Open Licensing and Science Policy

Victoria Stodden
Department of Statistics
Columbia University

Guest Lecture

Columbia University

April 16, 2014

Agenda

1. Creating Reliable Computational Science: Updating Research Paradigms
2. Statistical Inference in "Big Data" Settings
3. Rethinking Inference and Scientific Knowledge Production
4. Some Efforts to Address these Challenges

1. Creating Reliable Computational Science: Updating Research Paradigms

Reproducibility in Science

- not a new concept, rooted in *skepticism*
- Transactions of the Royal Society 1660's
- Transparency, knowledge transfer -> goal to perfect the *scholarly record*.
- Technology has changed the nature of experimentation, data, and communication.

Computation is Becoming Central to Scientific Research

1. enormous, and increasing, amounts of data collection:
   - CMS project at LHC: 300 "events" per second, 5.2M seconds of runtime per year, .5MB per event = 780TB/yr => several PB when data processed,
   - Sloan Digital Sky Survey: 9th data release (SDSS-III 2012), 60TB,
   - quantitative revolution in social science due to abundance of social network data (Lazier et al, *Science*, 2009)
   - *Science* survey of peer reviewers: 340 researchers regularly work with datasets >100GB; 119 regularly work with datasets >1TB (N=1700, Feb 11, 2011, p. 692)
2. massive simulations of the complete evolution of a physical system, systematically varying parameters,
3. deep intellectual contributions now encoded in software.

Credibility Crisis

JASA June	Computational Articles	Code Publicly Available
1996	9 of 20	0%
2006	33 of 35	9%
2009	32 of 32	16%
2011	29 of 29	21%

Ioannidis (2011): 9% of authors studied made data available.

Generally, data and code not made available at the time of publication, insufficient information in the publication for verification, replication of results. *A Credibility Crisis*

Scientific Perspective

“Really Reproducible Research” pioneered 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.”

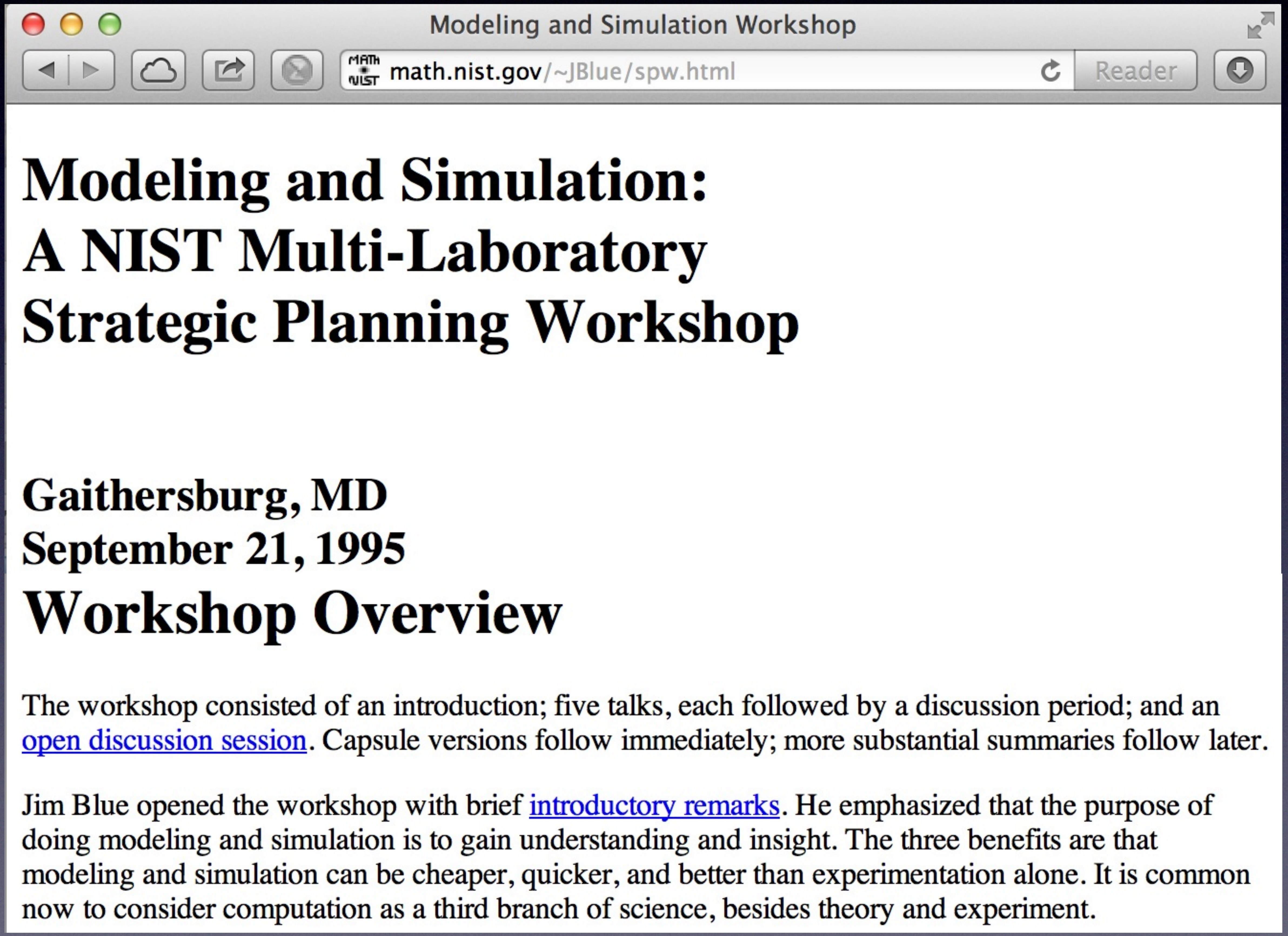
paraphrased by David Donoho, 1998.

