Towards Reproducible Research

Miguel Colom[1], José Armando Hernández[1], Bertrand Kerautret[2]

October 8, 2024

[1]Centre Borelli, École Normale Supérieure Paris-Saclay, France
[2]Univ Lyon, Univ Lyon 2, CNRS, INSA Lyon, UCBL, LIRIS, UMR5205, F-69676 Bron, France

Overview

1. Introduction

Introduction: some definitions[1]

Repeatability and Reproducibility

Capacity to perform the same experiment as many times as needed.

→ **Repeatability**: Same team, same experimental setup

→ **Reproducibility**: Different team, same experimental setup

Example: is distilled water electrically conductive? Is salt water conductive? We can perform the experiment many times and get results (https://www.dailymotion.com/video/x2lcg6a).

[1] There are different terminologies, see ACM Artifact Review and Badging, we use the version 1.0

Introduction: some definitions[1]

Repeatability and Reproducibility

Capacity to perform the same experiment as many times as needed.

→ **Repeatability**: Same team, same experimental setup

→ **Reproducibility**: Different team, same experimental setup

Example: is distilled water electrically conductive? Is salt water conductive? We can perform the experiment many times and get results (https://www.dailymotion.com/video/x2lcg6a).

Replicability

Capacity to obtain the same results when repeating an experiment by following a detailed procedure.

→ Different team, different experimental setup

In computational sciences (deterministic code, digital data): results obtained by following a detailed and correct pseudo-code description must be equivalent if the same input data is provided.

[1] There are different terminologies, see ACM Artifact Review and Badging, we use the version 1.0

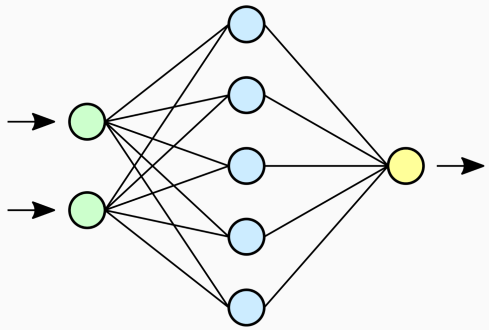
Definition: Repeatability Examples

Repeatable

Obtaining the classification results with a neural network.

We can repeat the experiment as many times as we want.

We just need the weights of the network and the input data.



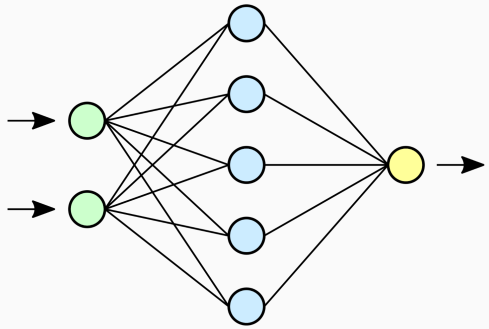
Definition: Repeatability Examples

Repeatable

Obtaining the classification results with a neural network.

We can repeat the experiment as many times as we want.

We just need the weights of the network and the input data.



Not repeatable:

Detection of the merger of two black holes from gravitational waves. We can't repeat the experiment as needed.

Definition: Reproducibility Examples

Reproducible:

Given:

- a detailed pseudo-code (or the source code itself),
- any associated learning or initialization data,
- the input data,

we should obtain exactly the same results each time we run the algorithm.

$\Rightarrow$ Exactly the same denoised image, classification results, etc.

Definition: Reproducibility Examples

Reproducible:

Given:

- a detailed pseudo-code (or the source code itself),
- any associated learning or initialization data,
- the input data,

we should obtain exactly the same results each time we run the algorithm.

$\Rightarrow$ Exactly the same denoised image, classification results, etc.

Not reproducible

In a paper that shows

- a pseudo-code without all the details, or its initialization,
- the source code is not available,
- neither the learning data,

other researchers can't compare with the proposed method.

$\Rightarrow$ We can't be sure about anything on the method, nor test it with their own data.

Motivation: Example on Biomedical Research

Main Keys Points

- 2009: David Donoho points out a credibility crisis in scientific research

Motivation: Example on Biomedical Research

Main Keys Points

- 2009: David Donoho points out a credibility crisis in scientific research
- 2012: the director of the oncology division at Amgen: tried to reproduce 53 of the most important papers on oncology. He failed to reproduce 47 of them.
  Sources:
  https://www.nature.com/articles/nature.2016.19269

  https://www.nature.com/articles/483531a

Motivation: Example on Biomedical Research

Main Keys Points

- 2009: David Donoho points out a credibility crisis in scientific research
- 2012: the director of the oncology division at Amgen: tried to reproduce 53 of the most important papers on oncology. He failed to reproduce 47 of them.
  Sources:
  https://www.nature.com/articles/nature.2016.19269

  https://www.nature.com/articles/483531a
- Bayer HealthCare Germany confirmed: only 25% of cancer research is reproducible.
  Source: https://www.nature.com/articles/nrd3439-c1

Disciplines and RR

Different disciplines

- **Computational sciences**
- Biomedical, biology
- Social sciences
- ...

Each of them with their own requirements and protocols.

Implementation of Reproducible Research

- Non-exact sciences (biology, medicine, . . .): difficult (but *desirable*). Hard to have exactly the same conditions along experiments.

Implementation of Reproducible Research

- Non-exact sciences (biology, medicine, . . .): difficult (but *desirable*). Hard to have exactly the same conditions along experiments.
- Computational sciences: no excuse!

What about non-deterministic algorithms?

Some algorithms might be chaotic, or a procedure be non-deterministic.

What about non-deterministic algorithms?

Some algorithms might be chaotic, or a procedure be non-deterministic.

- Training of large neural networks. Random initialization, random dropout, ...
- Unique data from observations. Example: merger of two black holes.
- ...

What about non-deterministic algorithms?

Some algorithms might be chaotic, or a procedure be non-deterministic.

- Training of large neural networks. Random initialization, random dropout, ...
- Unique data from observations. Example: merger of two black holes.
- ...

How can we perform RR with them?

Statistics

In **reproducibility** we're interested in the validity of the method or the procedure.

Statistics

In **reproducibility** we're interested in the validity of the method or the procedure.

- Results should be comparable and equivalent
- Two different training processes will produce different weights
- However, the results of the inferenec need to be statistically equivalent

Statistics

In **reproducibility** we're interested in the validity of the method or the procedure.

- Results should be comparable and equivalent
- Two different training processes will produce different weights
- However, the results of the inferenec need to be statistically equivalent

Statistical methods to decide if results are equivalent. I.e. hypotesis testing.

Why are not we all doing reproducible research?

Why are not we all doing reproducible research?

Several reasons in general:

- Some researchers don't want to make public working code
  - doesn't correspond to any version of the pseudo-codes,
  - low software quality,
  - quality software takes more time to produce: testing, documentation, objective quality metrics, ...
- Results of a method do not generalize
- ... (For the discussion later!)

Why are not we all doing reproducible research?

Several reasons in general:

- Some researchers don't want to make public working code
  - doesn't correspond to any version of the pseudo-codes,
  - low software quality,
  - quality software takes more time to produce: testing, documentation, objective quality metrics, ...
- Results of a method do not generalize
- ... (For the discussion later!)

Not really considered for career advance

- Classic metrics: "number of high impact-factor classic publications"
- Software is considered as a 2nd class citizen

Good news:

The French government is promoting a Plan for Open Science

- https://www.ouvrirlascience.fr/second-national-plan-for-open-scienc
- "Path three: opening up and promoting source code produced by research"
- Not directly on reproducible research, but certainly fosters it
- → Let's share in the discussion later other initiatives or experiences you know

Good news:

The French government is promoting a Plan for Open Science

- https://www.ouvrirlascience.fr/second-national-plan-for-open-scienc
- "Path three: opening up and promoting source code produced by research"
- Not directly on reproducible research, but certainly fosters it
- → Let's share in the discussion later other initiatives or experiences you know

Initiatives at European level

- European Open Science Cloud (EOSC): recommendations to the European Commission

  https://www.eosc.eu
- Publication: "Scholarly infrastructures for research software"

  https://op.europa.eu/en/publication-detail/-/publication/145fd0f3-3907-11eb-b27b-01aa75ed71a1/language-en

2. Efforts on Software Useful for Reproducible Research

2. Efforts on software useful for reproducible research

2.1 Updated Preview of Reproducible Research Platforms

- Synthetic recall of main platforms [CKK19].
- Updates with news elements from last publication.

