

# General Outlook and Call for Action

Software for Reproducibility, Software Endangered, Software Ignored

Roberto Di Cosmo (Software Heritage, INRIA)

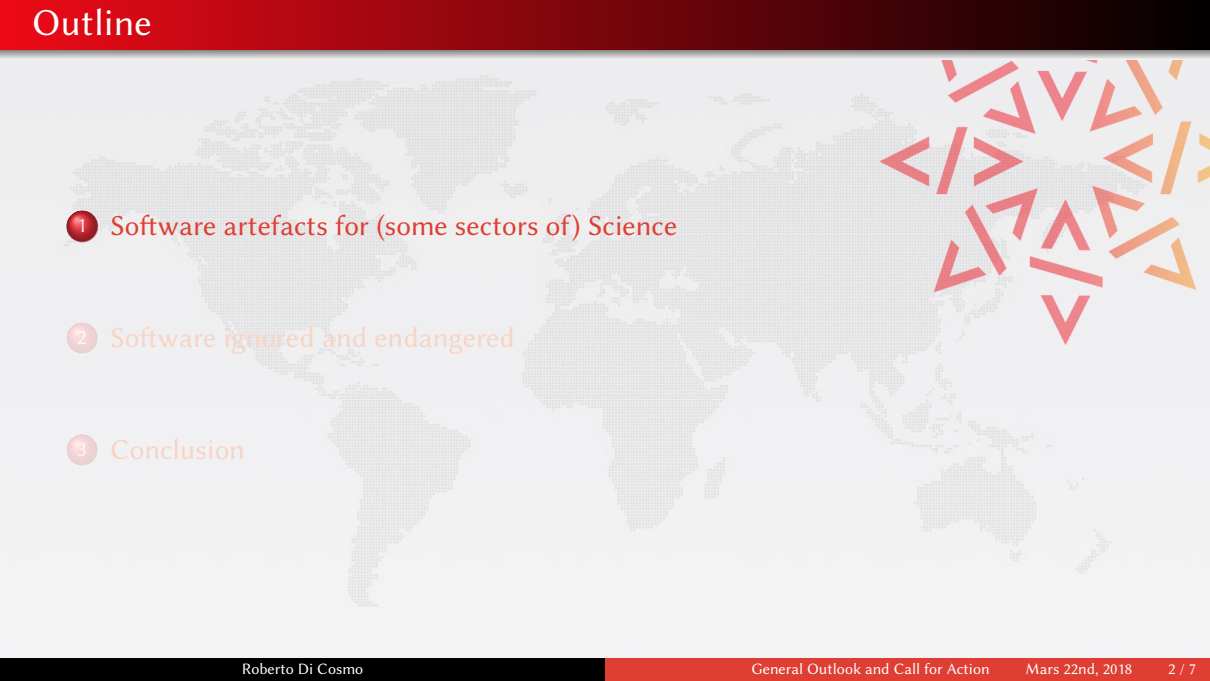
`roberto@dicosmo.org`

Mars 22nd, 2018



# Software Heritage

THE GREAT LIBRARY OF SOURCE CODE



1 Software artefacts for (some sectors of) Science

2 Software ignored and endangered

3 Conclusion

## Why

Necessary to

- *reproduce*,
- *modify and evolve*, **building new experiments** from old ones

# Pressure to make the source code available is raising

## Why

Necessary to

- *reproduce*,
- *modify and evolve*, **building new experiments** from old ones

## When and where

- debate started end of first 2000 decade (biology, statistics, medicine, etc.)
- growing in Computer Science since the ESEC/FSE 2011 Artifact Evaluation context (winner: Vouillon and Di Cosmo); see <http://www.artifact-eval.org/>

## ACM policies: Artifact Review and Badging

### Terminology (not consensual yet!)

- **Repeatability** \ same team, same experimental setup
- **Replicability** \ different team, same experimental setup
- **Reproducibility** \ different team, different experimental setup

## ACM policies: Artifact Review and Badging

### Terminology (not consensual yet!)

- **Repeatability** \ same team, same experimental setup
- **Replicability** \ different team, same experimental setup
- **Reproducibility** \ different team, different experimental setup

### Badging software artefacts



Functional



Reusable



Available



Results replicated



Results reproduced

## ACM policies: Artifact Review and Badging

### Terminology (not consensual yet!)

- **Repeatability** \ same team, same experimental setup
- **Replicability** \ different team, same experimental setup
- **Reproducibility** \ different team, different experimental setup