Updating the Scientific Method

Argument: computation presents only a *potential* third branch of the scientific method (Stodden et al 2009):

- Branch 1 (deductive): mathematics, formal logic,
- Branch 2 (empirical): statistical analysis of controlled experiments,
- Branch 3,4? (computational): large scale simulations / data driven computational science.

Commonly asserted...


Modeling and Simulation Workshop

math.nist.gov/~JBlue/spw.html

Reader

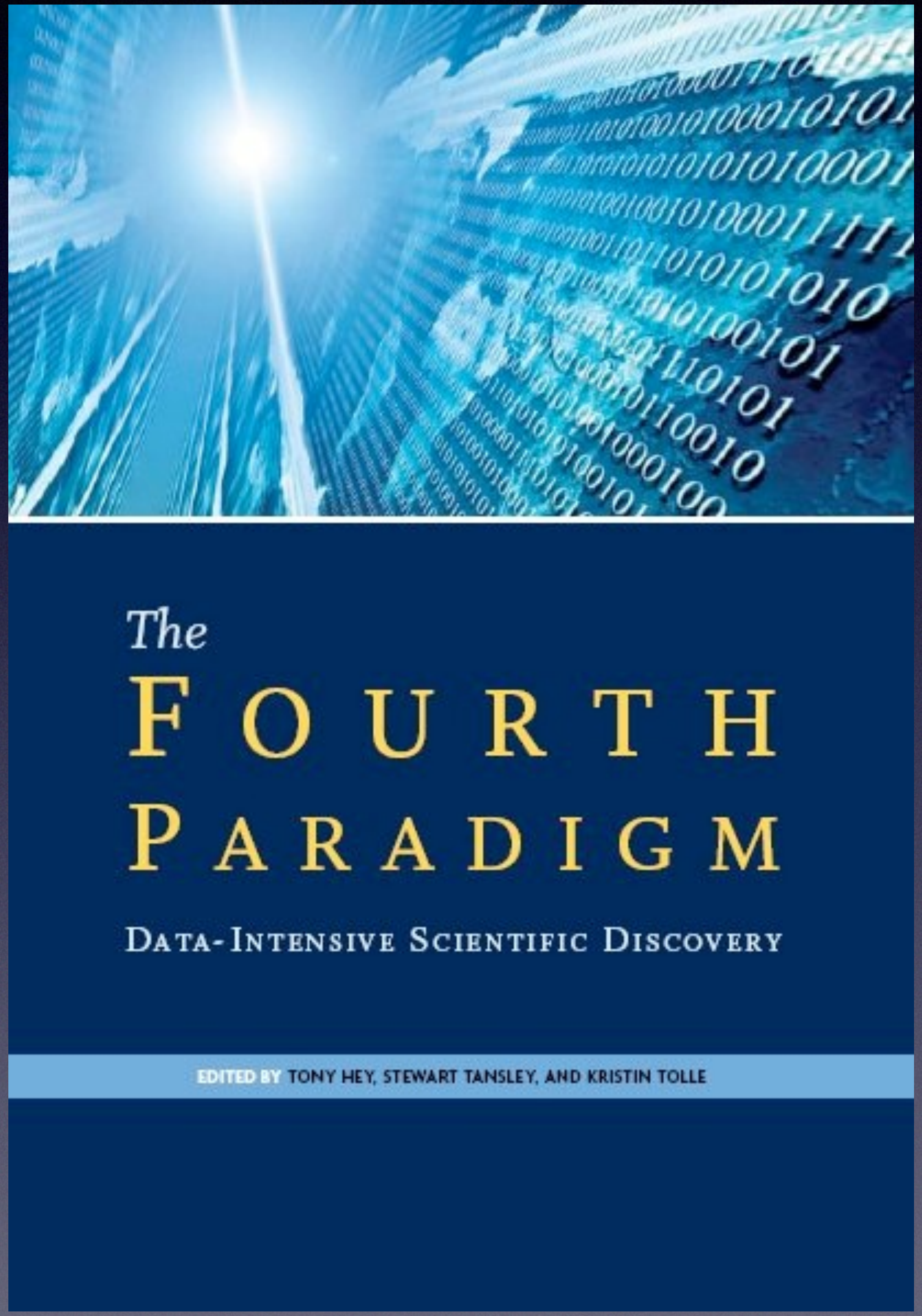
Modeling and Simulation: A NIST Multi-Laboratory Strategic Planning Workshop

Gaithersburg, MD
September 21, 1995

Workshop Overview

The workshop consisted of an introduction; five talks, each followed by a discussion period; and an open discussion session. Capsule versions follow immediately; more substantial summaries follow later.