2. Efforts on software useful for reproducible research

2.1 Updated Preview of Reproducible Research Platforms

- Synthetic recall of main platforms [CKK19].
- Updates with news elements from last publication.

2.2 Library development

- Another ways to help for RR.
- Examples of library efforts from image domains.

2.1 Updated Preview of Reproducible Research Platforms

Different types of platforms [CKK19]

- Online execution platforms.
- Dissemination platforms.
- Peer-reviewed journals.

- **Galaxy** - https://galaxyproject.org
- **IPython** - https://ipython.org
- **Jupyter** - http://jupyter.org
- **RunMyCode** - http://www.runmycode.org
- **Code Ocean** - https://codeocean.com
- **DAE** - http://dae.cse.lehigh.edu/DAE
- **IPOL** - https://www.ipol.im
- **Research Compendia** - ResearchCompendia.org
- **MLOSS** - https://mloss.org/software
- **DataHub** - https://datahub.io/
- **PaperWithCode** - https://paperswithcode.com

- **ReScience Journal** - http://rescience.github.io
- **JOSS Journal** - https://joss.theoj.org
- **Insight J Journal** - https://insight-journal.org

2.1 Updated Preview of Reproducible Research Platforms

Different types of platforms [CKK19]

- Online execution platforms.
- Dissemination platforms.
- Peer-reviewed journals.

- **Galaxy** - https://galaxyproject.org
- **IPython** - https://ipython.org
- **Jupyter** - http://jupyter.org
- **RunMyCode** - http://www.runmycode.org
- **Code Ocean** - https://codeocean.com
- **DAE** - http://dae.cse.lehigh.edu/DAE
- **IPOL** - https://www.ipol.im
- **Research Compendia** - ResearchCompendia.org
- **MLOSS** - https://mloss.org/software
- **DataHub** - https://datahub.io/
- **PaperWithCode** - https://paperswithcode.com

- **ReScience Journal** - http://rescience.github.io
- **JOSS Journal** - https://joss.theoj.org
- **Insight J Journal** - https://insight-journal.org

2.1 Updated Preview of Reproducible Research Platforms

Different types of platforms [CKK19]

- Online execution platforms.
- Dissemination platforms.
- Peer-reviewed journals.

- **Galaxy** - https://galaxyproject.org
- **IPython** - https://ipython.org
- **Jupyter** - http://jupyter.org
- **RunMyCode** - http://www.runmycode.org
- **Code Ocean** - https://codeocean.com
- **DAE** - http://dae.cse.lehigh.edu/DAE
- **IPOL** - https://www.ipol.im
- **Research Compendia** - ResearchCompendia.org
- **MLOSS** - https://mloss.org/software
- **DataHub** - https://datahub.io/
- **PaperWithCode** - https://paperswithcode.com

- **ReScience Journal** - http://rescience.github.io
- **JOSS Journal** - https://joss.theoj.org
- **Insight J Journal** - https://insight-journal.org

New Reproducible Research Platform (1)

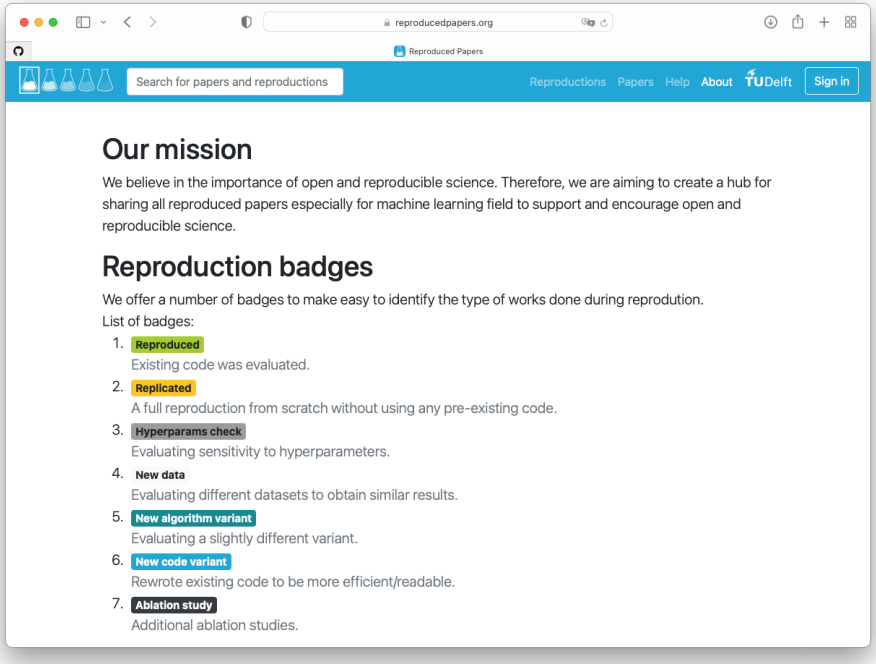
Platform https://reproducedpapers.org **[Dissemination platforms]**

- Presented during previous edition of RRPR 2020 [YHK+21].
- Oriented machine learning reproducibility (in teaching context - master degree -).

New Reproducible Research Platform (1)

Platform https://reproducedpapers.org **[Dissemination platforms]**

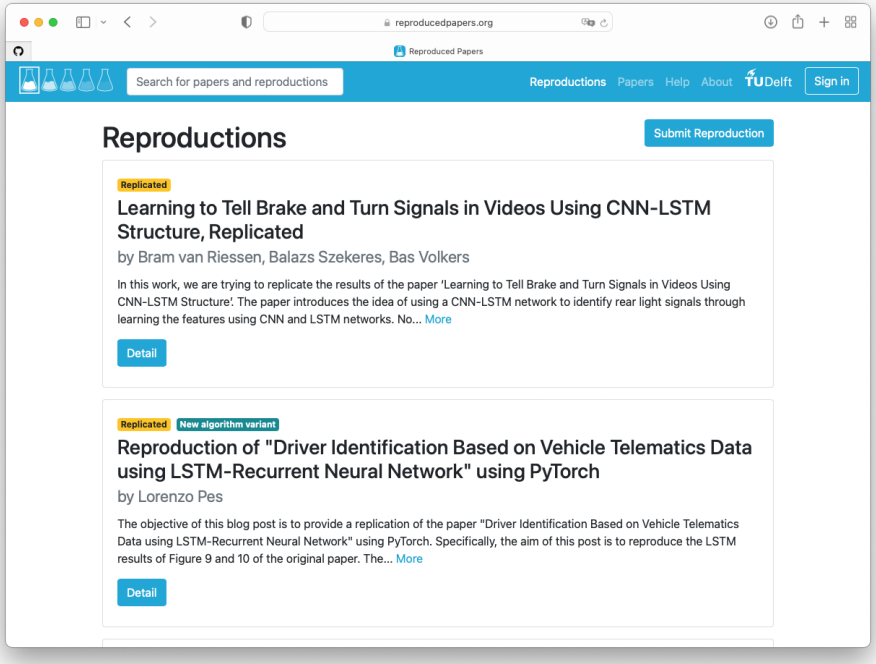
- Presented during previous edition of RRPR 2020 [YHK+21].
- Oriented machine learning reproducibility (in teaching context - master degree -).
- Main ideas:
  - Suggest a paper to be reproduced.
  - Propose a reproduction/replication/new variant/new data.
  - A paper/web page describing reproduction.



New Reproducible Research Platform (1)

Platform https://reproducedpapers.org **[Dissemination platforms]**

- Presented during previous edition of RRPR 2020 [YHK+21].
- Oriented machine learning reproducibility (in teaching context - master degree -).
- Main ideas:
  - Suggest a paper to be reproduced.
  - Propose a reproduction/replication/new variant/new data.
  - A paper/web page describing reproduction.



