Implementing Reproducibility in Computational Science

Victoria Stodden
School of Information Sciences
University of Illinois at Urbana-Champaign

SIAM Annual Meeting: Invited Talk
Boston, MA

July 14, 2016

Agenda

1. Framing Notions of Reproducibility
2. A Collective Action Problem: Example Responses
3. Responses 2: Tools
4. Responses 3: Policy
5. Developing Best Practices

The Impact of Technology I

1. Big Data / Data Driven Discovery: high dimensional data, p >> n,

2. Computational Power: simulation of the complete evolution of a physical system, systematically varying parameters,

3. Deep intellectual contributions now encoded only in software.

The software contains “ideas that enable biology...”
Stories from the Supplement, 2013

The Impact of Technology II

1. **Communication**: nearly all aspects of research becoming *digitized* and *accessible* due to the Internet.
    - myriad examples.. including the Open Access movement.

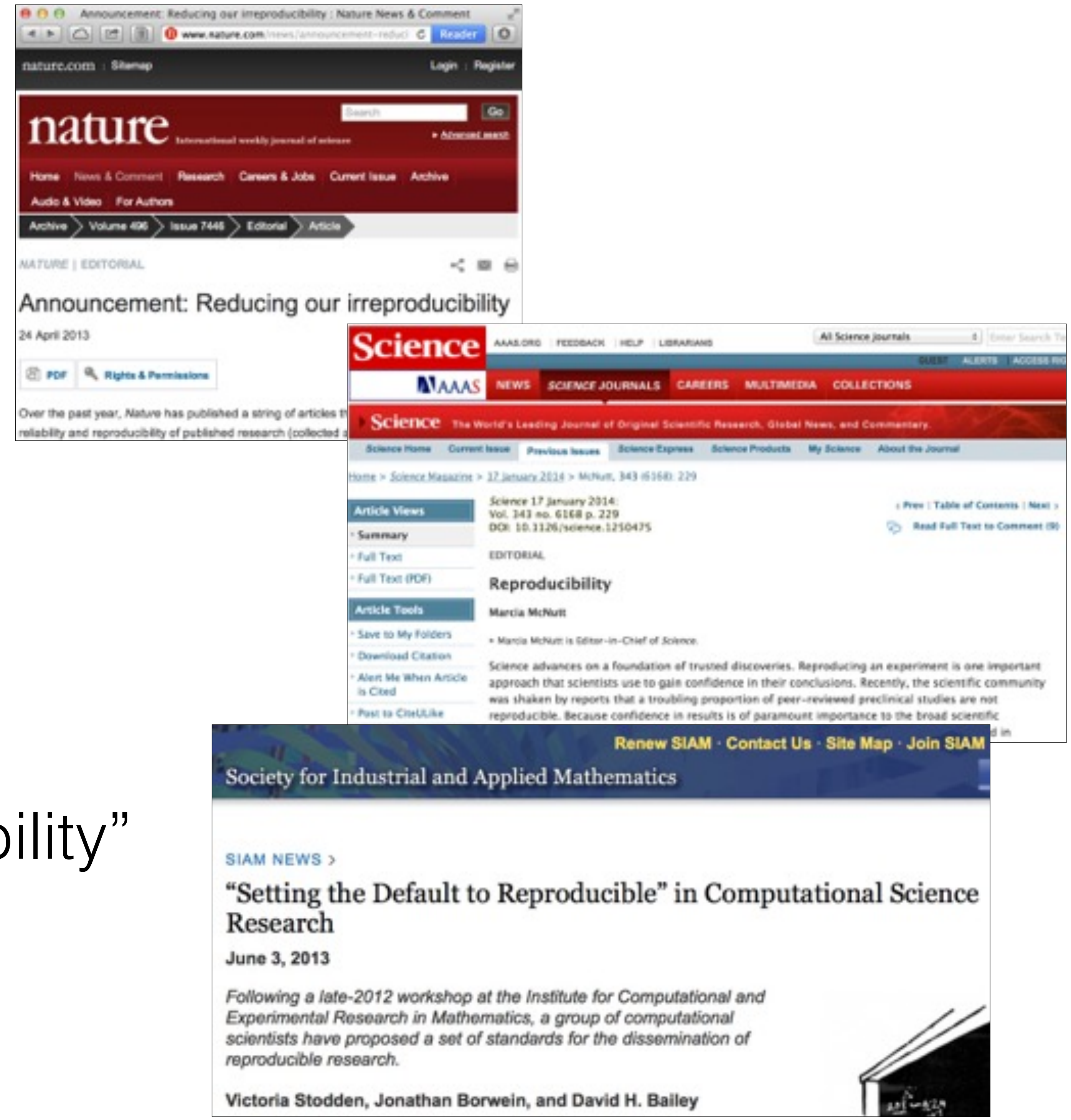
2. **Intellectual Property Law**: digitally shared objects often have more and more easily enforceable IP rights associated.
    - Reproducible Research Standard (Stodden 2009).

Parsing Reproducibility I

"Empirical Reproducibility"

"Statistical Reproducibility"

"Computational Reproducibility"



V. Stodden, IMS Bulletin (2013)

Empirical Reproducibility

Cell Reports
Commentary

Cell PRESS Open ACCESS

Sorting Out the FACS: A Devil in the Details

William C. Hines,[1,5,*] Ying Su,[2,3,4,5,*] Irene Kuhn,[1] Kornelia Polyak,[2,3,4,5] and Mina J. Bissell[1,5]
[1]Life Sciences Division, Lawrence Berkeley National Laboratory, Mailstop 977R225A, 1 Cyclotron Road, Berkeley, CA 94720, USA
[2]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02215, USA
[3]Department of Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA
[4]Department of Medicine, Harvard Medical School, Boston, MA 02115, USA
[5]These authors contributed equally to this work
*Correspondence: chines@lbl.gov (W.C.H.), ying_su@dfci.harvard.edu (Y.S.)
http://dx.doi.org/10.1016/j.celrep.2014.02.021

The reproduction of results is the cornerstone of science; yet, at times, reproducing the results of others can be a difficult challenge. Our two laboratories, one on the East and the other on the West Coast of the United States, decided to collaborate on a problem of mutual interest—namely, the heterogeneity of the human breast. Despite using seemingly identical methods, reagents, and specimens, our two laboratories quite reproducibly were unable to replicate each other's fluorescence-activated cell sorting (FACS) profiles of primary breast cells. Frustration of studying cells close to their context in vivo makes the exercise even more challenging.

Paired with in situ characterizations, FACS has emerged as the technology most suitable for distinguishing diversity among different cell populations in the mammary gland. Flow instruments have evolved from being able to detect only a few parameters to those now capable of measuring up to—and beyond—an astonishing 50 individual markers per cell (Cheung and Utz, 2011). As with any exponential increase in data complexity, breast reduction mammoplasties. Molecular analysis of separated fractions was to be performed in Boston (K.P.'s laboratory, Dana-Farber Cancer Institute, Harvard Medical School), whereas functional analysis of separated cell populations grown in 3D matrices was to take place in Berkeley (M.J.B.'s laboratory, Lawrence Berkeley National Lab, University of California, Berkeley). Both our laboratories have decades of experience and established protocols for isolating cells from primary normal breast tissues as well as the capabilities required for

NATIONAL ACADEMY OF SCIENCES | NATIONAL ACADEMY OF ENGINEERING | INSTITUTE OF MEDICINE | NATIONAL RESEARCH COUNCIL

ILAR Roundtable

Home | About | Roundtable Members | Roundtable Activities | What's New at the ILAR Roundtable

Reproducibility Issues in Research with Animals and Animal Models

The missing "R": Reproducibility in a Changing Research Landscape

A workshop of the Roundtable on Science and Welfare in Laboratory Animal Use

National Academy of Sciences, NAS 125
2100 C Street NW, Washington DC
June 4-5, 2014

The ability to reproduce an experiment is one important approach that scientists use to gain confidence in their conclusions. Studies that show that a number of significant peer-reviewed studies are not reproducible has alarmed the scientific community. Research that uses animals and animal models seems to be one of the most susceptible to reproducibility issues.

Evidence indicates that there are many factors that may be contributing to scientific irreproducibility, including insufficient reporting of details pertaining to study design and planning; inappropriate interpretation of results; and author, reviewer, and editor abstracted reporting, assessing, and accepting studies for publication.

In this workshop, speakers from around the world will explore the many facets of the issue and potential pathways to reducing the problems. Audience participation portions of the workshop are designed to facilitate understanding of the issue.

Tweet #ilar
Get updates!
Search Site

Upcoming Events
April 20-21, 2015
Design, Implementation, Monitoring and Sharing of Performance Standards

Past Events
September 3-4, 2014
Transportation of Laboratory Animals
• Presentations and videos online

June 4-5, 2014
Reproducibility Issues in Research with Animals and Animal Models
• Presentations and videos online

Statistical Reproducibility

- False discovery, p-hacking (Simonsohn 2012), file drawer problem, overuse and mis-use of p-values, lack of multiple testing adjustments.
- Low power, poor experimental design, nonrandom sampling,
- Data preparation, treatment of outliers, re-combination of datasets, insufficient reporting/tracking practices,
- inappropriate tests or models, model misspecification,
- Model robustness to parameter changes and data perturbations,
- Investigator bias toward previous findings; conflicts of interest.
- …

Computational Reproducibility

Traditionally two branches to the scientific method:

- Branch 1 (deductive): mathematics, formal logic,
- Branch 2 (empirical): statistical analysis of controlled experiments.

