must be written for people to read, and only incidentally for machines to execute."

Apollo 11 source code (excerpt)

P63SP0T3 CA BIT6 # IS THE LR ANTENNA IN POSITION 1 YET RAND CHAN33 EXTEND EXTEND CAF CODE500 # BRANCH IF ANTENNA ALREADY IN POSITION 1 CAF CODE500 # ASTRONAUT: PLEASE CRANK THE TC BANKCALL # SILLY THING AROUND CADR GOTOPOOH # TERMINATE TCF 603P0704 # TERMINATE TCF 803SP073 # PROCEED SEE IF HE'S LYING P83SP074 CAB BANKCALL # ENTER INITIALIZE LANDING RADAR				
PASSPOTA CAANA3 EZF PGSSPOTA # BRANCH IF ANTENNA ALREADY IN POSITION 1 CAF CODE500 # ASTRONAUT: PLEASE CRANK THE TC BANKCALL # SILLY THING AROUND CADR GOTOPODH # TERMINATE TCF 0030P073 # PROCEED SEE IF HE'S LYING P83SP074 CAB SANKCALL # ENTER INITIALIZE LANDING RADAR	P63SP0T3	CA EXTEND	BIT6	# IS THE LR ANTENNA IN POSITION 1 YET
EXTEND EZF P03SP074 # BRANCH IF ANTENNA ALREADY IN POSITION 1 CAF CODES00 # ASTRONAUT: PLASE CRANK THE TC BANKCALL # SILLY THING AROUND CADR GOPERF1 TCF 0GTOPOOH # TERMINATE TCF P03SP074 # PROCEED SEE IF HE'S LYING P03SP074 TC BANKCALL # ENTER INITIALIZE LANDING RADAR CADR SETPOS1		RAND	CHAN33	
BZF P63SP014 # BRANCH IF ANTENNA ALREADY IN POSITION 1 CAF CODE500 # ASTRONAUT: PLEASE CRANK THE TC BANKCALL # SILLY THING AROUND CAPR GOTOPOOH # TERMINATE TCF P03SP014 # PROCEED SEE IF HE'S LYING P63SP074 TC BANKCALL # ENTER CAPR SEE IF HE'S LYING		EXTEND		
CAF CODE500 # ASTRONAUT: PLEASE CRANK THE TC BANKCALL # SILLY THING AROUND CADR GOTOPOOH # TERMINATE TCF GOTOPOOH # TERMINATE TCF PGOSPOT3 # PROCEED P63SP0T4 TC BANKCALL # ENTER CADR SETFOS1 *		BZF	P63SP0T4	# BRANCH IF ANTENNA ALREADY IN POSITION 1
TC BANKCALL # SILLY THING AROUND CADR GOPERFI FORCE TCF FOGSPOOH # TERMINATE TCF POGSPOT3 # PROCEED PG3SPOT4 TC BANKCALL # ENTER PG3SPOT4 TC BANKCALL # ENTER		CAF	CODE500	# ASTRONAUT: PLEASE CRANK THE
CADR GOPERF1 TCF GOTOPOOH # TERMINATE TCF P63SP0T3 # PROCEED SEE IF HE'S LYING P63SP0T4 TC BANKCALL # ENTER INITIALIZE LANDING RADAR CADR SETPOS1		TC	BANKCALL	# SILLY THING AROUND
TCF GOTOPOON # TERMINATE TCF P03SP0T3 # PROCEED SEE IF HE'S LYING P03SP0T4 TC BANKCALL # ENTER INITIALIZE LANDING RADAR CADR SETPOS1		CADR	GOPERF1	
TCF P63SPOT3 # PROCEED SEE IF HE'S LYING P63SPOT4 TC BANKCALL # ENTER INITIALIZE LANDING RADAR CADR SETPOS1		TCF	GOTOPOOH	# TERMINATE
P63SP0T4 TC BANKCALL # ENTER INITIALIZE LANDING RADAR CADR SETPOS1		TCF	P63SP0T3	# PROCEED SEE IF HE'S LYING
CADR SETPOS1	P63SP0T4	тс	BANKCALL	# ENTER INITIALIZE LANDING RADAR
		CADR	SETP0S1	
TC POSTJUMP # OFF TO SEE THE WIZARD		TC	POSTJUMP	# OFF TO SEE THE WIZARD
CADR BURNBABY		CADR	BURNBABY	