New Reproducible Research Platform (2)

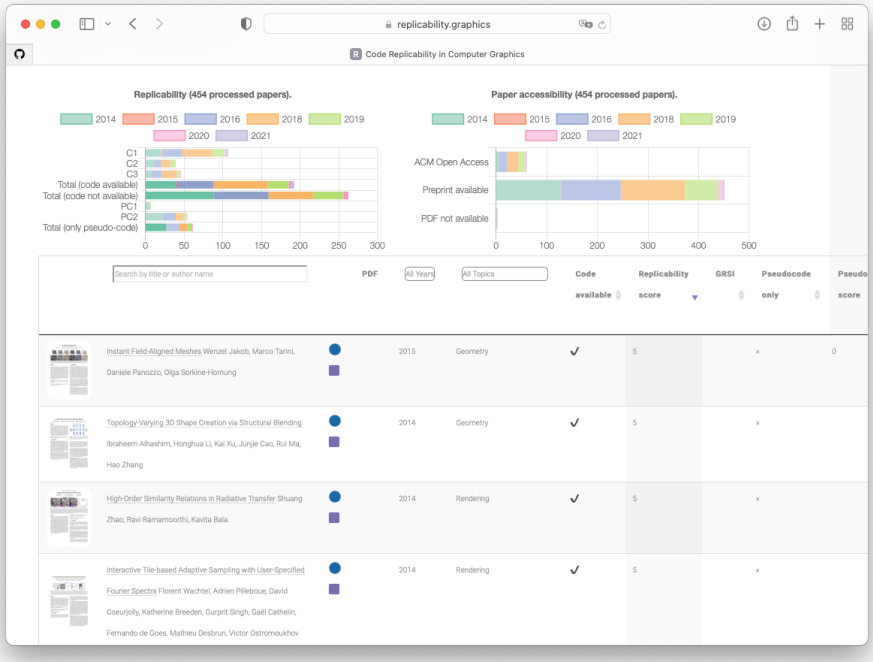
Platform `https://replicability.graphics` **[Dissemination platforms]**

- Invited in previous edition of RRPR [BCDM20].
- Computer Graphics
- Replicate and evaluate difficult-to-reproduce SIGGRAPH papers
- Light formated review rating the success of replication

New Reproducible Research Platform (2)

Platform https://replicability.graphics **[Dissemination platforms]**

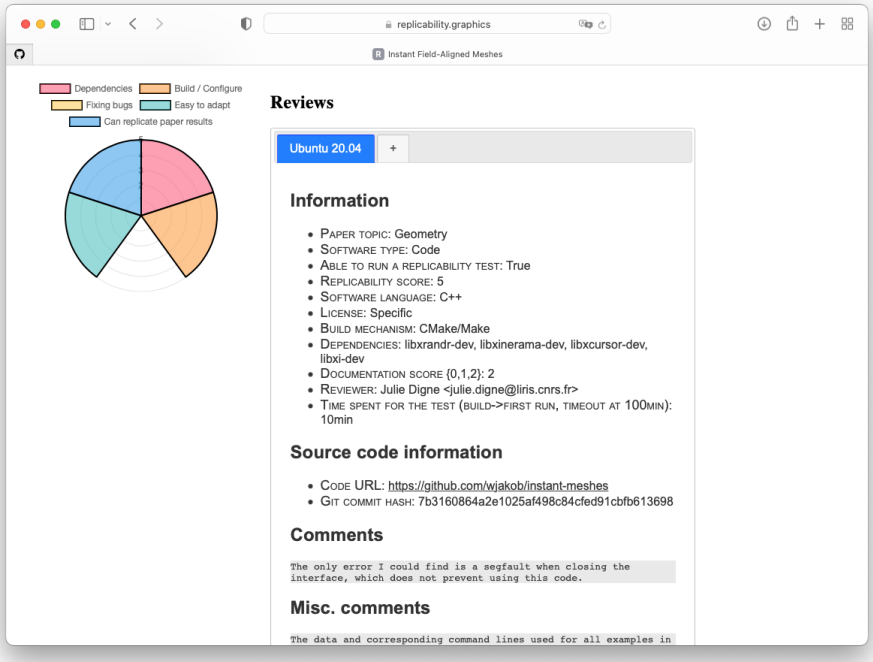
- Invited in previous edition of RRPR [BCDM20].
- Computer Graphics
- Replicate and evaluate difficult-to-reproduce SIGGRAPH papers
- Light formated review rating the success of replication



New Reproducible Research Platform (2)

Platform https://replicability.graphics **[Dissemination platforms]**

- Invited in previous edition of RRPR [BCDM20].
- Computer Graphics
- Replicate and evaluate difficult-to-reproduce SIGGRAPH papers
- Light formated review rating the success of replication



New Reproducible Research Platform (3)

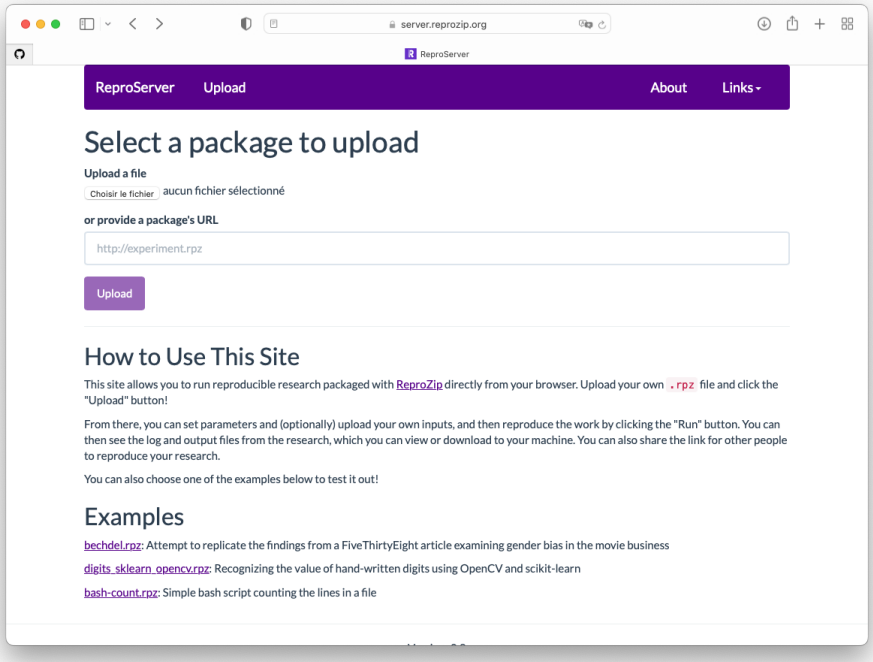
ReproZip: https://server.reprozip.org **[Online execution platforms]**

- Open source tool to bundles everything necessary to reproduce your work.
- Way to share reproducible work from the ReproServer.

New Reproducible Research Platform (3)

ReproZip: https://server.reprozip.org **[Online execution platforms]**

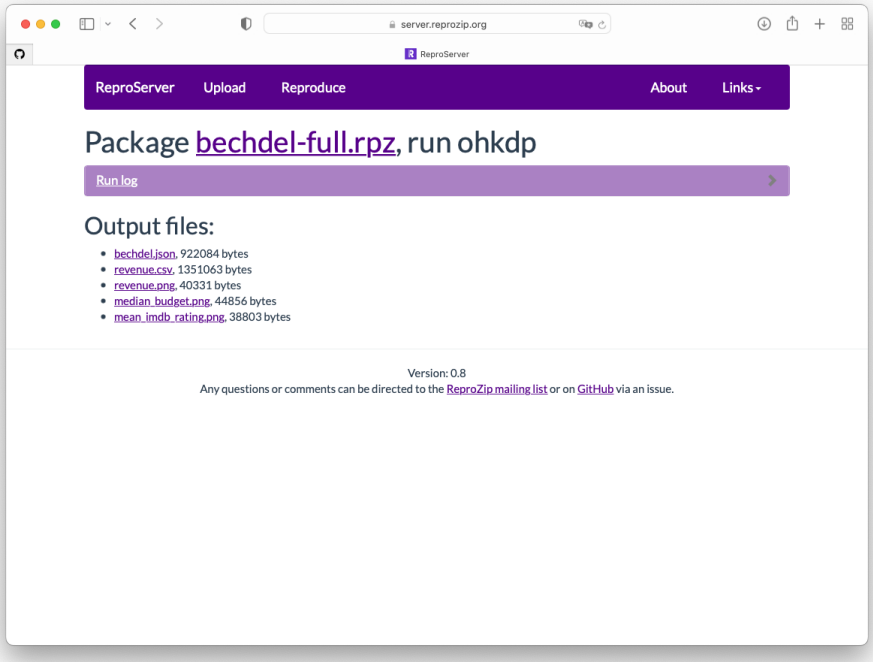
- Open source tool to bundles everything necessary to reproduce your work.
- Way to share reproducible work from the ReproServer.