Now, new branches due to technological changes?

- Branch 3,4? (computational): large scale simulations / data driven computational science.

The Ubiquity of Error

The central motivation for the scientific method is to root out error:

- Deductive branch: the well-defined concept of the proof,
- Empirical branch: the machinery of hypothesis testing, appropriate statistical methods, structured communication of methods and protocols.

Claim: Computation presents only a *potential* third/fourth branch of the scientific method (Donoho et al. 2009), until the development of comparable standards.

Really Reproducible Research

“Really Reproducible Research” (1992) inspired by Stanford Professor Jon Claerbout:

> “The idea is: An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete ... set of instructions [and data] which generated the figures.” David Donoho, 1998

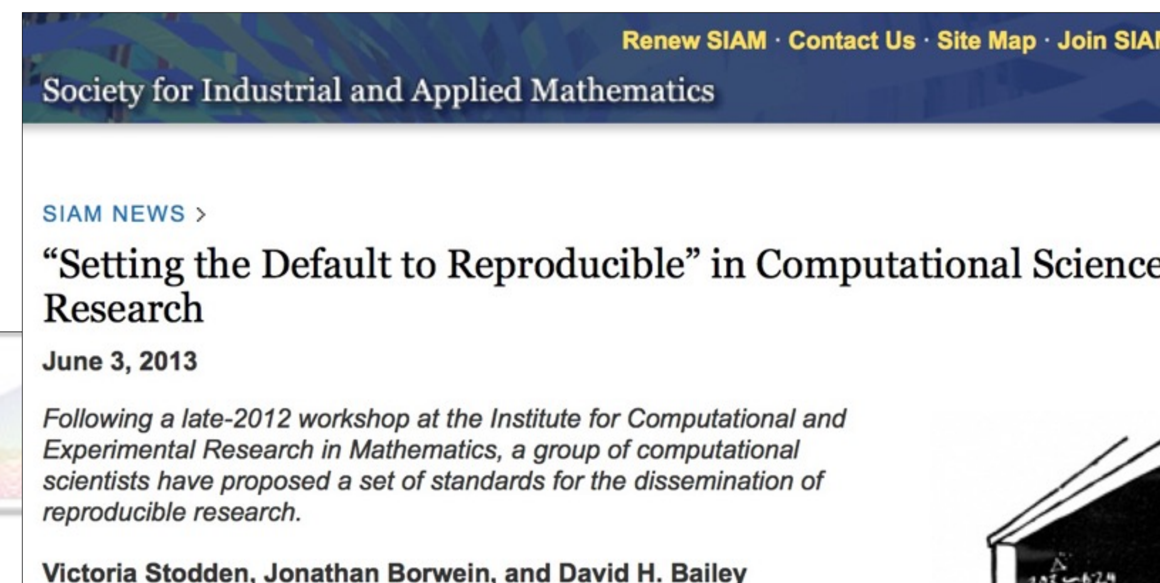
Note the difference between: reproducing the computational steps and, replicating the experiments independently including data collection and software implementation. (Both required)

Community Responses

Declarations and Documents:

- Yale Declaration 2009

- ICERM 2012

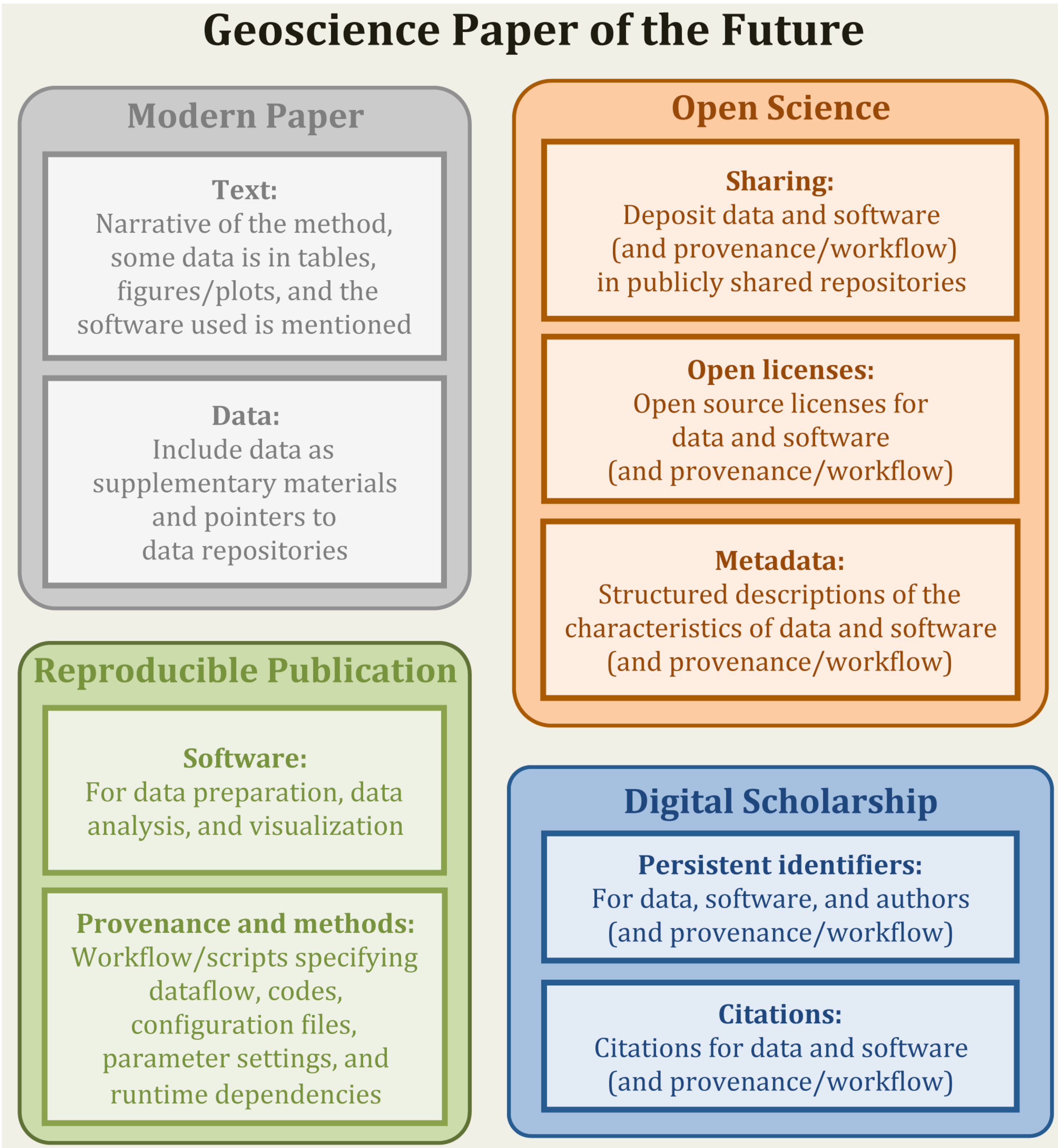


- XSEDE 2014

Defining Reproducibility

ICERM Criterion	Definition
Reviewable	The descriptions permit the research methods to be independently assessed and the results judged credible.
Confirmable	The main conclusions of the research can be attained independently without the use of software provided by the author (using the complete description of algorithms and methodology provided).
Replicable	Tools are made available that would allow one to duplicate the results of the research.
Auditable	Sufficient records (including data and software) have been archived so that the research can be defended later if necessary or differences between independent confirmations resolved. The archive might be private, as with traditional laboratory notebooks.
Reproducible	Auditable research made openly available. This comprises well-documented and fully open code and data that are publicly available that would allow one to (a) fully audit the computational procedure, (b) replicate and also independently reproduce the results of the research, and (c) extend the results or apply the method to new problems.

Collective Action Problem



Robinson, Gil, Duffy, Mattmann, Peckham 2015 ctuning.org/ae/

Artifact Evaluation for Computer Systems' Research

We work with the community and ACM to improve our methodology and tools for reproducible experimentation, artifact submission and rev

Upcoming AE

- PACT 2016
- PPoPP 2017
- CGO 2017

Recently completed Artifact Evaluation

PPoPP 2016 – see accepted artifacts here.

Highest ranked artifact: "A Wait-free Queue as Fast as Fetch-and-Add", Chaoran Yang and John Mellor-Crummey (award by dividiti)

CGO 2016 – see accepted artifacts here.

Highest ranked artifact: "A Basic Linear Algebra Compiler for Structured Matrices", Daniele G. Spampinato and Markus Puschel (award by the cTuning foundation)

Recent events

- 4 May 2016 – we've participated in the ACM workshop on reproducibiliy to unify artifact evaluation across various SIGs, and will update our artifact submission/reviewing procedures soon!
- 20 March 2016 – our Dagstuhl report on Artifact Evaluation for Publications (Bruce R. Childers, Grigori Fursin, Shriram Krishnamurthi and Andreas Zeller) is now available online.
- March 2016 – we are discussing with ACM a possibility to let the authors add up to 2 pages of their AE appendix to the camera ready paper.
- 14 March 2016 (Monday, 18:00–18:30) – we have arranged a public AE discussion (results, distinguished artifact award, issues, future work).