Quake III source code (excerpt)

```
float 0_rsqrt( float number )
{
    long i;
    float x2, y;
    const float threehalfs = 1.5F;
```

```
x2 = number * 0.5F;
y = number;
i = * ( long * ) &y; // evil floating point bit level hacking
i = 0x5f3759df - ( i >> 1); // what the fuck?
y = % ( float * ) &i;
y = y * ( threehalfs - ( x2 * y * y ) ); // Ist iteration
// y = y * ( threehalfs - ( x2 * y * y ) ); // Ist iteration, this
can be removed
```

return y;

Len Shustek, Computer History Museum

2006

1985

"Source code provides a view into the mind of the designer."

R. Di Cosmo roberto@dicosmo.org (CC-BY 4.0)

Outline

- Software as a pillar for Open Science
- An emerging policy framework
- Towards implementation: assessing the needs for a software pillar
- 6 Focus on ARDC and infrastructures
- 🕖 Demo time!
- 8 Actions

International highlights

Paris Call on Software Source code (2019, UNESCO)



40 international experts call to "promote software development as a valuable research activity, and research software as a key enabler for Open Science/Open Research, [...] recognising in the careers of academics their contributions to high quality software development, in all their forms"

International highlights

Paris Call on Software Source code (2019, UNESCO)



40 international experts call to "promote software development as a valuable research activity, and research software as a key enabler for Open Science/Open Research, [...] recognising in the careers of academics their contributions to high quality software development, in all their forms"

UNESCO recommendations for Open Science, 2018-2021

"The source code must be included in the software release and [...] the license must allow modifications, derivative works and sharing [...]" "Open science infrastructures should be [...] essentially not-for-profit and long-term"

International highlights

Paris Call on Software Source code (2019, UNESCO)



40 international experts call to "promote software development as a valuable research activity, and research software as a key enabler for Open Science/Open Research, [...] recognising in the careers of academics their contributions to high quality software development, in all their forms"

UNESCO recommendations for Open Science, 2018-2021

"The source code must be included in the software release and [...] the license must allow modifications, derivative works and sharing [...]" "Open science infrastructures should be [...] essentially not-for-profit and long-term"

EOSC SIRS report: Software Source Code and Open Science, 2020



- connect scholarly ecosystem via Software Heritage
- use open non profit infrastructures
- open source first: "all research software should be made available under an Open Source license by default"

French National plan for Open Science, 2021-2024

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR, DE LA RECHERCHE ET DE L'INNOVATION John Autorit Applitie Fauradit



SECOND FRENCH PLAN FOR OPEN SCIENCE

Generalising open science in France 2021-2024



CONTRACTOR

Second French Plan for Open Science



Launch on 6 July 2021 by Frédérique Vidal, Minister for Higher Education, Research and Innovation

- Multiplying the levers for change in order to generalise open science practices
- Structuring the policy for opening up or sharing research data
- New commitments to the opening of source code produced by research
- European and international inclusion in the context of the French Presidency of the European Union

,

Disciplinary and thematic variations: open science policies must be adapted to disciplinary specificities



French National plan for Open Science, 2021-2024

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR, DE LA RECHERCHE ET DE L'INNOVATION Livret Egolini Fauraid



SECOND FRENCH PLAN FOR OPEN SCIENCE

Generalising open science in France 2021-2024





Second French Plan for Open Science



Launch on 6 July 2021 by Frédérique Vidal, Minister for Higher Education, Research and Innovation

- Multiplying the levers for change in order to generalise open science practices
- Structuring the policy for opening up or sharing research data
- New commitments to the opening of source code produced by research
- European and international inclusion in the context of the French Presidency of the European Union
- Disciplinary and thematic variations: open science policies must be adapted to disciplinary specificities





Define and promote an open source software policy

- Produce a National Charter for Open Source Software coming from higher education, research and innovation
- Develop the link between data and software through a network of Chief Data Officers in the various universities and research performing organisations.
- Develop the economic models of open source software and make them known within commercialization services
- * Support Software Heritage and recommend it for the archiving and referencing of source code

Recognise source code as a contribution to research

- Create an open source research software prize
- Provide greater recognition for software production in the career of researchers, research support staff

Build an ecosystem that connects code, data and publications

 Develop proper coordination between software forges, open publication archives, data repositories and the scientific publishing sector.