New Reproducible Research Platform (3)

ReproZip: https://server.reprozip.org **[Online execution platforms]**

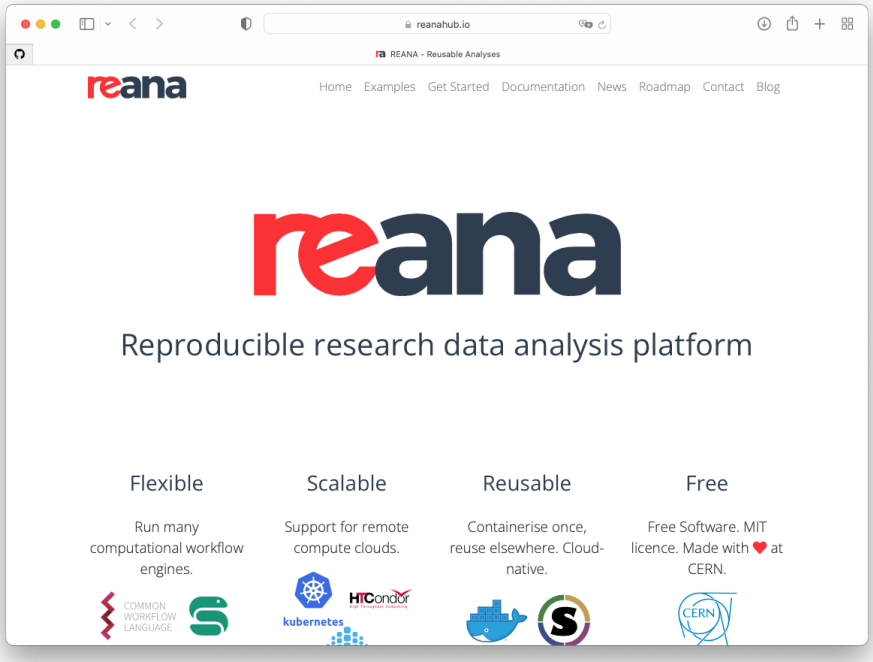
- Open source tool to bundles everything necessary to reproduce your work.
- Way to share reproducible work from the ReproServer.



New Reproducible Research Platform (4)

Platform REANA : https://reanahub.io [Online execution platforms]

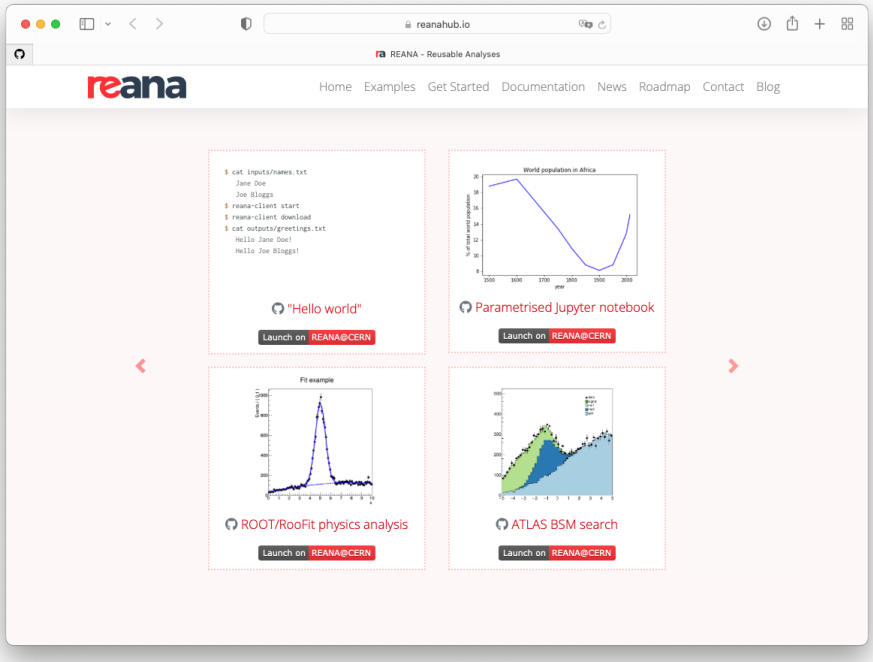
- Reproducible research data analysis platform.



New Reproducible Research Platform (4)

Platform REANA : https://reanahub.io **[Online execution platforms]**

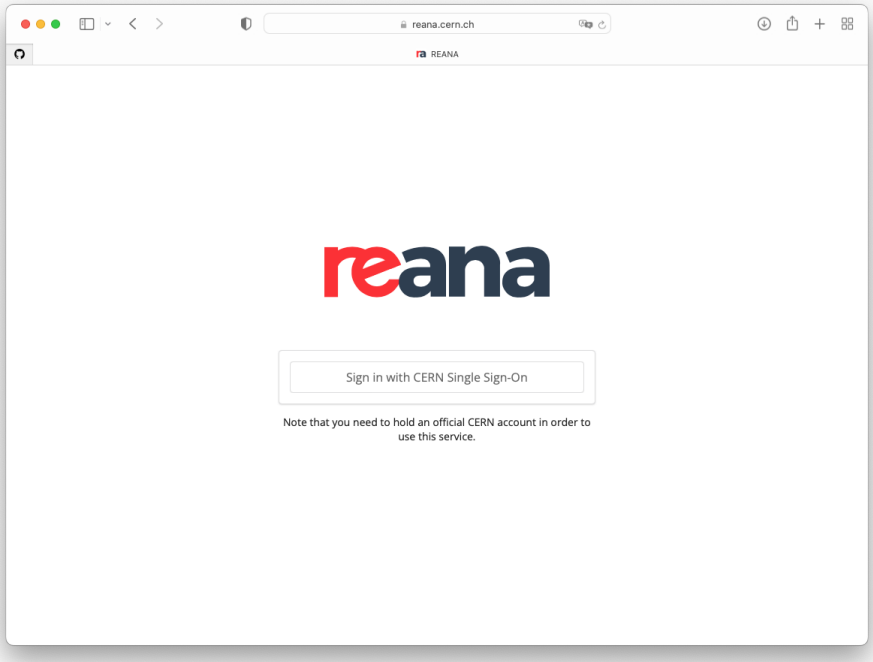
- Reproducible research data analysis platform.
- Allow to launch workflows from external sources (GitHub, GitLab).



New Reproducible Research Platform (4)

Platform REANA : https://reanahub.io **[Online execution platforms]**

- Reproducible research data analysis platform.
- Allow to launch workflows from external sources (GitHub, GitLab).
- Reserved to CERN members only.
  $\Rightarrow$ limited impact



New Reproducible Research Platform (5)

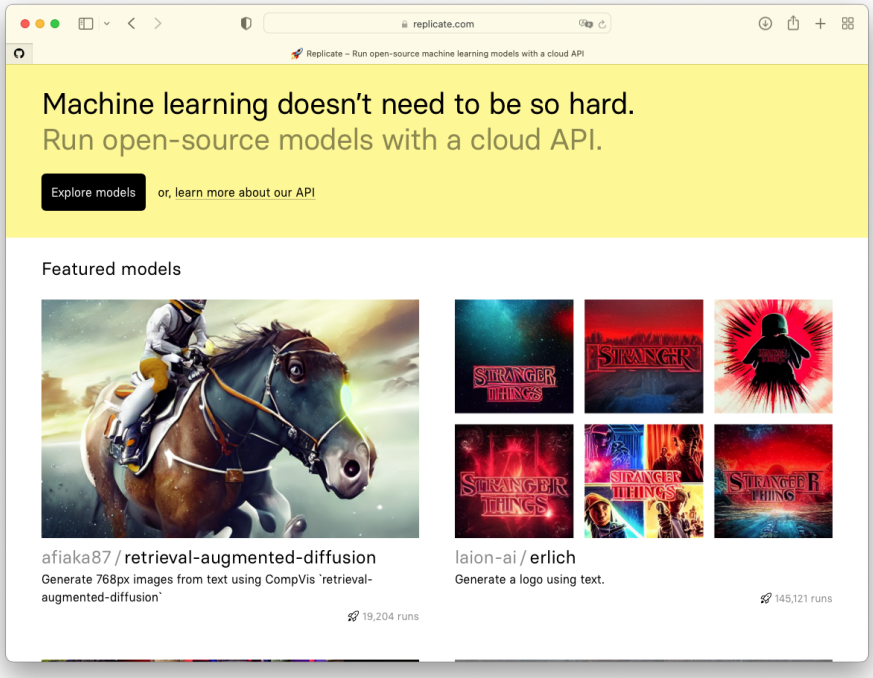
Platform https://replicate.com **[Online execution platforms]**

- Machine learning code replication.
- Applying ML from real-world problem without complex code installation.
- Propose model, run online, run with API or from your computer.
  $\Rightarrow$ base on COG (open source tool to package a production ready container)
  https://github.com/replicate/cog