Tweets by @cresearch

cResearch
@cresearch
Artifact submission website for PAC
(only for accepted papers):
ctuning.org/ae/pact2016.ht... - dead
July!

cResearch Retweeted
CSAIL at MIT
@MIT_CSAIL
Which programming language sho

Embed

Sponsors and sup

cTuning Foundation

acm

arxiv.org/abs/1501.05387

Cornell University Library

We gratefully acknowledge support from the Simons Foundation and member institutions

arXiv.org > cs > arXiv:1501.05387

Search or Article-id (Help | Advanced search) All papers Go!

Computer Science > Distributed, Parallel, and Cluster Computing

Gunrock: A High-Performance Graph Processing Library on the GPU

Yangzihao Wang, Andrew Davidson, Yuechao Pan, Yuduo Wu, Andy Riffel, John D. Owens

(Submitted on 22 Jan 2015 (v1), last revised 22 Feb 2016 (this version, v6))

For large-scale graph analytics on the GPU, the irregularity of data access and control flow, and the complexity of programming GPUs have been two significant challenges for developing a programmable high-performance graph library. "Gunrock", our graph-processing system designed specifically for the GPU, uses a high-level, bulk-synchronous, data-centric abstraction focused on operations on a vertex or edge frontier. Gunrock achieves a balance between performance and expressiveness by coupling high performance GPU computing primitives and optimization strategies with a high-level programming model that allows programmers to quickly develop new graph primitives with small code size and minimal GPU programming knowledge. We evaluate Gunrock on five key graph primitives and show that Gunrock has on average at least an order of magnitude speedup over Boost and PowerGraph, comparable performance to the fastest GPU hardwired primitives, and better performance than any other GPU high-level graph library.

Comments: 14 pages, accepted by PPoPP'16 (removed the text repetition in the previous version v5)

Subjects: **Distributed, Parallel, and Cluster Computing (cs.DC)**

ACM classes: D.1.3

DOI: 10.1145/2851141.2851145

Cite as: arXiv:1501.05387 [cs.DC]

(or arXiv:1501.05387v6 [cs.DC] for this version)

Download:

- PDF
- Other formats

Current browse context:

cs.DC

< prev | next >

new | recent | 1501

Change to browse by:

cs

References & Citations

- NASA ADS

DBLP - CS Bibliography

listing | bibtex

Yangzihao Wang
Andrew A. Davidson
Yuechao Pan
Yuduo Wu
Andy Riffel
...

Bookmark (what is this?)

A. Artifact description

A.1 Abstract

The artifact contains all the executables of the current existing graph primitives in Gunrock's latest version on github, as well as the shell scripts of running them. It can support the runtime and/or edge throughput results in Table 3 of our PPoPP'2016 paper ***Gunrock: A High-Performance Graph Processing Library on the GPU****. To validate the results, run the test scripts and check the results piped in the according text output files.*

A.2 Description

A.2.1 Check-list (artifact meta information)

- **Algorithm: breadth-first search, single-source shortest path, betweenness centrality, Pagerank, connected component**
- **Program: CUDA and C/C++ code**
- **Compilation: Host code: gcc 4.8.4 with the -O3 flag; device code: nvcc 7.0.27 with the -O3 flag**
- **Binary: CUDA executables**
- **Data set: Publicly available matrix market files**
- **Run-time environment: Ubuntu 12.04 with CUDA and GPU Computing SDK installed**
- **Hardware: Any GPU with compute capability $\geq$ 3.0 (Recommended GPU: NVIDIA K40c GPU)**
- **Output: Runtime and/or edge throughput**
- **Experiment workflow: Git clone project; download the datasets; run the test scripts; observe the results**
- **Publicly available?: Yes**

A.2.2 How delivered

Gunrock is an open source library under Apache 2.0 license and is hosted with code, API specifications, build instructions, and design documentations on Github.

A.2.3 Hardware dependencies

Gunrock requires NVIDIA GPU with the compute capability of no less than 3.0.

A.2.4 Software dependencies

Gunrock requires Boost (for CPU reference) and CUDA with version no less than 5.5. Gunrock has been tested on Ubuntu 12.04/14.04, and is expected to run correctly under other Linux distributions.

A.2.5 Datasets

All datasets are either publicly available or generated using standard graph generation software. Users will be able to run script to get these datasets once they built Gunrock code. The rgg graph is generated by Gunrock team. The download link is provided here: `https://drive.google.com/uc?export=download&id=0Bw6LwCuER0a3VWNrVUV6eTZyeFU`. Please located the unzipped `rgg_n_2_24_s0.mtx` file under `gunrock_dir/datasets/large/rgg_n_2_24_s0/`. Users are welcom to try other datasets or generate rgg/R-MAT graphs using the command line option during the test. We currently only support matrix market format files as input.

A.3 Installation

Follow the build instruction on Gunrock's github page (`http://gunrock.github.io/`), users can build Gunrock and generate the necessary executables for the experiments.

A.4 Experiment workflow

For the convenience of the artifact evaluation, we provide a series of shell scripts which run the graph primitives we have described in the paper and store the results in the output text files. Below are the steps to download Gunrock code, build, run the experiments, and observe the results.

- Clone Gunrock code to the local machine:

```
$ git clone https://github.com/gunrock/gunrock.git
$ cd gunrock
$ git submodule init && git submodule update
```

- Use CMake to build Gunrock. Make sure that boost and CUDA is correctly installed before this step:

```
$ cd /path/to/gunrock/../
$ mkdir gunrock_build && cd gunrock_build
$ cmake ../gunrock/
$ make -j16
```

  The last comand will build Gunrock's executables under `gunrock_build/bin` and shared library under `gunrock_build/lib`.

- Prepare the dataset. First step into Gunrock directory:

```
$ cd /path/to/gunrock/
$ cd dataset/large/ && make
```

  This will download and extract all the large datasets, including the 6 datasets in the paper.

- Step into the test script directory and run scripts for five graph primitives:

```
$ cd ../test-scripts
$ sh ppopp16-test.sh
```

- Observe the results for each dataset under five directories: BFS, SSSP, BC, PR, and CC.

A.5 Evaluation and expected result

For BFS and SSSP, the expected results include both runtime and edge throughput. For BC, Pagerank, and CC, the expected results contain runtime only.

A.6 Notes

To know more about our library, send feedback, or file issues, please visit our github page (`http://gunrock.github.io/`).

arxiv.org/abs/1110.3649v1

Algorithms to automatically quantify the geometric similarity of anatomical surfaces

D. Boyer, Y. Lipman, E. St. Clair, J. Puente, T. Funkhouser, B. Patel, J. Jernvall, I. Daubechies

(Submitted on 17 Oct 2011 (this version), latest version 15 Mar 2012 (v3))

We describe new approaches for distances between pairs of 2-dimensional surfaces (embedded in 3-dimensional space) that use local structures and global information contained in inter-structure geometric relationships. We present algorithms to automatically determine these distances as well as geometric correspondences. This is motivated by the aspiration of students of natural science to understand the continuity of form that unites the diversity of life. At present, scientists using physical traits to study evolutionary relationships among living and extinct animals analyze data extracted from carefully defined anatomical correspondence points (landmarks). Identifying and recording these landmarks is time consuming and can be done accurately only by trained morphologists. This renders these studies inaccessible to non-morphologists, and causes phenomics to lag behind genomics in elucidating evolutionary patterns. Unlike other algorithms presented for morphological correspondences our approach does not require any preliminary marking of special features or landmarks by the user. It also differs from other seminal work in computational geometry in that our algorithms are polynomial in nature and thus faster, making pairwise comparisons feasible for significantly larger numbers of digitized surfaces. We illustrate our approach using three datasets representing teeth and different bones of primates and humans, and show that it leads to highly accurate results.

Subjects: **Numerical Analysis (math.NA)**; Computer Vision and Pattern Recognition (cs.CV); Graphics (cs.GR)

Cite as: arXiv:1110.3649 **[math.NA]**

(or arXiv:1110.3649v1 **[math.NA]** for this version)

Infrastructure Responses

Tools and software to enhance reproducibility and disseminate the scholarly record:

Dissemination Platforms

ResearchCompendia.org	IPOL	Madagascar
MLOSS.org	thedatahub.org	nanoHUB.org
Open Science Framework		RunMyCode.org

Workflow Tracking and Research Environments

Vistrails	Kepler	CDE	Jupyter	torch.ch
Galaxy	GenePattern	Sumatra	Taverna	DataCenterHub
Pegasus	Kurator			RCloud

Embedded Publishing

Verifiable Computational Research	SOLE	knitR
Collage Authoring Environment	SHARE	Sweave

grants.nih.gov/reproducibility/index.htm

NIH National Institutes of Health
Office of Extramural Research

Grants & Funding
NIH's Central Resource for Grants and Funding Information

Entire Site | Search this Site

eRA | Glossary & Acronyms | FAQs | Help

HOME | ABOUT GRANTS | FUNDING | POLICY & COMPLIANCE | NEWS & EVENTS | ABOUT OER

Home » Policy & Compliance » Rigor and Reproducibility

NIH Grants Policy Statement

Notices of Policy Changes

Compliance & Oversight

Select Policy Topics

+

Rigor and Reproducibility

Scientific rigor and transparency in conducting biomedical research is key to the successful application of knowledge toward improving health outcomes. The information provided on this website is designed to assist the extramural community in addressing rigor and transparency in NIH grant applications and progress reports.