Jim Blue opened the workshop with brief introductory remarks. He emphasized that the purpose of doing modeling and simulation is to gain understanding and insight. The three benefits are that modeling and simulation can be cheaper, quicker, and better than experimentation alone. It is common now to consider computation as a third branch of science, besides theory and experiment.



"It is common now to consider computation as a third branch of science, besides theory and experiment."

"This book is about a new, fourth paradigm for science based on data-intensive computing."

Credibility Crisis

Los Angeles Times | BUSINESS

LOCAL U.S. WORLD BUSINESS SPORTS ENTERTAINMENT HEALTH STYLE TRAV

Science has lost its way, at a big cost to humanity

Researchers are rewarded for splashy findings, not for double-checking accuracy. So many scientists looking for cures to diseases have been building on ideas that aren't even true.

Science AAAS.ORG | FEEDBACK | HELP | LIBRARIANS All Science Journals Enter Search Ter

GUEST ALERTS ACCESS RIGH

AAAS NEWS SCIENCE JOURNALS CAREERS MULTIMEDIA COLLECTIONS

Science The World's Leading Journal of Original Scientific Research, Global News, and Commentary.

Science Home | Current Issue | Previous Issues | *Science* Express | *Science* Products | My *Science* | About the Journal

Home > *Science* Magazine > 17 January 2014 > McNutt, **343** (6168): 229

Article Views
- Summary
- Full Text
- Full Text (PDF)

Article Tools
- Save to My Folders
- Download Citation
- Alert Me When Article is Cited
- Post to CiteULike
- E-mail This Page
- Rights & Permissions
- Commercial Reprints and E-Prints
- View PubMed Citation

Related Content

Science 17 January 2014:
Vol. 343 no. 6168 p. 229
DOI: 10.1126/science.1250475

‹ Prev | Table of Contents | Next ›
Read Full Text to Comment (8)

EDITORIAL

Reproducibility

Marcia McNutt

» Marcia McNutt is Editor-in-Chief of *Science*.

Science advances on a foundation of trusted di
approach that scientists use to gain confidence
community was shaken by reports that a troub
not reproducible. Because confidence in results
community, we are announcing new initiatives
Science. For preclinical studies (one of the targ
recommendations of the U.S. National Institute
increasing transparency.* Authors will indicate
handling (such as how to deal with outliers), wh
ensure a sufficient signal-to-noise ratio, wheth
experimenter was blind to the conduct of the e
guidelines.

Announcement: Reducing our irreproducibility : Nature News & Comment
www.nature.com/news/announcement-reduci Reader

nature.com : Sitemap Login : Register

nature International weekly journal of science Search Go ▸ Advanced search

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For Authors

Archive › Volume 496 › Issue 7446 › Editorial › Article

NATURE | EDITORIAL

Announcement: Reducing our irreproducibility

24 April 2013

PDF | Rights & Permissions

Over the past year, *Nature* has published a string of articles tha
reliability and reproducibility of published research (collected an

TheScientist
EXPLORING LIFE. INSPIRING INNOVATION

NIH Tackles Irreproducibility

The federal agency speaks out about how to improve the quality of scientific research.