New Reproducible Research Platform (5)

Platform https://replicate.com **[Online execution platforms]**

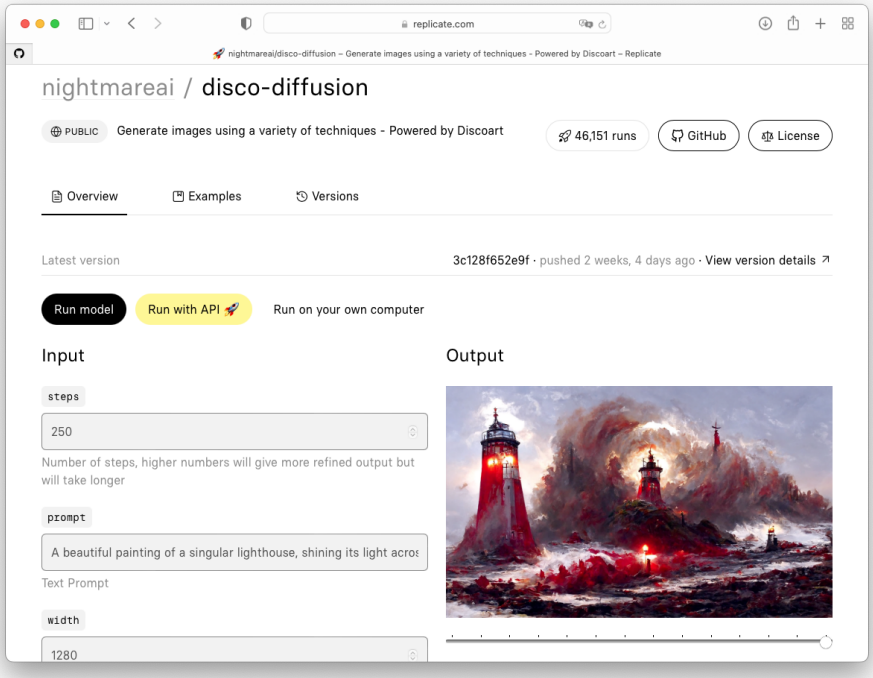
- Machine learning code replication.
- Applying ML from real-world problem without complex code installation.
- Propose model, run online, run with API or from your computer.
  ⇒ base on COG (open source tool to package a production ready container)
  https://github.com/replicate/cog



New Reproducible Research Platform (5)

Platform https://replicate.com **[Online execution platforms]**

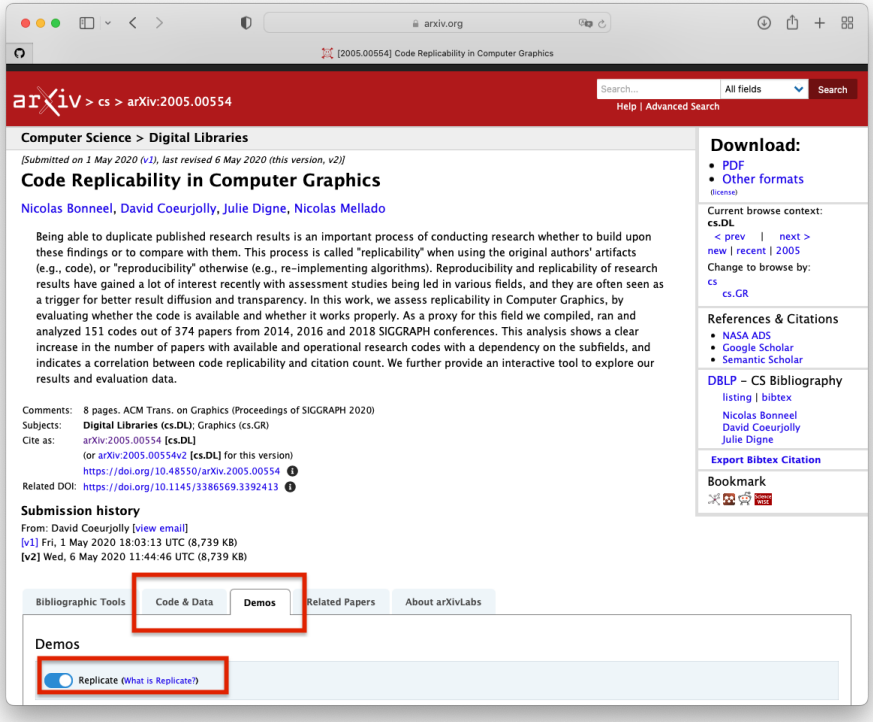
- Machine learning code replication.
- Applying ML from real-world problem without complex code installation.
- Propose model, run online, run with API or from your computer.
  $\Rightarrow$ base on COG (open source tool to package a production ready container)
  https://github.com/replicate/cog



New Reproducible Research Platform (5)

Platform https://replicate.com [Online execution platforms]

- Machine learning code replication.
- Applying ML from real-world problem without complex code installation.
- Propose model, run online, run with API or from your computer.
  $\Rightarrow$ base on COG (open source tool to package a production ready container)
  https://github.com/replicate/cog



2.2 Library Development for RR

Libraries and contribution to RR

- Way to diffuse and reuse research methods
- Documented and referenced algorithms
- Extend the classic academic results to real applications
- Gather algorithms and make easier comparisons and use in other context
- Make easier software demonstrator or online demonstrations

2.2 Library Development for RR

Libraries and contribution to RR

- Way to diffuse and reuse research methods
- Documented and referenced algorithms
- Extend the classic academic results to real applications
- Gather algorithms and make easier comparisons and use in other context
- Make easier software demonstrator or online demonstrations

Attentive key point towards reproducibility

- Compatibility problems: libraries evolve rapidly (compilation issues).
- Implementation change: can change numerical results from a version to another.
- Implies the use of container solution like *Docker* to ensure longer terms reproducibility. **Discussion**: docker containers, replicability, reproducibility, maintainability.

2.2 Library Development for RR: examples (1)

Example of library of vision/image/geometry domains

Library	ref	domain	langage	version	#auth.	date	funding
OpenCV	[Bra00]	Comp. Vision	C++	4.5.5	1,383	1999	Willow Garage
ITK	[MLJ+14]	Image Processing	C++/Pyt.	5.2.1	265	2000	Kitware
PCL	[RC11]	Point clouds	C++	1.12.1	464	2010	Willow Garage
CGal	[The22]	Geometry proc.	C++	5.4	123	1996	Acad./GeometryFactory
CImg	[Tsc12]	Image processing	C++	3.1.2	72	1999	Acad.
Geogram	[Geo]	Geometric algorith.	C++	1.7.8	7	1998	Acad./INRIA/ERC
Olena	[RCG18]	Image processing	C++/Pyt.	2.1	50	2001	Acad. / Project
Tulip	[Aub04]	huge graph visualiz	C++/Pyt.	5.6.2	9	2001	Acad./private
Vigra	[Vig]	Comp. Vision	C++	1.11	50	2008	Acad.
DGtal	[DGt]	Digital geometry	C++/Pyt.	1.2	27	2011	Acad. / Project
OpenMVG	[MMPM17]	Mult. View Geom.	C++	2.0	86	2013	Acad./Mikros/Foxel
TTK	[TFL+17]	Topology ToolKit	C++	1.0	36	2017	Acad. / Project
Higra	[HIG]	Graph analysis	C++/Pyt.	0.6.5	4	2018	Acad. / Project

2.2 Library Development for RR: examples (1)

Example of library of vision/image/geometry domains

- Some libraries were initiated/funded by private company.