### Badging software artefacts



Functional



Reusable



Available



Results replicated



Results reproduced

# Why it is important

## Metadata

Evaluated software artifacts must be properly **described**

## Archival

Evaluated software artifacts must be properly **archived**

## Identification

Evaluated software artifacts must be properly **referenced**

## Citation

Evaluated software artifacts must be properly **cited** (*not the same as referenced*)



# Why it is important

## Metadata

Evaluated software artifacts must be properly **described**

## Archival

Evaluated software artifacts must be properly **archived**

## Identification

Evaluated software artifacts must be properly **referenced**

## Citation

Evaluated software artifacts must be properly **cited** (*not the same as referenced*)

This is happening *outside* of RDA



1 Software artefacts for (some sectors of) Science

2 Software ignored and endangered

3 Conclusion

## The points of attention

<http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2016:593:FIN>

**Art. 3 (TDM)** Limitations to text and data mining

**Art. 13 (Upload filtering)** threatens Open Access/Data/Source/Science as a whole!



## The points of attention

<http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2016:593:FIN>


**Art. 3 (TDM)** Limitations to text and data mining

**Art. 13 (Upload filtering)** threatens Open Access/Data/Source/Science as a whole!

## Calendar and actions

- vote of EU parliament expected end of April, **please act now**
  - Research <http://sparceurope.org/copyrightreform>
  - Free Software  
<https://fsfe.org/news/2017/news-20170908-01.en.html>
  - C4C <https://copyright4creativity.eu>
  - VoxScientia <https://voxscientia.eu/>

*More info offline*



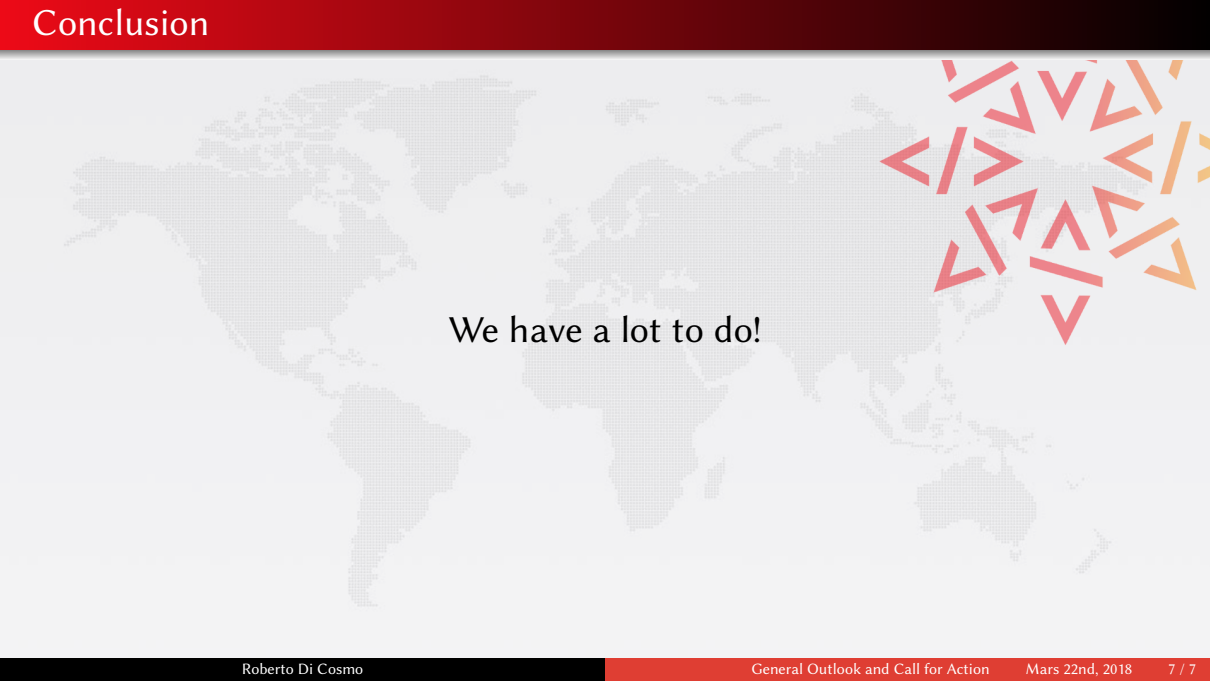
Off the record



1 Software artefacts for (some sectors of) Science

2 Software ignored and endangered

3 Conclusion



We have a lot to do!