By Jef Akst | January 28, 2014

nature International weekly journal of science

Menu ▸ Advanced search Search Go

archive ▸ volume 483 ▸ issue 7391 ▸ editorials ▸ article

NATURE | EDITORIAL

Must try harder

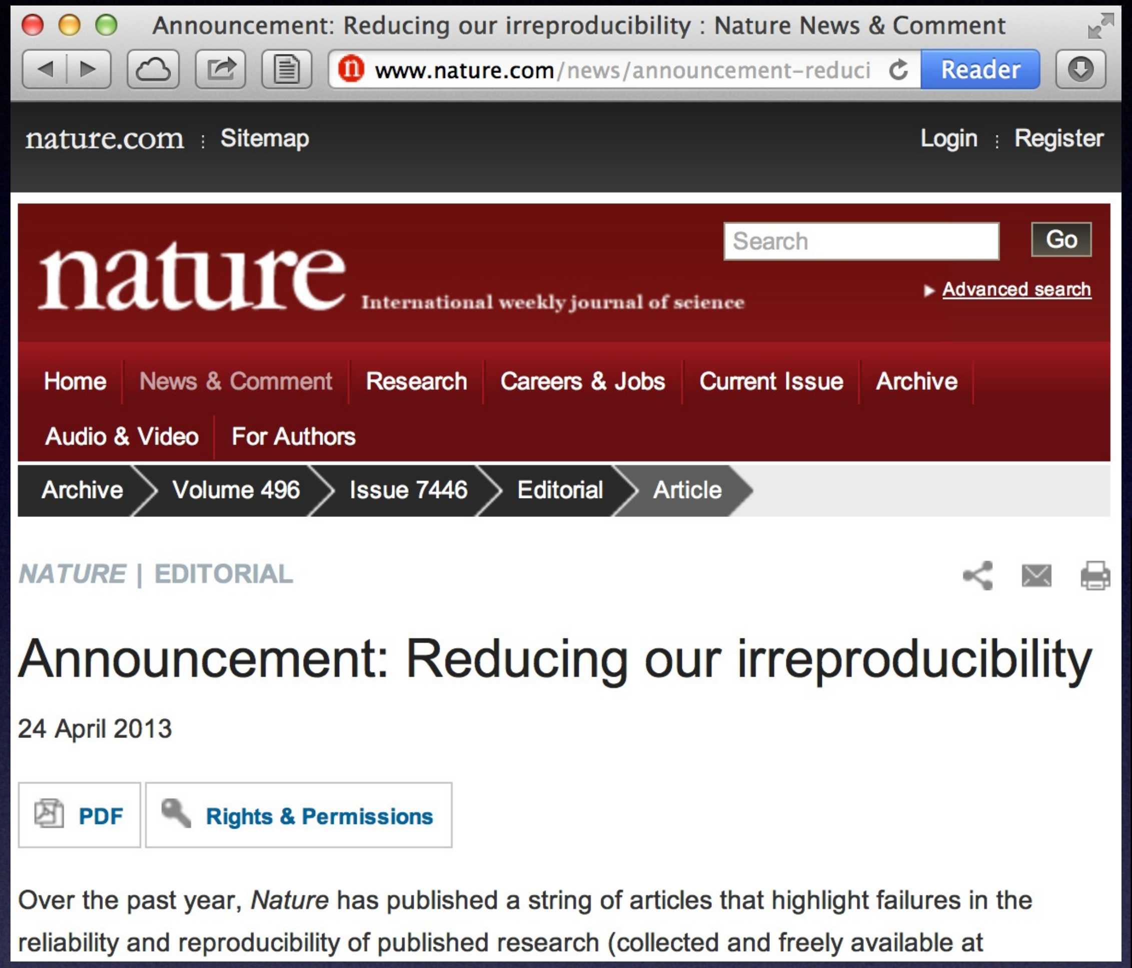
Nature **483**, 509 (29 March 2012) | doi:10.1038/483509a
Published online 28 March 2012

PDF | Citation | Reprints | Rights & permissions | Article metrics

Too many sloppy mistakes are creeping into scientific papers. Lab heads must look more rigorously at the data — and at themselves.

Parsing Reproducibility

"Empirical Reproducibility"