Library	ref	domain	langage	version	#auth.	date	funding
OpenCV	[Bra00]	Comp. Vision	C++	4.5.5	1,383	1999	Willow Garage
ITK	[MLJ+14]	Image Processing	C++/Pyt.	5.2.1	265	2000	Kitware
PCL	[RC11]	Point clouds	C++	1.12.1	464	2010	Willow Garage
CGal	[The22]	Geometry proc.	C++	5.4	123	1996	Acad./GeometryFactory
CImg	[Tsc12]	Image processing	C++	3.1.2	72	1999	Acad.
Geogram	[Geo]	Geometric algorith.	C++	1.7.8	7	1998	Acad./INRIA/ERC
Olena	[RCG18]	Image processing	C++/Pyt.	2.1	50	2001	Acad. / Project
Tulip	[Aub04]	huge graph visualiz	C++/Pyt.	5.6.2	9	2001	Acad./private
Vigra	[Vig]	Comp. Vision	C++	1.11	50	2008	Acad.
DGtal	[DGt]	Digital geometry	C++/Pyt.	1.2	27	2011	Acad. / Project
OpenMVG	[MMPM17]	Mult. View Geom.	C++	2.0	86	2013	Acad./Mikros/Foxel
TTK	[TFL+17]	Topology ToolKit	C++	1.0	36	2017	Acad. / Project
Higra	[HIG]	Graph analysis	C++/Pyt.	0.6.5	4	2018	Acad. / Project

2.2 Library Development for RR: examples (1)

Example of library of vision/image/geometry domains

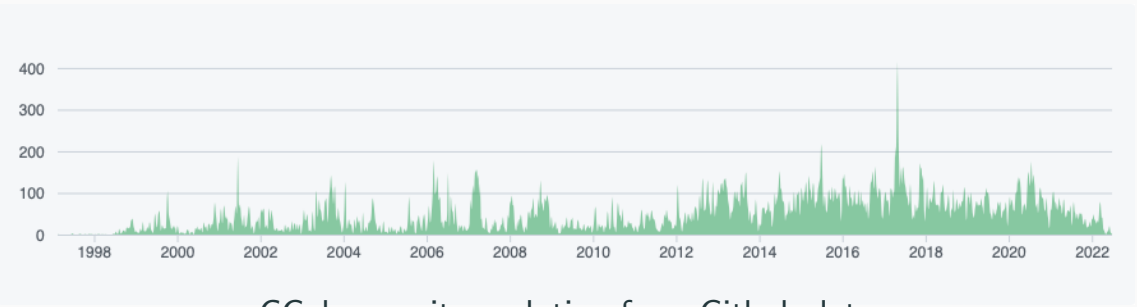
- Some libraries were initiated/funded by private company.
- The others come mainly from university initiatives.

Library	ref	domain	langage	version	#auth.	date	funding
OpenCV	[Bra00]	Comp. Vision	C++	4.5.5	1,383	1999	Willow Garage
ITK	[MLJ+14]	Image Processing	C++/Pyt.	5.2.1	265	2000	Kitware
PCL	[RC11]	Point clouds	C++	1.12.1	464	2010	Willow Garage
CGal	[The22]	Geometry proc.	C++	5.4	123	1996	Acad./GeometryFactory
CImg	[Tsc12]	Image processing	C++	3.1.2	72	1999	Acad.
Geogram	[Geo]	Geometric algorith.	C++	1.7.8	7	1998	Acad./INRIA/ERC
Olena	[RCG18]	Image processing	C++/Pyt.	2.1	50	2001	Acad. / Project
Tulip	[Aub04]	huge graph visualiz	C++/Pyt.	5.6.2	9	2001	Acad./private
Vigra	[Vig]	Comp. Vision	C++	1.11	50	2008	Acad.
DGtal	[DGt]	Digital geometry	C++/Pyt.	1.2	27	2011	Acad. / Project
OpenMVG	[MMPM17]	Mult. View Geom.	C++	2.0	86	2013	Acad./Mikros/Foxel
TTK	[TFL+17]	Topology ToolKit	C++	1.0	36	2017	Acad. / Project
Higra	[HIG]	Graph analysis	C++/Pyt.	0.6.5	4	2018	Acad. / Project

2.2 Library Development for RR: examples (2)

Long time efforts

- Commit evolution over 25 years from the CGal library [The22].

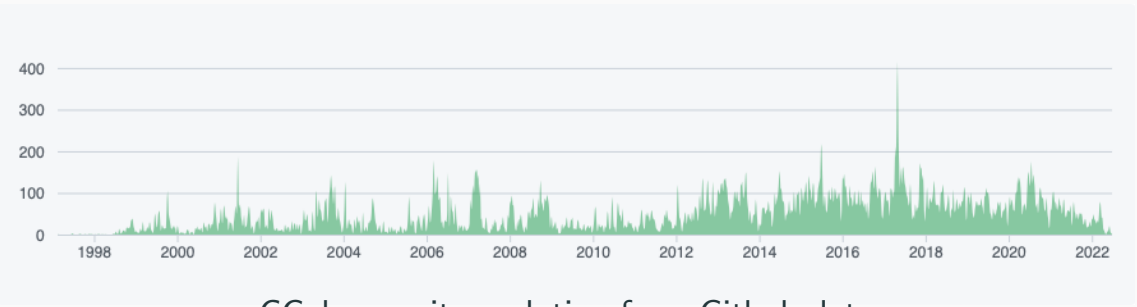


CGal commits evolution from Github data

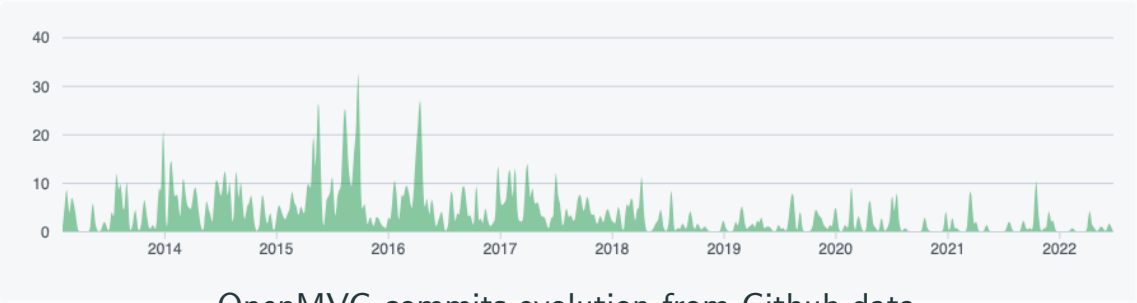
2.2 Library Development for RR: examples (2)

Long time efforts

- Commit evolution over 25 years from the CGal library [The22].
- Commit evolution over 10 years from the OpenMVG [MMPM17].



CGal commits evolution from Github data

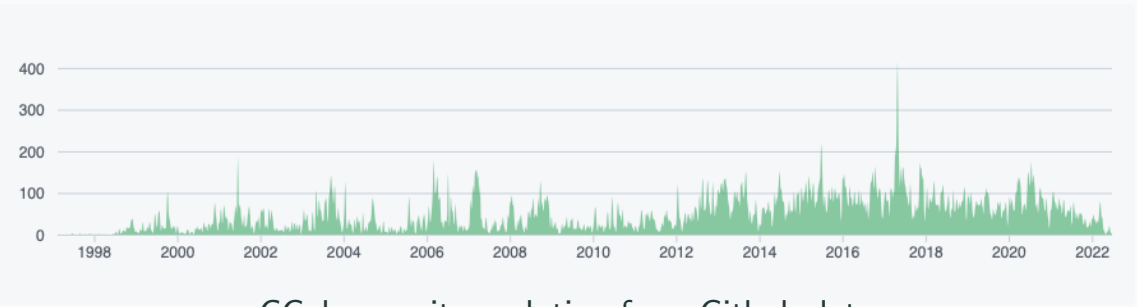


OpenMVG commits evolution from Github data

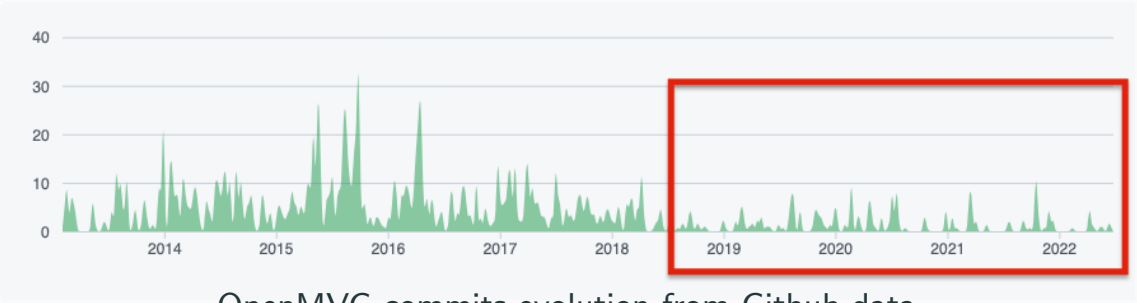
2.2 Library Development for RR: examples (2)

Long time efforts

- Commit evolution over 25 years from the CGal library [The22].
- Commit evolution over 10 years from the OpenMVG [MMPM17].
  $\Rightarrow$ no more funding from 2016 (open source only)