On This Page:

- Goals
- Guidance: Rigor and Reproducibility in Grant Applications
- Resources
- News
- References

Goals

The NIH strives to exemplify and promote the highest level of scientific integrity, public accountability, and social responsibility in the conduct of science. Updates to grant applications instructions and review language are intended to:

Related Resources

FAQs

ORWH Studying Sex to Strengthen Science (S4)

NIH Rigor and Reproducibility

NIGMS Training Modules

Intranet Resources on Rigor and Transparency (NIH Staff Only)

Contact:
reproducibility@nih.gov

03/01/2013

SIAM Journals Introduce Supplementary Materials

Recently an ad hoc Committee on Supplementary Materials* formulated a number of recommendations for expanding the online options associated with papers published in SIAM journals. Some of these recommendations have now been implemented, and the options are available to authors, starting with the following journals: *SIAM Review*, *SIAM Journal on Matrix Analysis and Applications*, *SIAM Journal on Scientific Computing*, *SIAM Journal on Mathematical Analysis*, and *SIAM Journal on Numerical Analysis*. Other journals will soon join this group.

Supplementary Materials, which will be posted online with a link from the webpage for the paper, will consist of un-refereed materials that the author(s), referees, and editor agree are appropriate to accompany the publication. These might include animations of results shown in the paper, additional figures or examples that may be useful to the reader in understanding the paper, or computer code or data sets that were used in generating figures or tables in the paper. Archiving these as Supplementary Materials to accompany the paper will insure that they are available to readers of the journal at a stable URL, and can be cited using the DOI of the article. The refereed portion of the paper should stand on its own as the official publication, while the Supplementary Materials are intended to complement the paper.

All Supplementary Materials must be submitted along with the manuscript, accompanied by an index that lists each attachment and a justification for including it. SIAM submission forms for the journals affected have been redesigned to allow this; details on preparing and submitting this content can be found on the journal-specific Supplementary Materials webpages at http://www.siam.org/journals/.

Referees will be asked to give these materials at least a cursory look to insure that they are appropriate as material associated with the paper. Beyond this, Supplementary Materials are generally not refereed, but the referees or editor may suggest changes, including removing some extraneous Supplementary Materials or moving nonessential items from the main text to the Supplementary Materials.

By identifying a broad range of Supplementary Materials, we hope to encourage authors to submit data or computer code that is a critical component of the scholarship contained in the paper. This will go beyond aiding the reader who wants to understand the details of the work presented. Many funding agencies now require that data and/or computer code associated with published research results be made publicly available. The availability of electronic archives for material associated with SIAM publications may assist our authors in complying with such requirements.

On a related topic, authors are also encouraged to use appendices for traditional printed material that should be refereed and published along with the paper, but that need not be part of the main flow. Appendices will continue to be handled as in the past and will appear as part of the paper. We believe that increased use of appendices, together with the capability of attaching Supplementary Materials, will help authors streamline papers for readability while still including all the necessary components to fully

AMSTATNEWS

The Membership Magazine of the American Statistical Association

Reproducible Research in *JASA*

Montse Fuentes, Coordinating Editor of JASA and Editor of JASA ACS

Societal impact through scientific advances is predicated on discovery and new knowledge that is reliable and robust and provides a solid foundation on which further advances can be built. Unfortunately, there is evidence many published scientific results will not stand the test of time, in part due to the lack of good scientific practices for reproducibility.

Our statistical profession has a responsibility to establish publication standards that improve the transparency and robustness of what we publish and to promote awareness within the scientific community of the need for rigor in our statistical research to ensure reproducibility of our scientific results. *JASA* is committed to helping lead the effort by presenting solutions that can help improve research quality and reproducibility.

Starting September 1, *JASA ACS* will require code and data as a minimum standard for reproducibility of statistical scientific research. New infrastructure is being established to support this initiative. Each manuscript will go through the current review process managed by an associate editor (AE), who will assign to one of the reviewers the broad evaluation of the code. A new editorial role—associate editor for reproducibility (AER)—will be added to ensure we meet a standard of reproducibility.

> Reproducibility of scientific research is our ultimate goal, and the code and data requirement is a first step in that direction.

sc16.supercomputing.org/studentssc/student-cluster-competition/call-for-scc-application-proposal/

SC16 Salt Lake City, Utah | hpc matters.

Sponsors: IEEE computer society, acm, sighpc

Search this website ... Search

Exhibition: **November 14-17, 2016** | Salt Palace Convention Center
Conference: **November 13-18, 2016** | Salt Lake City, Utah

Home | Attendees | Submitters | Conference Components | Exhibitors | Students@SC | Media | SCinet | Diversity

Quick Links

- Submit Your Work
- Sample Submission Forms
- Important Deadlines
- Conference Schedule
- Support Students@SC

Call for SCC Application Proposal

Important Dates:

February 15, 2016: **Submissions open for Student Cluster Competition Team Proposals**

April 15, 2016 Deadline Extended - April 22, 2016: **Submissions close for Student Cluster Competition Team Proposals**

May 13, 2016: **Student Cluster Competition Team Invitation Notification**

As part of a major initiative that aims to increase the level of reproducibility and replicability of results, SC16 invites authors of technical papers submitted to the conference to volunteer to publish their methodology, code and data with the paper, if their paper is accepted to SC16. If you want your paper to be considered for this initiative, make sure you check the box in the Linklings form when you submit your final manuscript. For information on the Technical Program CFP click here: http://sc16.supercomputing.org/submitters/technical-papers/

As a benchmark of success the SC17 Student Cluster Completion plans to select one or more of these papers for reproduction. If the paper is selected for reproduction, the authors must be willing to assist the student cluster organizers by answering questions throughout the planning phase of the competition. In addition, one of the paper authors must agree to serve as the application expert for the Student Cluster Competition at SC17.

Result and Artifact Review and Badging

An experimental result is not fully established unless it can be independently reproduced. A variety of recent studies, primarily in the biomedical field, have revealed that an uncomfortably large number of research results found in the literature fail this test, because of sloppy experimental methods, flawed statistical analyses, or in rare cases, fraud. Publishers can promote the integrity of the research ecosystem by developing review processes that increase the likelihood that results can be independently replicated and reproduced. An extreme approach would be to require completely independent reproduction of results as part of the refereeing process. An intermediate approach is to require that artifacts associated with the work undergo a formal audit. By "artifact" we mean a digital object that was either created by the authors to be used as part of the study or generated by the experiment itself. For example, artifacts can be software systems, scripts used to run experiments, input datasets, raw data collected in the experiment, or scripts used to analyze results.

Additional benefits ensue if the research artifacts are themselves made publically available so that any interested party may audit them. This also enables replication experiments to be performed, which, because they inevitably are done under slightly different conditions, serve to verify the robustness of the original results. And perhaps more importantly, well-formed and documented artifacts allow others to build directly upon the previous work through reuse and repurposing.

A number of ACM conferences and journals have already instituted formal processes for artifact review. Here we provide terminology and standards for review processes of these types in order to promote a base level of uniformity which will enable labeling of successfully reviewed papers across ACM publications choosing to adopt such practices.

Of course, there remain many circumstances in which such enhanced review will be either infeasible or not possible. As a result, such review processes are encouraged, but remain completely optional for ACM journals and conferences, and when they are made available, it is recommended that participation by authors also be made optional. Authors who do agree to such additional review, and whose work meets established standards, will be rewarded with appropriate labeling both in the text of the article and in the metadata displayed in the ACM Digital Library. Specific labels, or badges, are proposed below.

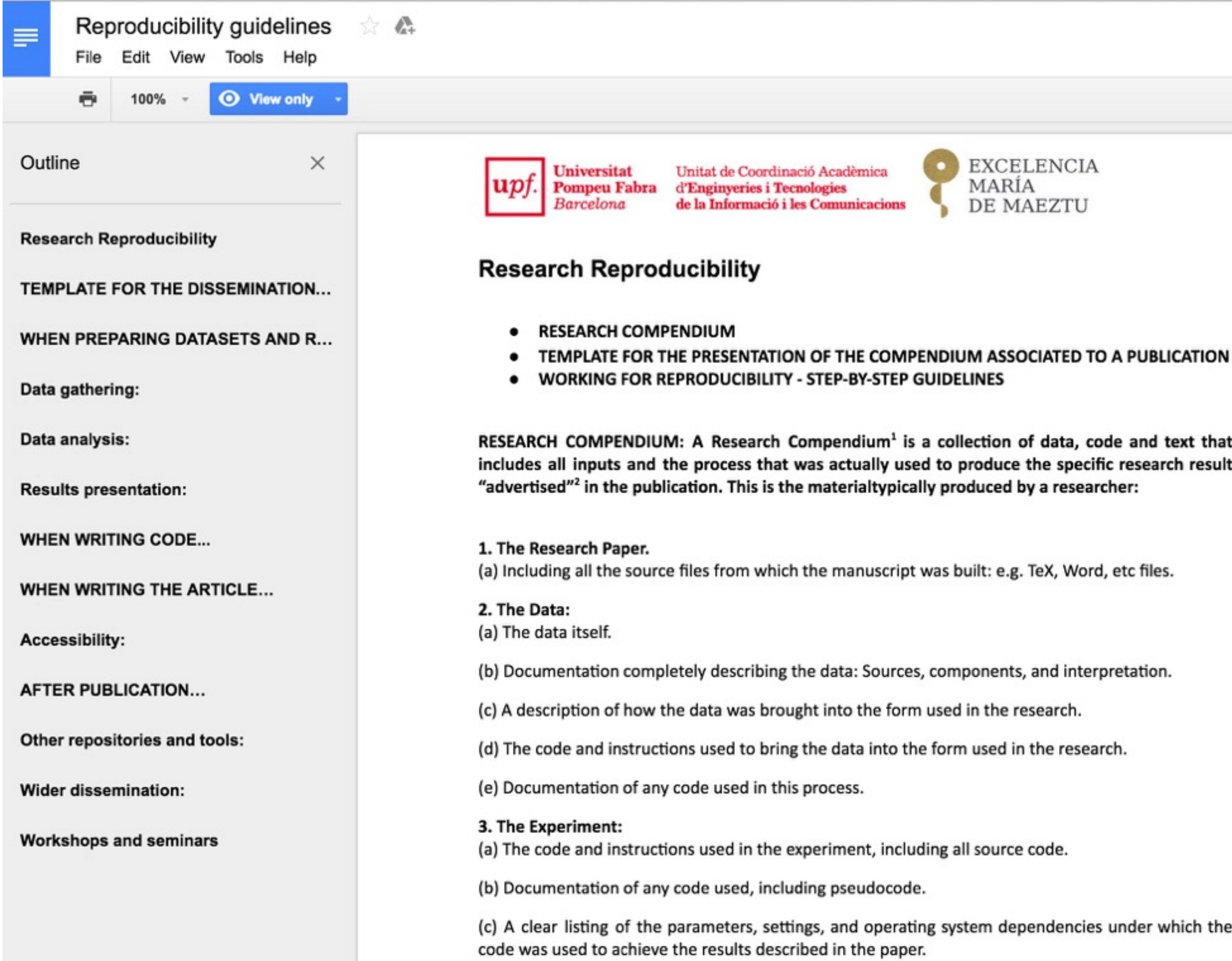
DTIC-MdM Strategic Program: Data-Driven Knowledge Extraction