Announcement: Reducing our irreproducibility : Nature News & Comment

www.nature.com/news/announcement-reduci

nature.com : Sitemap Login : Register

nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For Authors

Archive > Volume 496 > Issue 7446 > Editorial > Article

NATURE | EDITORIAL

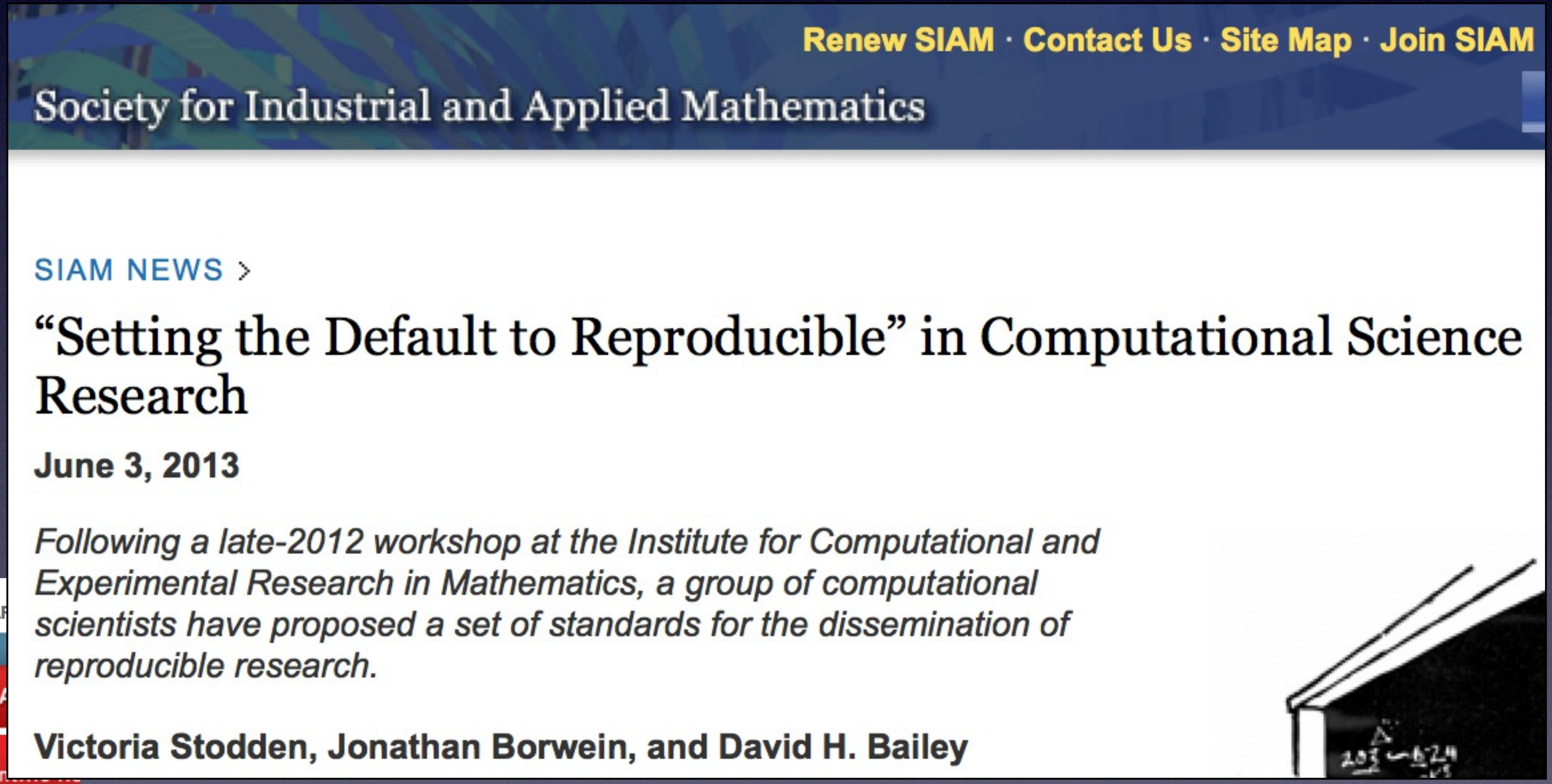
Announcement: Reducing our irreproducibility

24 April 2013

PDF | Rights & Permissions

Over the past year, *Nature* has published a string of articles that highlight failures in the reliability and reproducibility of published research (collected and freely available at

"Computational Reproducibility"


Renew SIAM · Contact Us · Site Map · Join SIAM

Society for Industrial and Applied Mathematics

SIAM NEWS >

"Setting the Default to Reproducible" in Computational Science Research

June 3, 2013

Following a late-2012 workshop at the Institute for Computational and Experimental Research in Mathematics, a group of computational scientists have proposed a set of standards for the dissemination of reproducible research.

Victoria Stodden, Jonathan Borwein, and David H. Bailey

"Statistical Reproducibility"


Science | AAAS.ORG | FEEDBACK | HELP | LIBRA

AAAS | NEWS | SCIENCE JOURNALS | CA

Science The World's Leading Journal of Original Scie

Science Home | Current Issue | Previous Issues | Science Express | Sci

Home > *Science* Magazine > 17 January 2014 > McNutt, **343** (6168): 229

Article Views
- Summary
- Full Text
- Full Text (PDF)

Article Tools
- Save to My Folders

Science 17 January 2014:
Vol. 343 no. 6168 p. 229
DOI: 10.1126/science.1250475

EDITORIAL

Reproducibility

Marcia McNutt

» Marcia McNutt is Editor-in-Chief of *Science*.