CGal commits evolution from Github data



OpenMVG commits evolution from Github data

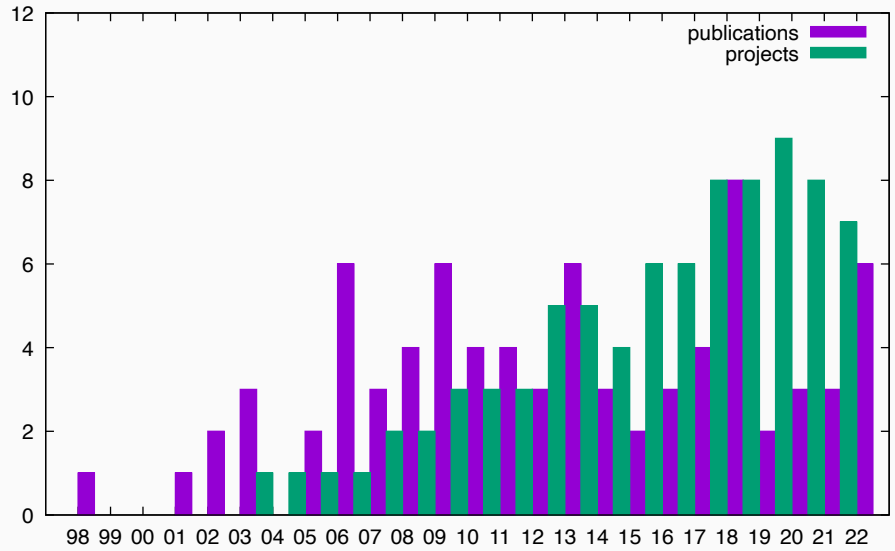
2.2 Library Development for RR: examples (2)

Long time efforts

- Commit evolution over 25 years from the CGal library [The22].
- Commit evolution over 10 years from the OpenMVG [MMPM17].
  $\Rightarrow$ no more funding from 2016 (open source only)

Example of impact measure

- Geogram library [Geo] (7 authors).
- Estimated from the authors declaring using the library in projects/publications.



2.2 Library Development for RR

Elements to increase reproducibility across libraries

- Limit dependancies to other libraries or include them inside the library itself.
- Orient to an header only library (if C++ like CImg or CGal since version 5).
- From a user use, extract all dependent header files.

2.2 Library Development for RR

Elements to increase reproducibility across libraries

- Limit dependancies to other libraries or include them inside the library itself.
- Orient to an header only library (if C++ like CImg or CGal since version 5).
- From a user use, extract all dependent header files.

Examples in C++:

It is possible to use a special option in order to list all header files:

```
g++ hello.c -o hello -I/specialHeaderLibPath  -MD
```

Then command will generate a `hello.d` file containing all header files that are really used by the compiler to produce the executable.

2.2 Library Development for RR

Elements to increase reproducibility across libraries

- Limit dependancies to other libraries or include them inside the library itself.
- Orient to an header only library (if C++ like CImg or CGal since version 5).
- From a user use, extract all dependent header files.

Examples in C++:

It is possible to use a special option in order to list all header files:

```
g++ hello.c -o hello -I/specialHeaderLibPath  -MD
```

Then command will generate a `hello.d` file containing all header files that are really used by the compiler to produce the executable.

$\Rightarrow$ then an archive can be constructed including only needed files and independently of library or system header file.

3. Advanced Editorial Efforts

3. Advanced Editorial Investment: IPOL Publications

Peer-reviewed

- Both the article (PDF) and the source code.
- Reproducibility: the reviewers check carefully that the source code matches the pseudo-code.

3. Advanced Editorial Investment: IPOL Publications

Peer-reviewed

- Both the article (PDF) and the source code.
- Reproducibility: the reviewers check carefully that the source code matches the pseudo-code.

Each publication:

- A text describing the method in detail, including pseudo-codes.
- The source code, under an open-source software license.
- An online demo which allows users to test the method with their own data.
- An archive of experiments.

3. Advanced Editorial Investment: IPOL Publications

Peer-reviewed

- Both the article (PDF) and the source code.
- Reproducibility: the reviewers check carefully that the source code matches the pseudo-code.

Each publication:

- A text describing the method in detail, including pseudo-codes.
- The source code, under an open-source software license.
- An online demo which allows users to test the method with their own data.
- An archive of experiments.

- No need to be an original work. We're interested in the math details, reproducibility, and understanding.
- ISSN, DOI, indexed by SCOPUS. Not yet an "Impact Factor".

3. Advanced Editorial Investment: IPOL Publications

Peer-reviewed

- Both the article (PDF) and the source code.
- Reproducibility: the reviewers check carefully that the source code matches the pseudo-code.

Each publication:

- A text describing the method in detail, including pseudo-codes.
- The source code, under an open-source software license.
- An online demo which allows users to test the method with their own data.
- An archive of experiments.

- No need to be an original work. We're interested in the math details, reproducibility, and understanding.
- ISSN, DOI, indexed by SCOPUS. Not yet an "Impact Factor".

Let's have a look! ⇒ http://www.ipol.im/pub/art/2017/201/

3.1 IPOL publications: Benefits of RR, from our experience

If a method is worth it, the impact is large

- Users worldwide can test the algorithm with their own data.
- Increase in the number of citations: other researchers can now compare to you.

3.1 IPOL publications: Benefits of RR, from our experience

If a method is worth it, the impact is large

- Users worldwide can test the algorithm with their own data.
- Increase in the number of citations: other researchers can now compare to you.

Scientific acceleration:

$\Rightarrow$ other researchers can reuse text, source code, data

3.1 IPOL publications: Benefits of RR, from our experience

If a method is worth it, the impact is large

- Users worldwide can test the algorithm with their own data.
- Increase in the number of citations: other researchers can now compare to you.

Scientific acceleration:

$\Rightarrow$ other researchers can reuse text, source code, data

Useful for Centre Borelli to show a landscape of our scientific activity

A way of working!

3.2 New initiative with OVD-SaaS

Main idea of the project

- New long-term project at **Centre Borelli**
- Inspired on the experience of IPOL
- Opening industrial software as a service (SaaS) to final users
- Needless to say: based on reproducible research and sharing data
- Do not focus on technical details of low-level services, but solving a concrete and final **industrial need**

3.2 New initiative with OVD-SaaS

Main idea of the project

- New long-term project at **Centre Borelli**
- Inspired on the experience of IPOL
- Opening industrial software as a service (SaaS) to final users
- Needless to say: based on reproducible research and sharing data
- Do not focus on technical details of low-level services, but solving a concrete and final **industrial need**
- Examples of application domains:
  $\Rightarrow$ Security, quality management, health and biometry, predictive maintenance, ...

3.2 New initiative with OVD-SaaS

Main idea of the project

- New long-term project at **Centre Borelli**
- Inspired on the experience of IPOL
- Opening industrial software as a service (SaaS) to final users
- Needless to say: based on reproducible research and sharing data
- Do not focus on technical details of low-level services, but solving a concrete and final **industrial need**
- Examples of application domains:
  $\Rightarrow$ Security, quality management, health and biometry, predictive maintenance, ...

Even more !

SaclAI-School, MVA track on reproducible research, ...

3.2 New initiative with OVD-SaaS: feedback

Preliminary feedback

- First "**ML-Briefs**" in April 2022. Second in October 2022.
- Publication of short ML papers.

3.2 New initiative with OVD-SaaS: feedback

Preliminary feedback

- First "**ML-Briefs**" in April 2022. Second in October 2022.
- Publication of short ML papers.
- Some challenges:
  - **Evaluation** of **neural networks**. They're not "*classic*" code. Evaluate type of architecture, how the training was conducted (initialization, which datasets. etc). Black-box model to check same outputs.
  - **Instability** of libraries. Use of Python's *venv* not enough. Other dependencies at OS and hardware levels. Using Docker containers.

3.3 Reviewing Reproducible Research: How?

Main attention points:

- Consider **source code** as part of the publication, not supplementary material
- Different **levels of evaluation**:
  - Lowest: **black box** (same inputs $\rightarrow$ same outputs)
  - ...
  - Highest: deep understanding of the method and checking that the source code matches exactly the implementation.

3.3 Reviewing Reproducible Research: How?

Main attention points:

- Consider **source code** as part of the publication, not supplementary material
- Different **levels of evaluation**:
  - Lowest: **black box** (same inputs $\rightarrow$ same outputs)
  - ...
  - Highest: deep understanding of the method and checking that the source code matches exactly the implementation.

Difficulty

- **Software** is not easy to review.
- Not all researchers are software engineers!