R. Di Cosmo roberto@dicosmo.org (CC-BY 4.0)

Software Chapter in the CoSO

Five action lines

- Identifying and highlighting research software production
- Technical and social tools and best practices
- Valorization and sustainability
- Liaison and animation at national, European, and international levels
- Recognition and careers



Software Chapter in the CoSO

Five action lines

- Identifying and highlighting research software production
- Technical and social tools and best practices
- Valorization and sustainability
- Liaison and animation at national, European, and international levels
- Recognition and careers

Leveraging experience and connections

- Open Source thematic group in Systematic (since 2007, more on demand)
- Collaboration with DINUM, Eclipse Foundation, OW2, ...

Software Chapter in the CoSO

Five action lines

- Identifying and highlighting research software production
- Technical and social tools and best practices
- Valorization and sustainability
- Liaison and animation at national, European, and international levels
- Recognition and careers

Leveraging experience and connections

- Open Source thematic group in Systematic (since 2007, more on demand)
- Collaboration with DINUM, Eclipse Foundation, OW2, ...

The Open Science award for Open Source research software

See the official page at MESRI

Outline

- Software as a pillar for Open Science
- An emerging policy framework
- 5 Towards implementation: assessing the needs for a software pillar
- 6 Focus on ARDC and infrastructures
- 🕖 Demo time!
- 8 Actions

What is at stake

ARDC

- Archive for retrieval (*reproducibility*)
- Reference for identification (*reproducibility*)
- Describe for discovery and reuse
- Cite/Credit for credit and evaluation



What is at stake

ARDC

- Archive for retrieval (*reproducibility*)
- Reference for identification (reproducibility)
- Describe for discovery and reuse
- Cite/Credit for credit and evaluation

Before ARDC

- Development practices and tools (VCS, build system, test suites, CI, ...)
- Opening up towards a community (documentation, organization, communication)

Need training, best practices



What is at stake

ARDC

- Archive for retrieval (*reproducibility*)
- Reference for identification (*reproducibility*)
- Describe for discovery and reuse
- Cite/Credit for credit and evaluation

Before ARDC

- Development practices and tools (VCS, build system, test suites, CI, ...)
- Opening up towards a community (documentation, organization, communication)

Need training, best practices

Beyond ARDC

- Policies (dissemination, reuse, careers!)
- Sustainability (legal, economic etc.)
- Technology transfer
- Advanced technologies and tools (quality, traceability, etc.)

Outline

- Software as a pillar for Open Science
- An emerging policy framework
- Towards implementation: assessing the needs for a software pillar
- 6 Focus on ARDC and infrastructures
- 🕖 Demo time!
- 8 Actions

2015: the first big bad news

Google Code and Gitorious.org shutdown: ~1M endangered repositories

• broken links in the web of knowledge (my papers too)



2015: the first big bad news

Google Code and Gitorious.org shutdown: ~1M endangered repositories

• broken links in the web of knowledge (my papers too)

Big bad news keep coming in

- summer 2019: BitBucket announces Mercurial VCS sunset
- july 2020: BitBucket erases 250.000+ repositories (including research software)
- summer 2022: GitLab.com considers erasing all projects that are inactive for a year

2015: the first big bad news

Google Code and Gitorious.org shutdown: ~1M endangered repositories

• broken links in the web of knowledge (my papers too)

Big bad news keep coming in

- summer 2019: BitBucket announces Mercurial VCS sunset
- july 2020: BitBucket erases 250.000+ repositories (including research software)
- summer 2022: GitLab.com considers erasing all projects that are inactive for a year

In Academia too!

• 2021: Inria's old gforge is unplugged... breaks the Opam build chain for OCaml

2015: the first big bad news

Google Code and Gitorious.org shutdown: ~1M endangered repositories

• broken links in the web of knowledge (my papers too)

Big bad news keep coming in

- summer 2019: BitBucket announces Mercurial VCS sunset
- july 2020: BitBucket erases 250.000+ repositories (including research software)
- summer 2022: GitLab.com considers erasing all projects that are inactive for a year

In Academia too!