V. Stodden, IMS Bulletin (2013)

ICERM Workshop Dec 2012

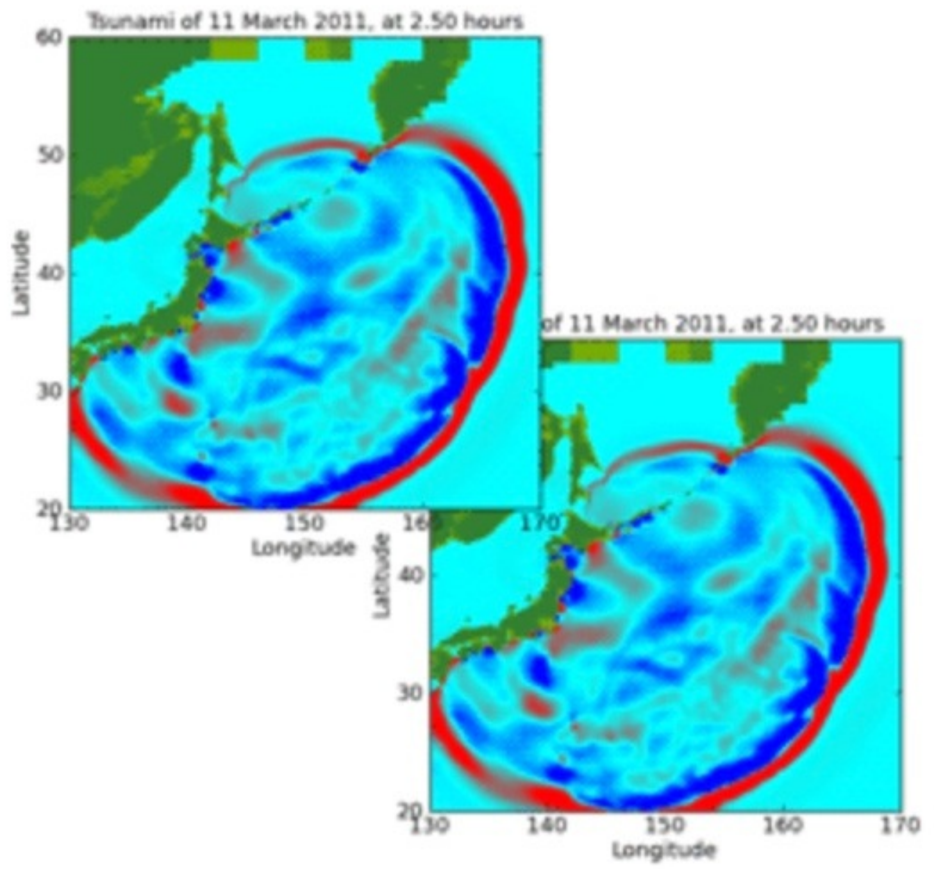
icerm

Search

Home | Programs & Events | Participate | Proposals | Resources | For Visitors | People | News | Diversity | Support ICERM

Reproducibility in Computational and Experimental Mathematics *(December 10-14, 2012)*

Description



Click for code to create this image.

In addition to advancing research and discovery in pure and applied mathematics, computation is pervasive across the sciences and now computational research results are more crucial than ever for public policy, risk management, and national security. Reproducibility of carefully documented experiments is a cornerstone of the scientific method, and yet is often lacking in computational mathematics, science, and engineering. Setting and achieving appropriate standards for reproducibility in computation poses a number of interesting technological and social challenges. The purpose of this workshop is to discuss aspects of reproducibility most relevant to the mathematical sciences among researchers from pure and applied mathematics from academics and other settings, together with interested parties from funding agencies, national laboratories, professional societies, and publishers. This will be a working workshop, with relatively few talks and dedicated time for breakout group discussions on the current state of the art and the tools, policies, and infrastructure that are needed to improve the situation. The groups will be charged with developing guides to current best practices and/or white papers on desirable advances.

Organizing Committee

- David H. Bailey (Lawrence Berkeley National Laboratory)
- Jon Borwein (Centre for Computer Assisted Research Mathematics and its Applications)
- Randall J. LeVeque (University of Washington)
- Bill Rider (Sandia National Laboratory)
- William Stein (University of Washington)
- Victoria Stodden (Columbia University)

ICERM Workshop Report

Setting the Default to Reproducible

Reproducibility in Computational and Experimental Mathematics

Developed collaboratively by the ICERM workshop participants[1]

Compiled and edited by the Organizers

V. Stodden, D. H. Bailey, J. Borwein, R. J. LeVeque, W. Rider, and W. Stein

Abstract

Science is built upon foundations of theory and experiment validated and improved through open, transparent communication. With the increasingly central role of computation in scientific discovery this means communicating all details of the computations needed for others to replicate the experiment, i.e. making available to others the associated data and code. The "reproducible research" movement recognizes that traditional scientific research and publication practices now fall short of this ideal, and encourages all those involved in the production of computational science – scientists who use computational methods and the institutions that employ them, journals and dissemination mechanisms, and funding agencies – to facilitate and practice really reproducible research.

Set the Default to "Open"

Reproducible Science in the Computer Age. Conventional wisdom sees computing as the "third leg" of science, complementing theory and experiment. That metaphor is outdated. Computing now pervades all of science. Massive computation is often required to reduce and analyze data; simulations are employed in fields as diverse as climate modeling and astrophysics. Unfortunately, scientific computing culture has not kept pace. Experimental researchers are taught early to keep notebooks or computer logs of every work detail: design, procedures, equipment, raw results, processing techniques, statistical methods of analysis, etc. In contrast, few computational experiments are performed with such care. Typically, there is no record of workflow, computer hardware and software configuration, or parameter settings. Often source code is lost. While crippling reproducibility of results, these practices ultimately impede the researcher's own productivity.