$\Rightarrow$ A possible solution (IPOL): use **at least two reviewers**, one of them being an expert reading source code.

3.3 Reviewing Reproducible Research: How?

Main attention points:

- Consider **source code** as part of the publication, not supplementary material
- Different **levels of evaluation**:
  - Lowest: **black box** (same inputs → same outputs)
  - ...
  - Highest: deep understanding of the method and checking that the source code matches exactly the implementation.

Difficulty

- **Software** is not easy to review.
- Not all researchers are software engineers!

⇒ A possible solution (IPOL): use **at least two reviewers**, one of them being an expert reading source code.

Special case of neural networks

Focus on the **architecture**, **training**, **understanding**, and **generalization**.

3.3 Reviewing Reproducible Research: How? (2)

Important point

- **Versioning** is important!
  - Research source code **changes** constantly
  - A publication is steady
- Much of the code is publicly available on platforms such as Gitlab, Github, and others

3.3 Reviewing Reproducible Research: How? (2)

Important point

- **Versioning** is important!
  - Research source code **changes** constantly
  - A publication is steady
- Much of the code is publicly available on platforms such as Gitlab, Github, and others

Need to identify software artifacts at different levels

- **DOI is not enough**: the pointed object might change (or even disappear!). It's just a **pointer**.
- Need of a **persistent** and **universal archive**
- A good candidate: UNESCO's **Software Heritage** project [DCGZ18]
- **SWHID**: SoftWare Heritage persistent IDentifiers [DCGZ19]
  - **At different levels**: contents, directories, revisions, releases, snapshots
  - Points to **perpetual objects** in an universal archive

A quick review of software and platforms for RR (1)

Scientific Workflows management systems [CBBC+17, Hat17]

- Taverna: open source scientific workflow management system that provides a workbench for designing workflows, used in **biology, chemistry, meteorology, social sciences**
- Galaxy: is very popular workflow system in the **bioinformatics** community and **life sciences**.
- OpenAlea is an **open source** scientific workflow system with a **visual programming environment** that enables users to design, execute and interact with workflows
- Nextflow: is a **command-line** based workflow system implemented in Groovy (a scripting language for the Java Virtual Machine), developed to **simplify the design of complex parallel scientific workflows**
- Chimera system: originated in support of **data-intensive physics computations** as a means to capture and automate a **complex pipeline of transformations** on the data by external software

A quick review of software and platforms for RR (2)

Project/Experiment management systems [Jen22]

- Neptune: is a **metadata store** for any **MLOps workflow**. It was built for both research and production teams that run a lot of experiments
- Weights & Biases: is a machine learning platform built for **experiment tracking**, **dataset versioning**, and **model management**
- Comet: ML platform that helps data scientists **track**, **compare**, **explain** and **optimize experiments** and **models** across the model's entire lifecycle
- Sacred + Omniboard: **open-source** software that allows machine learning researchers to **configure**, **organize**, **log**, and **replicate** experiments

A quick review of software and platforms for RR (3)

Project/Experiment management systems

- Tensorboard: **visualization toolkit** for TensorFlow
- Polyaxon: platform for **reproducible and scalable machine learning** and deep learning applications
- ClearML: is an **open-source** platform, a suite of tools to **streamline** your **ML workflow**
- Pachyderm: enterprise-grade, **open-source data science platform** that makes it possible for its users to **control an end-to-end machine learning cycle**
- MLflow: **open-source** platform that helps manage the **whole machine learning lifecycle**. This includes experimentation, but also model storage, reproducibility, and deployment

4. Conclusion

4. Conclusions

- **Reproducible research** is needed for the **advancement of science**
- The **article**, the **software**, and **data** must be all part of the same publication
- It has a **larger impact**, but it requires a **larger effort** too
- Researchers need to be rewarded

4. Conclusions

- **Reproducible research** is needed for the **advancement of science**
- The **article**, the **software**, and **data** must be all part of the same publication
- It has a **larger impact**, but it requires a **larger effort** too
- Researchers need to be rewarded

Many initiatives on this direction:

- **European**: European Open Science Cloud (EOSC)
- **National**: French *Plan for Open Science*. **Discussion**: others?
- **Centre Borelli**: IPOL, OVD-SaaS, MVA track on reproducibility, . . . **Discussion**: what about your research centers?

Thank you for your attention

References i


David Auber, *Tulip—a huge graph visualization framework*, Graph drawing software, Springer, 2004, pp. 105–126.


Nicolas Bonneel, David Coeurjolly, Julie Digne, and Nicolas Mellado, *Code Replicability in Computer Graphics*, ACM Transactions on Graphics (Proceedings of SIGGRAPH) **39** (2020), no. 4.


G. Bradski, *The OpenCV Library*, Dr. Dobb's Journal of Software Tools (2000).


Sarah Cohen-Boulakia, Khalid Belhajjame, Olivier Collin, Jérôme Chopard, Christine Froidevaux, Alban Gaignard, Konrad Hinsen, Pierre Larmande, Yvan Le Bras, Frédéric Lemoine, et al., *Scientific workflows for computational reproducibility in the life sciences: Status, challenges and opportunities*, Future Generation Computer Systems **75** (2017), 284–298.

References ii


Miguel Colom, Bertrand Kerautret, and Adrien Krähenbühl, *An Overview of Platforms for Reproducible Research and Augmented Publications*, Reproducible Research in Pattern Recognition (Bertrand Kerautret, Miguel Colom, Daniel Lopresti, Pascal Monasse, and Hugues Talbot, eds.), Lecture Notes in Computer Science, Springer International Publishing, 2019, pp. 25–39 (en).


Roberto Di Cosmo, Morane Gruenpeter, and Stefano Zacchiroli, *Identifiers for digital objects: the case of software source code preservation*, iPRES 2018-15th International Conference on Digital Preservation, 2018, pp. 1–9.


_____, *Referencing source code artifacts: a separate concern in software citation*, Computing in Science & Engineering **22** (2019), no. 2, 33–43.


Dgtal: Digital geometry tools and algorithms library, `http://dgtal.org`.


Geogram: A programming library with geometric algorithms, `https://github.com/BrunoLevy/geogram`.

References iii


Ben J Hatchwell, *Replication in behavioural ecology: a comment on ihle et al.*, Behavioral Ecology **28** (2017), no. 2, 360–360.


Higra: Hierarchical graph analysis.


Patrycja Jenkner, *Best tools for ml experiment tracking and management*, July 2022.


Matthew McCormick, Xiaoxiao Liu, Julien Jomier, Charles Marion, and Luis Ibanez, *Itk: enabling reproducible research and open science*, Frontiers in neuroinformatics **8** (2014), 13.


Pierre Moulon, Pascal Monasse, Romuald Perrot, and Renaud Marlet, *Openmvg: Open multiple view geometry*, Reproducible Research in Pattern Recognition (Cham) (Bertrand Kerautret, Miguel Colom, and Pascal Monasse, eds.), Springer International Publishing, 2017, pp. 60–74.

References iv


Radu Bogdan Rusu and Steve Cousins, *3D is here: Point Cloud Library (PCL)*, IEEE International Conference on Robotics and Automation (ICRA) (Shanghai, China), IEEE, May 9-13 2011.


Michaël Roynard, Edwin Carlinet, and Thierry Géraud, *An image processing library in modern c++: Getting simplicity and efficiency with generic programming*, International Workshop on Reproducible Research in Pattern Recognition, Springer, 2018, pp. 121–137.


Julien Tierny, Guillaume Favelier, Joshua A. Levine, Charles Gueunet, and Michael Michaux, *The Topology ToolKit*, IEEE Transactions on Visualization and Computer Graphics (Proc. of IEEE VIS) (2017), `https://topology-tool-kit.github.io/`.


The CGAL Project, *CGAL user and reference manual*, 5.4.1 ed., CGAL Editorial Board, 2022.


David Tschumperlé, *The cimg library*, IPOL 2012 Meeting on Image Processing Libraries, 2012, pp. 4–pp.

References

Vigra: Vision with generic algorithms, `https://ukoethe.github.io/vigra` last access on May 2022.

Burak Yildiz, Hayley Hung, Jesse H Krijthe, Cynthia Liem, Marco Loog, Gosia Migut, Frans A Oliehoek, Annibale Panichella, Przemysław Pawełczak, Stjepan Picek, et al., *Reproducedpapers. org: Openly teaching and structuring machine learning reproducibility*, International Workshop on Reproducible Research in Pattern Recognition, Springer, 2021, pp. 3–11.