• 2021: Inria's old gforge is unplugged... breaks the Opam build chain for OCaml

We need a universal archive of software source code:

2015: the first big bad news

Google Code and Gitorious.org shutdown: ~1M endangered repositories

• broken links in the web of knowledge (my papers too)

Big bad news keep coming in

- summer 2019: BitBucket announces Mercurial VCS sunset
- july 2020: BitBucket erases 250.000+ repositories (including research software)
- summer 2022: GitLab.com considers erasing all projects that are inactive for a year

In Academia too!

• 2021: Inria's old gforge is unplugged... breaks the Opam build chain for OCaml

We need a universal archive of software source code: now we have one!

www.softwareheritage.org



www.softwareheritage.org



www.softwareheritage.org



www.softwareheritage.org



R. Di Cosmo roberto@dicosmo.org (CC-BY 4.0)

Towards a Software Pillar of Open Science Septemb

September 9th 2022 11 / 17

The largest software archive, a shared infrastructure



The largest software archive, a shared infrastructure



R. Di Cosmo roberto@dicosmo.org (CC-BY 4.0)

Towards a Software Pillar of Open Science September 9th 2022 12 / 17

An international, non profit initiative

built for the long term



An international, non profit initiative

built for the long term









- save.softwareheritage.org
- deposit.softwareheritage.org

Archive (12B+ files, 170M+ projects)



- save.softwareheritage.org
- deposit.softwareheritage.org

Reference (20 billion SWHIDs)

Intrinsic, decentralised, cryptographically strong identifiers, SWHIDs



Now supported in SPDX 2.2, Wikidata etc.

Archive (12B+ files, 170M+ projects)



- save.softwareheritage.org
- deposit.softwareheritage.org

Describe

- Intrinsic metadata from source code
- Contributed the Codemeta generator

R. Di Cosmo roberto@dicosmo.org (CC-BY 4.0)

Reference (20 billion SWHIDs)

Intrinsic, decentralised, cryptographically strong identifiers, SWHIDs



Now supported in SPDX 2.2, Wikidata etc.

Archive (12B+ files, 170M+ projects)



- save.softwareheritage.org
- deposit.softwareheritage.org

Reference (20 billion SWHIDs)

Intrinsic, decentralised, cryptographically strong identifiers, SWHIDs



Now supported in SPDX 2.2, Wikidata etc.

Describe

- Intrinsic metadata from source code
- Contributed the Codemeta generator

Cite/Credit

• Contributed *software citation* style

biblatex-software, v 1.2-2 now on CTAN

Outline



A walkthrough

- Browse (e.g. Apollo 11, and your work may be already there !)
- Trigger archival, use the updateswh browser extension (GitHub action available too)
- Get and use SWHIDs (full specification available online)
- Cite software using the biblatex-software package from CTAN
- Example in a journal: an article from IPOL
- Example with Parmap: devel on Github, archive in SWH, curated deposit in HAL
- Extracting all the software products for Inria, for CNRS, for CNES, for LIRMM or for Rémi Gribonval using HalTools
- Curated deposit in SWH via HAL, see for example: LinBox, SLALOM, Givaro, NS2DDV, SumGra, Coq proof, ...
- Example use in research articles:
 - compare Fig. 1 and conclusions in the 2012 version and the updated version
 - SWHID in a replication experiment

Outline



Call to action: policy making

A working agenda

- avoid proprietarisation: set the default to open
 - publicly funded research software should be open source, exceptions must be justified
 - set up institutional support
 - build common knowledge base for technology transfer offices
- establish intelligent and effective incentives
 - count quality software contributions in careers, avoid purely numerical indicators, keep the human in the loop (mind Goodhart's law)
- avoid balkanisation, support mutualised common infrastructures
 - build on common, shared, open, non profit infrastructures, like Software Heritage
 - acknowledge the predominant human component of digital infrastructures
 - recurrent funding of their cost
 - proper evaluation of their service

Learn more



- UNESCO, *Draft recommendations on Open Science* 2021, (online)
- French Ministry of Research, *Second National Plan for Open Science* 2021, (online)
- EOSC SIRS Task Force, Scholarly Infrastructures for Research Software 2020, Publications office of the European Commission, (10.2777/28598)
- R. Di Cosmo, Archiving and Referencing Source Code with Software Heritage
 International Conference on Mathematical Software 2020 (10.1007/978-3-030-52200-1_36)
- J.F. Abramatic, R. Di Cosmo, S. Zacchiroli, *Building the Universal Archive of Source Code* CACM, October 2018 (10.1145/3183558)