The State of Experimental and Computational Mathematics. Experimental mathematics[1]—application of high-performance computing technology to research questions in pure and applied mathematics, including

ScienceCartoonsPlus.com.

"It says it's sick of doing things like inventories and payrolls, and it wants to make some breakthroughs in astrophysics."

physicists, legal scholars, journal editors, and funding agency officials representing academia, government labs, industry research, and all points in between. While

Renew SIAM · Contact Us · Site Map · Join SIAM

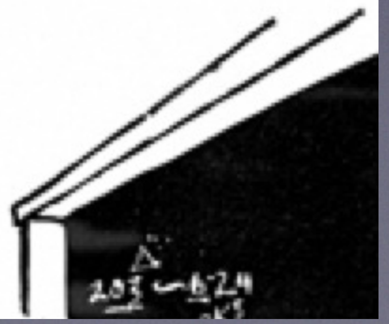
Society for Industrial and Applied Mathematics

SIAM NEWS >

"Setting the Default to Reproducible" in Computational Science Research

June 3, 2013

Following a late-2012 workshop at the Institute for Computational and Experimental Research in Mathematics, a group of computational scientists have proposed a set of standards for the dissemination of reproducible research.



Victoria Stodden, Jonathan Borwein, and David H. Bailey

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, Stodden, et al. 2009), until the development of comparable standards.

Data and Software

1. Digital databases require software and computing power for curation, filtering, analysis, visualization, and storage.

2. Computation requires software, and scientific investigation either produces data (e.g. simulations), or uses data as an input (e.g. data science), or both.

→ Data and software should be considered together as different but inextricably linked components of the computational science research process.

Intellectual Property Barriers

- Software is both copyrighted (by default) and patentable.
- Copyright: author sets terms of use using an open license:
  - Attribution only (ie. Modified BSD, MIT license, LGPL)
  - *Reproducible Research Standard (Stodden 2009)*
- Patents: Bayh-Dole (1980) vs reproducible research (Stodden 2012)
  - delays, barriers to software access
  - *Bilski v Kappos (2011)*

Legal Barriers: Copyright

"To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries." (U.S. Const. art. I, §8, cl. 8)

- Original expression of ideas falls under copyright *by default* (papers, code, figures, tables..)
- Copyright secures exclusive rights vested in the author to:
  - reproduce the work
  - prepare derivative works based upon the original

Exceptions and Limitations: Fair Use.

Responses Outside the Sciences 1: Open Source Software

- 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

Responses Outside the Sciences 2: Creative Commons

- Founded in 2001, by Stanford Law Professor Larry Lessig, MIT EECS Professor Hal Abelson, and advocate Eric Eldred.
- Adapts the Open Source Software approach to artistic and creative digital works.

Response from Within the Sciences

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.

Winner of the Access to Knowledge Kaltura Award 2008

Copyright and Data

- Copyright adheres to raw facts in Europe.
- In the US raw facts are not copyrightable, but the original "selection and arrangement" of these facts is copyrightable. (Feist Publns Inc. v. Rural Tel. Serv. Co., 499 U.S. 340 (1991)).
- the possibility of a residual copyright in data (attribution licensing or public domain certification).
- Law doesn't match reality on the ground: What constitutes a "raw" fact anyway?

GPL and Copyleft

"sharealike" See e.g. https://www.gnu.org/copyleft/

Bayh-Dole Act (1980)

- Promote the transfer of academic discoveries for commercial development, via licensing of patents (ie. Technology Transfer Offices),
- Bayh-Dole Act gave federal agency grantees and contractors title to government-funded inventions and charged them with using the patent system to aid disclosure and commercialization of the inventions.
- greatest impact in biomedical research collaborations and drug discovery.

Impact of Computation

In the computational sciences, disclosure of data and code are considered essential for reproducibility. Software can be patent-eligible (Bilski v. Kappos 130 S. Ct. 3218 2010), increasing the reach of Bayh-Dole in the sciences.

Universities can claim ownership over software developed in the course of research on this basis and potentially patent then license access to the code.

Hypothesis: The Bayh-Dole Act inhibits reproducibility in the computational sciences, and is a barrier to access to research inventions.

Sharing: Funding Agency Policy

- NSF grant guidelines: "NSF ... expects investigators to share with other researchers, at no more than incremental cost and within a reasonable time, the data, samples, physical collections and other supporting materials created or gathered in the course of the work. It also encourages grantees to share software and inventions or otherwise act to make the innovations they embody widely useful and usable." (2005 and earlier)
- NSF peer-reviewed Data Management Plan (DMP), January 2011.
- NIH (2003): "The NIH expects and supports the timely release and sharing of final research data from NIH-supported studies for use by other researchers." (>$500,000, include data sharing plan)

NSF Data Management Plan

“Proposals submitted or due on or after January 18, 2011, must include a supplementary document of no more than two pages labeled ‘Data Management Plan.’ This supplementary document should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results.” (http://www.nsf.gov/bfa/dias/policy/dmp.jsp)