Reproducible Research

We try to make and promote reproducible research, with the objective that publications are available including software and data online.

Other actions to promote it:

- **Seminars:**
  - At the Annual Event 2016 (June 28-29, free attendance), Reproducibility will be the subject of one of the keynote talks, by Victoria Stodden (University of Illinois at Urbana - Champaign)
  - DTIC Seminars
    - 19 May 2016: Malcolm Bain. Software licensing – from basic to advanced licensing and business models
    - 12 May 2016: Aurelio Ruiz. Reproducibility in Research.
- **María de Maeztu Reproducibility Award** - PhD workshop 2016
  - Additional details on this news
- **Award for Reproducibility in Software** - Best ICT Bachelor's Thesis in Spain 2016
- **Data management:** The UPF Library supports you in several aspects linked to Data Management, including the possibility to use the UPF repository to preserve your data. Check here for more details.



Reproducibility guidelines

File Edit View Tools Help

100% View only

Outline

- Research Reproducibility
- TEMPLATE FOR THE DISSEMINATION...
- WHEN PREPARING DATASETS AND R...
- Data gathering:
- Data analysis:
- Results presentation:
- WHEN WRITING CODE...
- WHEN WRITING THE ARTICLE...
- Accessibility:
- AFTER PUBLICATION...
- Other repositories and tools:
- Wider dissemination:
- Workshops and seminars

upf. Universitat Pompeu Fabra *Barcelona*

Unitat de Coordinació Acadèmica d'Enginyeries i Tecnologies de la Informació i les Comunicacions

EXCELENCIA MARÍA DE MAEZTU

Research Reproducibility

- **RESEARCH COMPENDIUM**
- **TEMPLATE FOR THE PRESENTATION OF THE COMPENDIUM ASSOCIATED TO A PUBLICATION**
- **WORKING FOR REPRODUCIBILITY - STEP-BY-STEP GUIDELINES**

RESEARCH COMPENDIUM: A Research Compendium[1] is a collection of data, code and text that includes all inputs and the process that was actually used to produce the specific research result "advertised"[2] in the publication. This is the materialtypically produced by a researcher:

1. The Research Paper.
(a) Including all the source files from which the manuscript was built: e.g. TeX, Word, etc files.

2. The Data:
(a) The data itself.

(b) Documentation completely describing the data: Sources, components, and interpretation.

(c) A description of how the data was brought into the form used in the research.

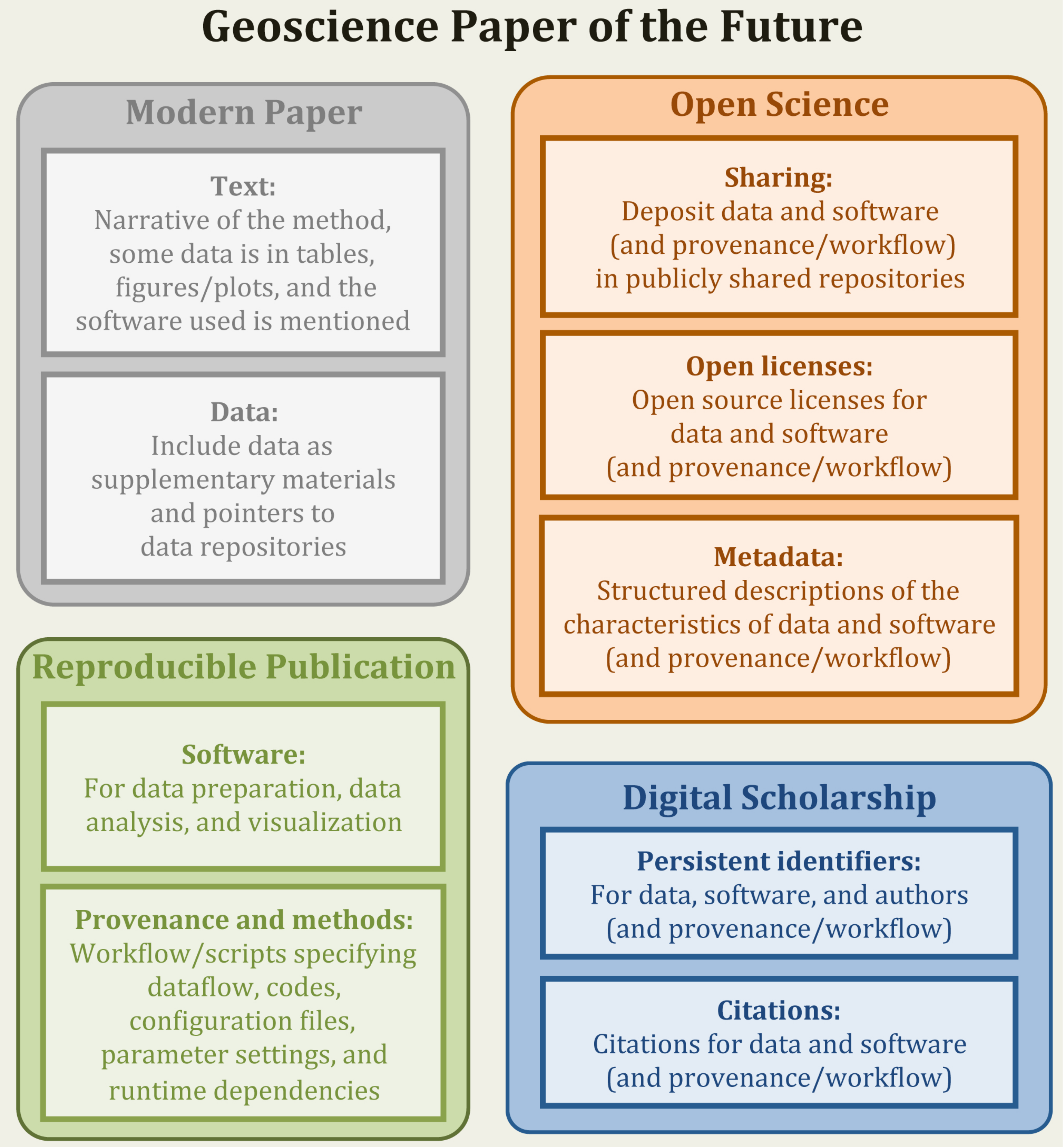
(d) The code and instructions used to bring the data into the form used in the research.

(e) Documentation of any code used in this process.

3. The Experiment:
(a) The code and instructions used in the experiment, including all source code.

(b) Documentation of any code used, including pseudocode.

(c) A clear listing of the parameters, settings, and operating system dependencies under which the code was used to achieve the results described in the paper.


Geoscience Paper of the Future
Modern Paper
Text:
Narrative of the method, some data is in tables, figures/plots, and the software used is mentioned
Data:
Include data as supplementary materials and pointers to data repositories
Open Science
Sharing:
Deposit data and software (and provenance/workflow) in publicly shared repositories
Open licenses:
Open source licenses for data and software (and provenance/workflow)
Metadata:
Structured descriptions of the characteristics of data and software (and provenance/workflow)
Reproducible Publication
Software:
For data preparation, data analysis, and visualization
Provenance and methods:
Workflow/scripts specifying dataflow, codes, configuration files, parameter settings, and runtime dependencies
Digital Scholarship
Persistent identifiers:
For data, software, and authors (and provenance/workflow)
Citations:
Citations for data and software (and provenance/workflow)

Background: Open Source Software

- Innovation: Open Licensing
    - ➡ Software with licenses that communicate alternative terms of use to code developers, rather than the copyright default.

- Hundreds of open source software licenses:
    - GNU Public License (GPL)
    - (Modified) BSD License
    - MIT License
    - Apache 2.0 License
    - ... see http://www.opensource.org/licenses/alphabetical

The Reproducible Research Standard

The *Reproducible Research Standard* (*RRS*) (Stodden, 2009)

- A suite of license recommendations for computational science:
- Release media components (text, figures) under CC BY,
- Release code components under Modified BSD or similar,
- Release data to public domain or attach attribution license.

➡ Remove copyright's barrier to reproducible research and,

➡ Realign the IP framework with longstanding scientific norms.

Journal of
open research software

Stodden, V and Miguez, S 2014 Best Practices for Computational Science: Software Infrastructure and Environments for Reproducible and Extensible Research. *Journal of Open Research Software*, 2(1): e21, pp. 1-6, DOI: http://dx.doi.org/10.5334/jors.ay

ISSUES IN RESEARCH SOFTWARE

Best Practices for Computational Science: Software Infrastructure and Environments for Reproducible and Extensible Research

Victoria Stodden* and Sheila Miguez*

The goal of this article is to coalesce a discussion around best practices for scholarly research that utilizes computational methods, by providing a formalized set of best practice recommendations to guide computational scientists and other stakeholders wishing to disseminate reproducible research, facilitate innovation by enabling data and code re-use, and enable broader communication of the output of computational scientific research. Scholarly dissemination and communication standards are changing to reflect the increasingly computational nature of scholarly research, primarily to include the sharing of the data and code associated with published results. We also present these Best Practices as a living, evolving, and changing document at http://wiki.stodden.net/Best_Practices.

Keywords: best practices; reproducible research; archiving; data sharing; code sharing; wiki; open science; computational science; scientific method

Introduction

The goal of this article is to coalesce a discussion around best practices for scholarly research that utilizes computational methods, by providing a formalized set of best

IEEE Computing in Science and Engineering focused on Reproducible Research [5] and called for "changing the culture" of scientific research [6]. A Roundtable at Yale Law School in 2009 focused on the issue of reproducibil-

Best Practice Principles

1. Open licensing should be used for data and code.
2. Workflow tracking should be carried out during the research process.
3. Data must be available and accessible.
    - Version Control for Data
    - Raw Data Availability
    - Data Types: small static files to large dynamic databases
4. Code and methods must be available and accessible
    - Version Control for Code / Making the Code Available Externally
    - Version Control for Environments / Making Environments Available and Documented
    - Code Samples and Test Data
    - “Really Big” Codebases
5. All 3rd party data and software should be cited.
6. Comply with funding agency and institutional requirements.

REPRODUCIBLE RESEARCH FOR SCIENTIFIC COMPUTING

Code Sharing Is Associated with Research Impact in Image Processing

In computational sciences such as image processing, publishing usually isn't enough to allow other researchers to verify results. Often, supplementary materials such as source code and measurement data are required. Yet most researchers choose not to make their code available because of the extra time required to prepare it. Are such efforts actually worthwhile, though?

How often have you attempted to implement and reproduce the results of another person's published paper? And when doing so, was this a straightforward process, similar to following a cookbook recipe, or rather a lengthy and painful

are imposed. Because of time pressure, we researchers often even forget to note the precise settings by which we obtained a figure's nice results. This makes it (almost) impossible, even for us as authors, to repeat the same experiments with the same results a year after the paper was written.

Vandewalle, 2012

Conclusion

Many steps toward resolving the multi-faceted and challenging problem of reproducibility in computationally enabled research.

No one size fits all solution.

Differential impact: junior vs tenured researchers.

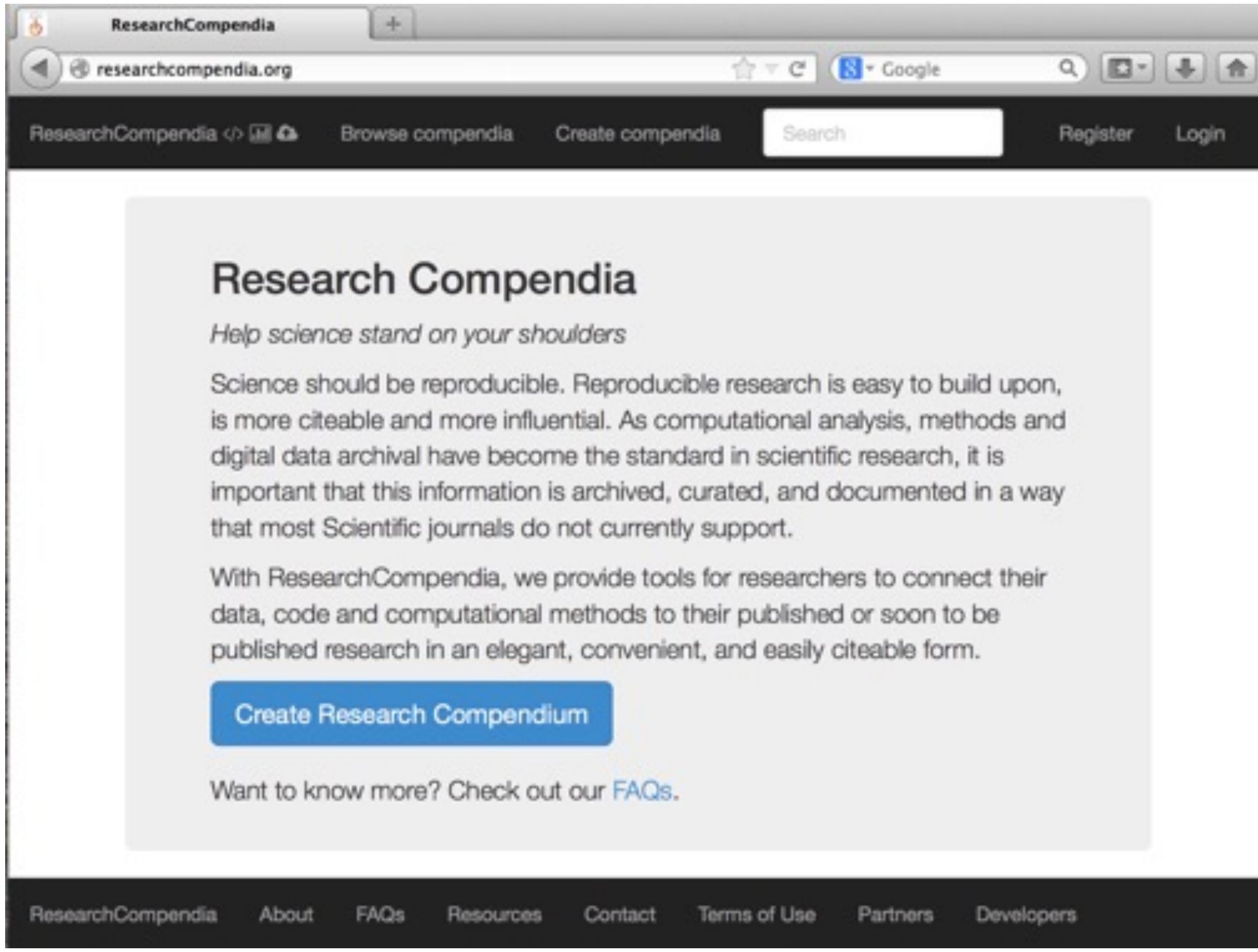
Recommendation 1: take some (more!) steps!

Recommendation 2: Develop a long term research agenda to understand the creation, dissemination, and use of really reproducible research.

Research Compendia

Pilot project: improve understanding of reproducible computational science, trace sources of error

- link data/code to published claims, re-use,
- a guide to empirical researchers,
- certifies results,
- large scale validation of findings,
- stability, sensitivity checks.



Is “Huh?” a Universal Word? Conversational Infrastructure and the Convergent Evolution of Linguistic Items

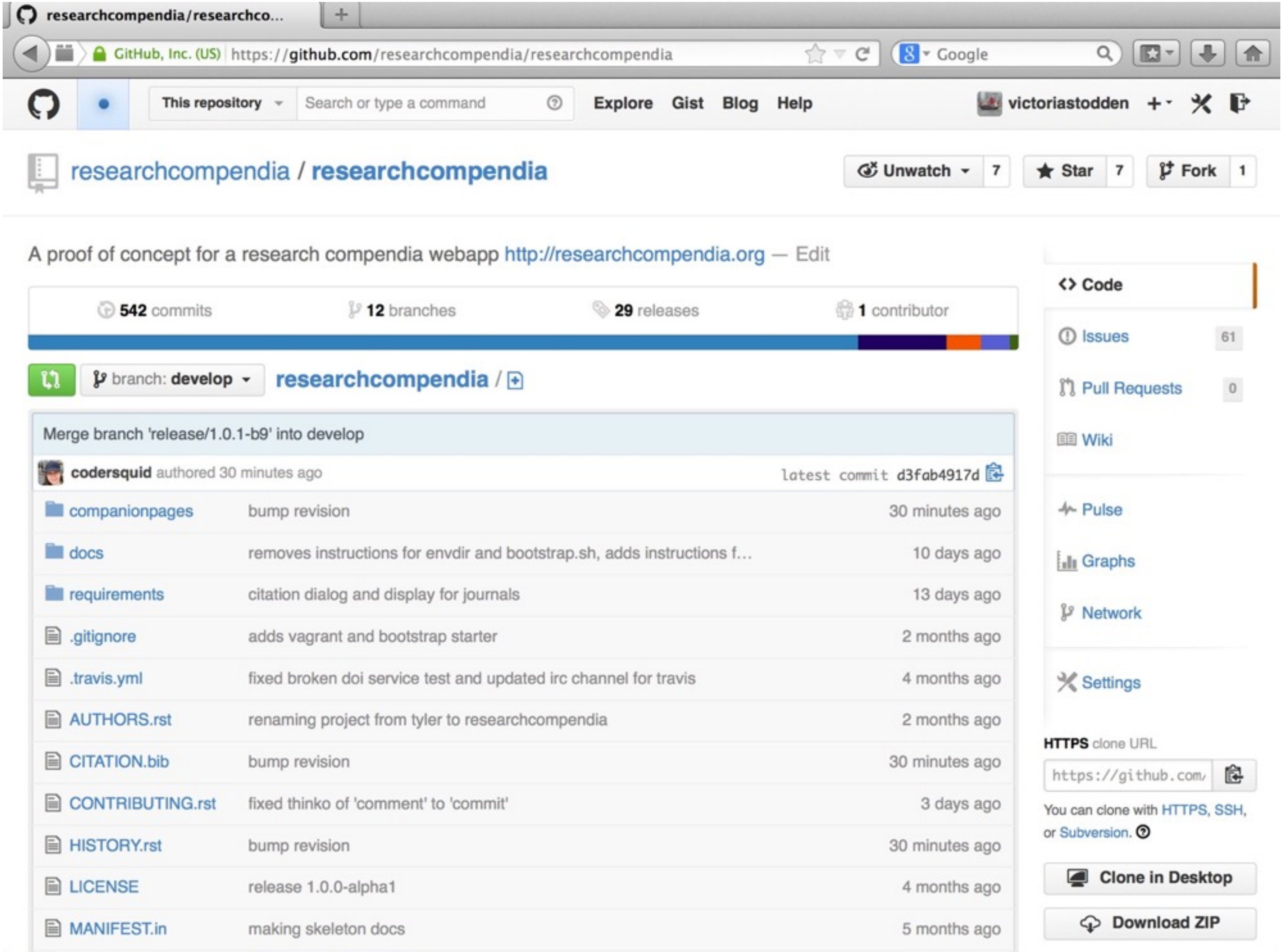
Mark Dingemanse, Francisco Torreira, N. J. Enfield, Johan J. Bolhuis

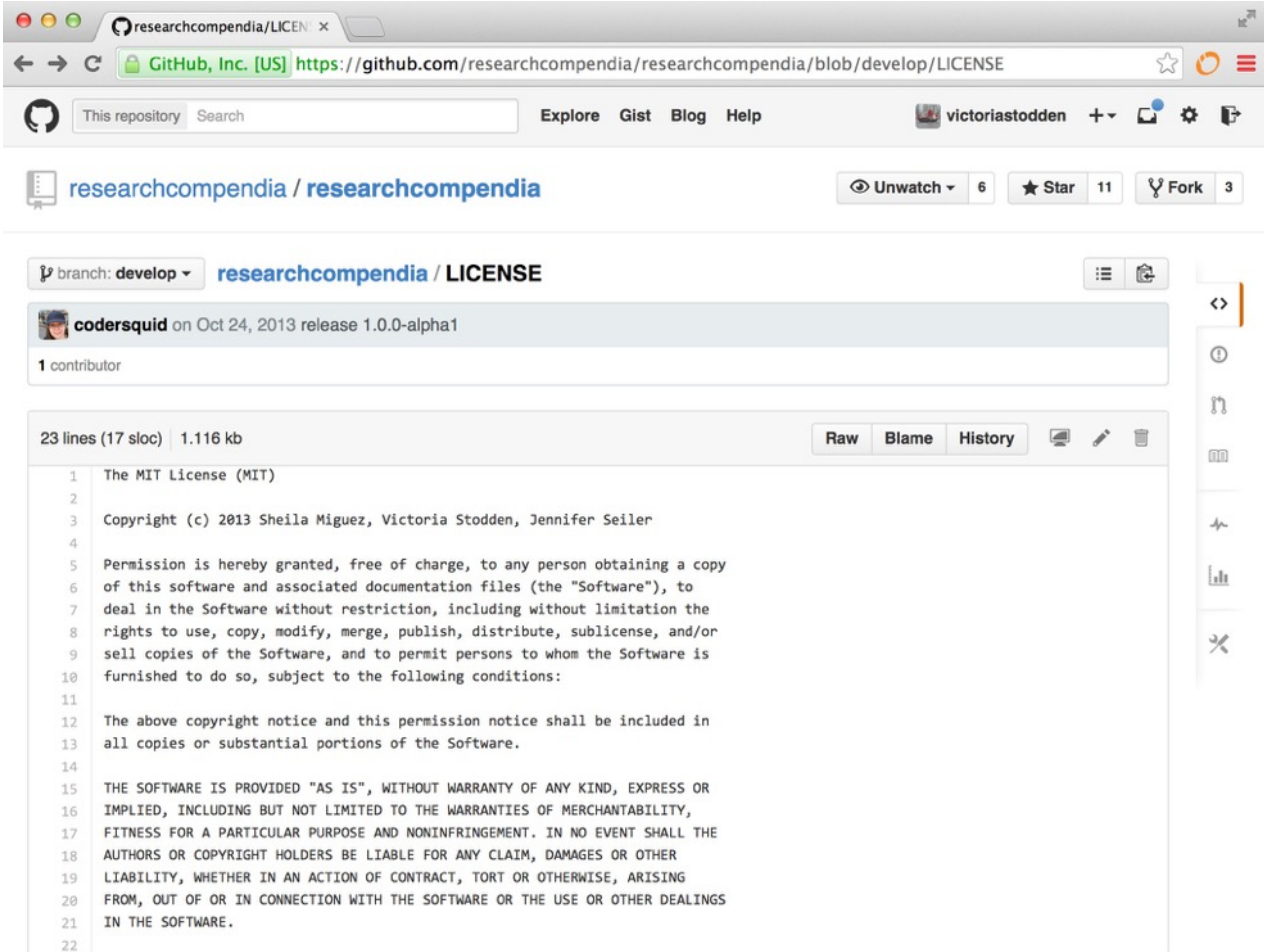
Code and Data Abstract

A word like Huh?–used as a repair initiator when, for example, one has not clearly heard what someone just said–is found in roughly the same form and function in spoken languages across the globe. We investigate it in naturally occurring conversations in ten languages and present evidence and arguments for two distinct claims: that Huh? is universal, and that it is a word. In support of the first, we show that the similarities in form and function of this interjection across languages are much greater than expected by chance. In support of the second claim we show that it is a lexical, conventionalised form that has to be learnt, unlike grunts or emotional cries. We discuss possible reasons for the cross-linguistic similarity and propose an account in terms of convergent evolution. Huh? is a universal word not because it is innate but because it is shaped by selective pressures in an interactional environment that all languages share: that of other-initiated repair. Our proposal enhances evolutionary models of language change by suggesting that conversational infrastructure can drive the convergent cultural evolution of linguistic items.

</> code | data | article

Compendium Type:: article
Content License:: CC0
Code License:: MIT


researchcompendia/researchco...
GitHub, Inc. (US) https://github.com/researchcompendia/researchcompendia
Google
This repository
Search or type a command
Explore
Gist
Blog
Help
victoriastodden
researchcompendia / researchcompendia
Unwatch 7
Star 7
Fork 1
A proof of concept for a research compendia webapp http://researchcompendia.org — Edit
542 commits
12 branches
29 releases
1 contributor
branch: develop
researchcompendia /
Merge branch 'release/1.0.1-b9' into develop
codersquid authored 30 minutes ago
latest commit d3fab4917d
companionpages bump revision 30 minutes ago
docs removes instructions for envdir and bootstrap.sh, adds instructions f… 10 days ago
requirements citation dialog and display for journals 13 days ago
.gitignore adds vagrant and bootstrap starter 2 months ago
.travis.yml fixed broken doi service test and updated irc channel for travis 4 months ago
AUTHORS.rst renaming project from tyler to researchcompendia 2 months ago
CITATION.bib bump revision 30 minutes ago
CONTRIBUTING.rst fixed thinko of 'comment' to 'commit' 3 days ago
HISTORY.rst bump revision 30 minutes ago
LICENSE release 1.0.0-alpha1 4 months ago
MANIFEST.in making skeleton docs 5 months ago
Code
Issues 61
Pull Requests 0
Wiki
Pulse
Graphs
Network
Settings
HTTPS clone URL
https://github.com/
You can clone with HTTPS, SSH, or Subversion.
Clone in Desktop
Download ZIP


researchcompendia/LICEN
GitHub, Inc. [US] https://github.com/researchcompendia/researchcompendia/blob/develop/LICENSE
This repository
Search
Explore
Gist
Blog
Help
victoriastodden
researchcompendia / researchcompendia
Unwatch 6
Star 11
Fork 3
branch: develop
researchcompendia / LICENSE
codersquid on Oct 24, 2013 release 1.0.0-alpha1
1 contributor
23 lines (17 sloc) 1.116 kb
Raw
Blame
History
The MIT License (MIT)
Copyright (c) 2013 Sheila Miguez, Victoria Stodden, Jennifer Seiler
Permission is hereby granted, free of charge, to any person obtaining a copy
of this software and associated documentation files (the "Software"), to
deal in the Software without restriction, including without limitation the
rights to use, copy, modify, merge, publish, distribute, sublicense, and/or
sell copies of the Software, and to permit persons to whom the Software is
furnished to do so, subject to the following conditions:
The above copyright notice and this permission notice shall be included in
all copies or substantial portions of the Software.
THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR
IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY,
FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE
AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER
LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING
FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS
IN THE SOFTWARE.

Querying the Scholarly Record

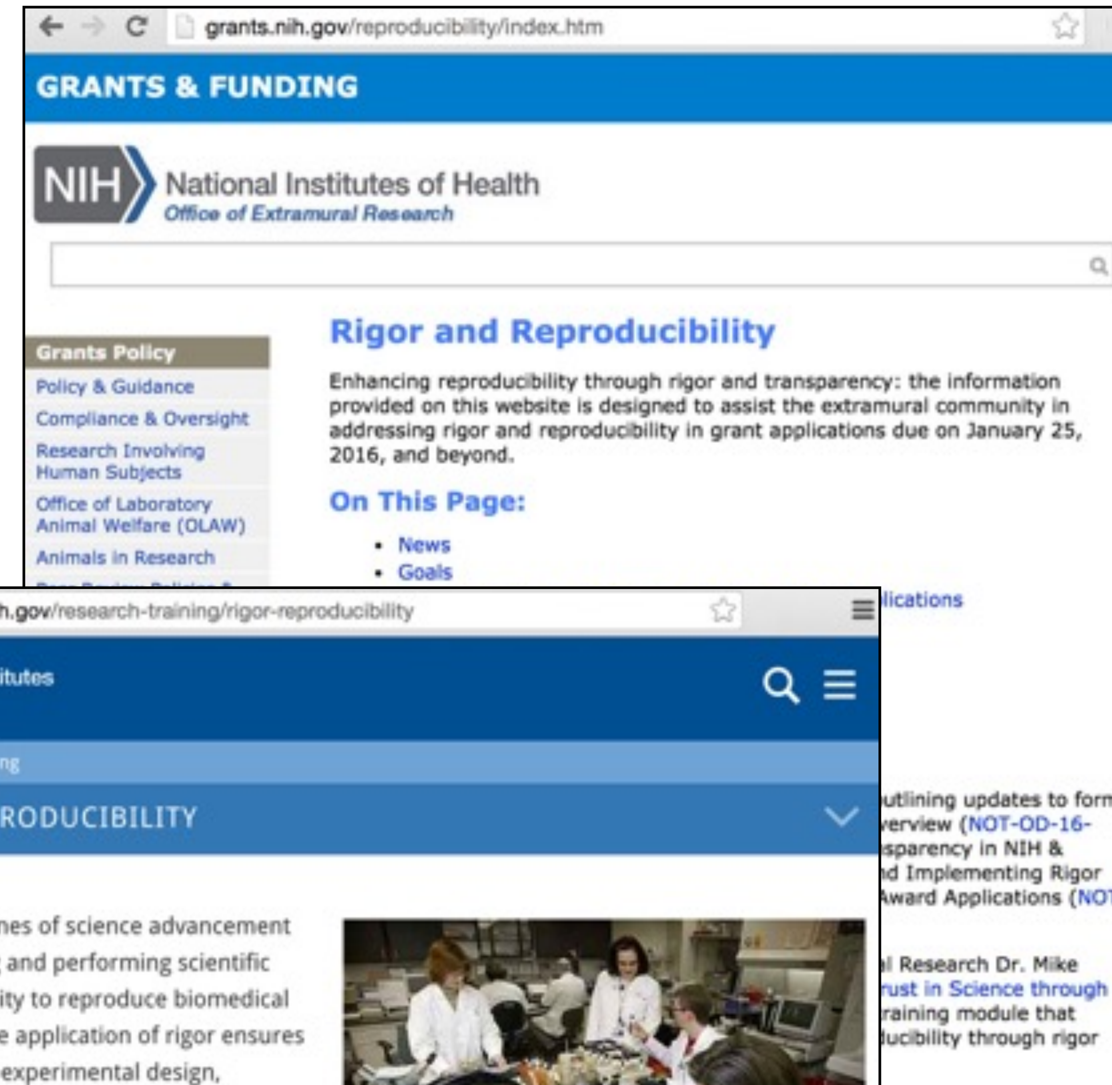
- Show a table of effect sizes and p-values in all phase-3 clinical trials for Melanoma published after 1994;
- Name all of the image denoising algorithms ever used to remove white noise from the famous “Barbara” image, with citations;
- List all of the classifiers applied to the famous acute lymphoblastic leukemia dataset, along with their type-1 and type-2 error rates;
- Create a unified dataset containing all published whole-genome sequences identified with mutation in the gene BRCA1;
- Randomly reassign treatment and control labels to cases in published clinical trial X and calculate effect size. Repeat many times and create a histogram of the effect sizes. Perform this for every clinical trial published in the year 2003 and list the trial name and histogram side by side.

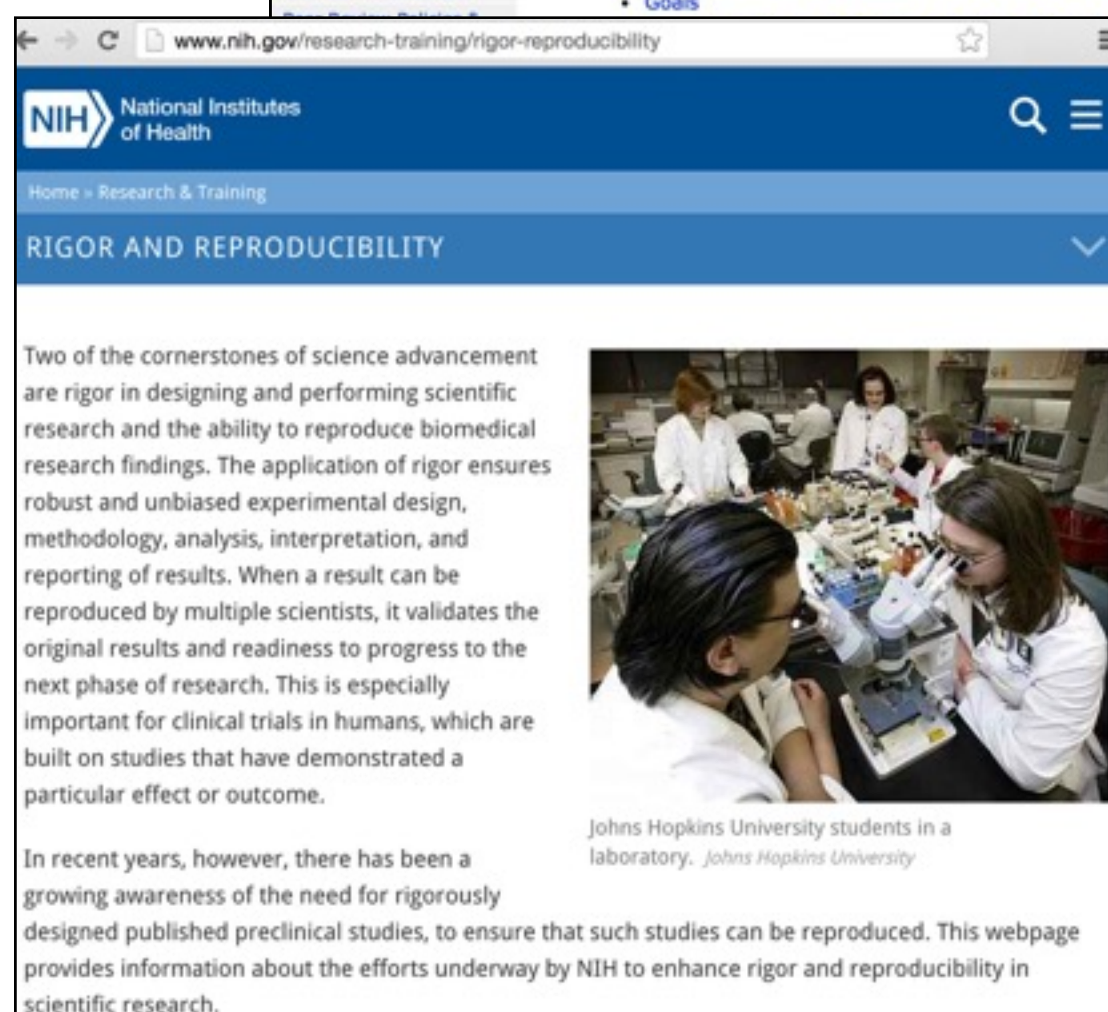
Courtesy of Donoho and Gavish 2012

Government Mandates

- OSTP 2013 Open Data and Open Access Executive Memorandum; Executive Order.
- "Public Access to Results of NSF-Funded Research"
- NOAA Data Management Plan, Data Sharing Plan
- NIST "Common Access Platform"
- …

Federal Agencies





Journal Requirements

- Science: code data sharing since 2011.
- Nature: data sharing.
- …

See also Stodden V, Guo P, Ma Z (2013) "T*oward Reproducible Computational Research: An Empirical Analysis of Data and Code Policy Adoption by Journals*." PLoS ONE 8(6): e67111. doi:10.1371/journal.pone.0067111

Three Principles for CI

1. *Supporting scientific norms*—not only should CI enable new discoveries, but it should also permit others to reproduce the computational findings, reuse and combine digital outputs such as datasets and code, and facilitate validation and comparisons to previous findings.

2. *Supporting best practices in science*—CI in support of science should embed and encourage best practices in scientific research and discovery.

3. *Taking a holistic approach to CI*—the complete end-to-end research pipeline should be considered to ensure interoperability and the effective implementation of 1 and 2.

Social and Political environment..

See Stodden, Miguez, Seiler, "ResearchCompendia.org: Cyberinfrastructure for Reproducibility and Collaboration in Computational Science" CiSE 2015