Software management plans appearing.. (BigData joint NSF/NIH solicitation)

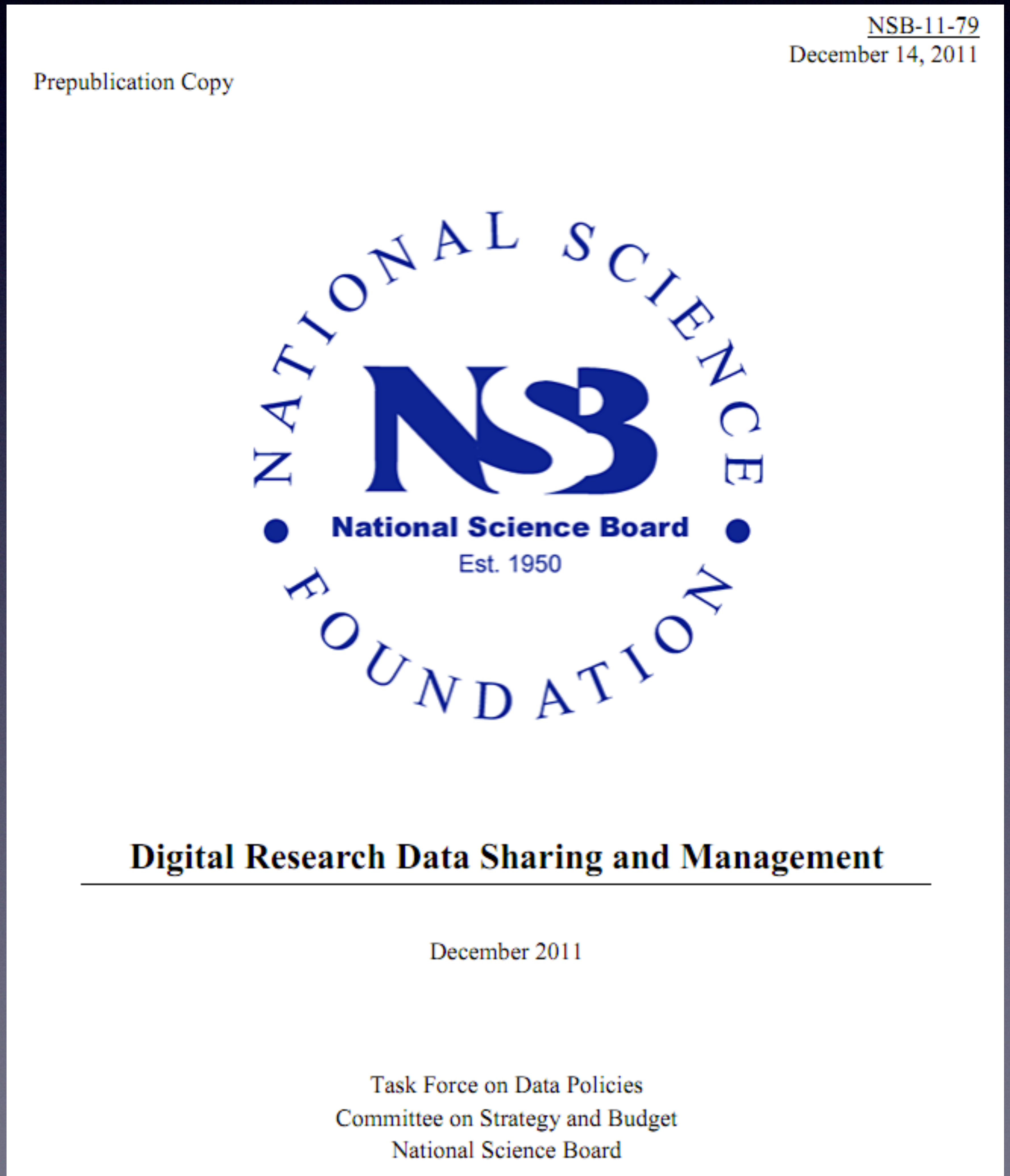
2013: Open Science in DC

- Feb 22: Executive Memorandum directing federal funding agencies to develop plans for public access to data and publications.
- May 9: Executive Order directing federal agencies to make their data publicly available.

Science Policy in Congress

- America COMPETES due to be reauthorized, drafting underway.
- Sensenbrenner introduced "Public Access to Science," Sept 19, 2013.
- Hearing on Research Integrity and Transparency by the House Science, Space, and Technology Committee (March 5).
- Reproducibility cannot be an unfunded mandate.

National Science Board Report


NSB-11-79
December 14, 2011

Prepublication Copy

NATIONAL SCIENCE FOUNDATION
NSB
National Science Board
Est. 1950

Digital Research Data Sharing and Management

December 2011

Task Force on Data Policies
Committee on Strategy and Budget
National Science Board

"Digital Research Data Sharing and Management," December 2011.

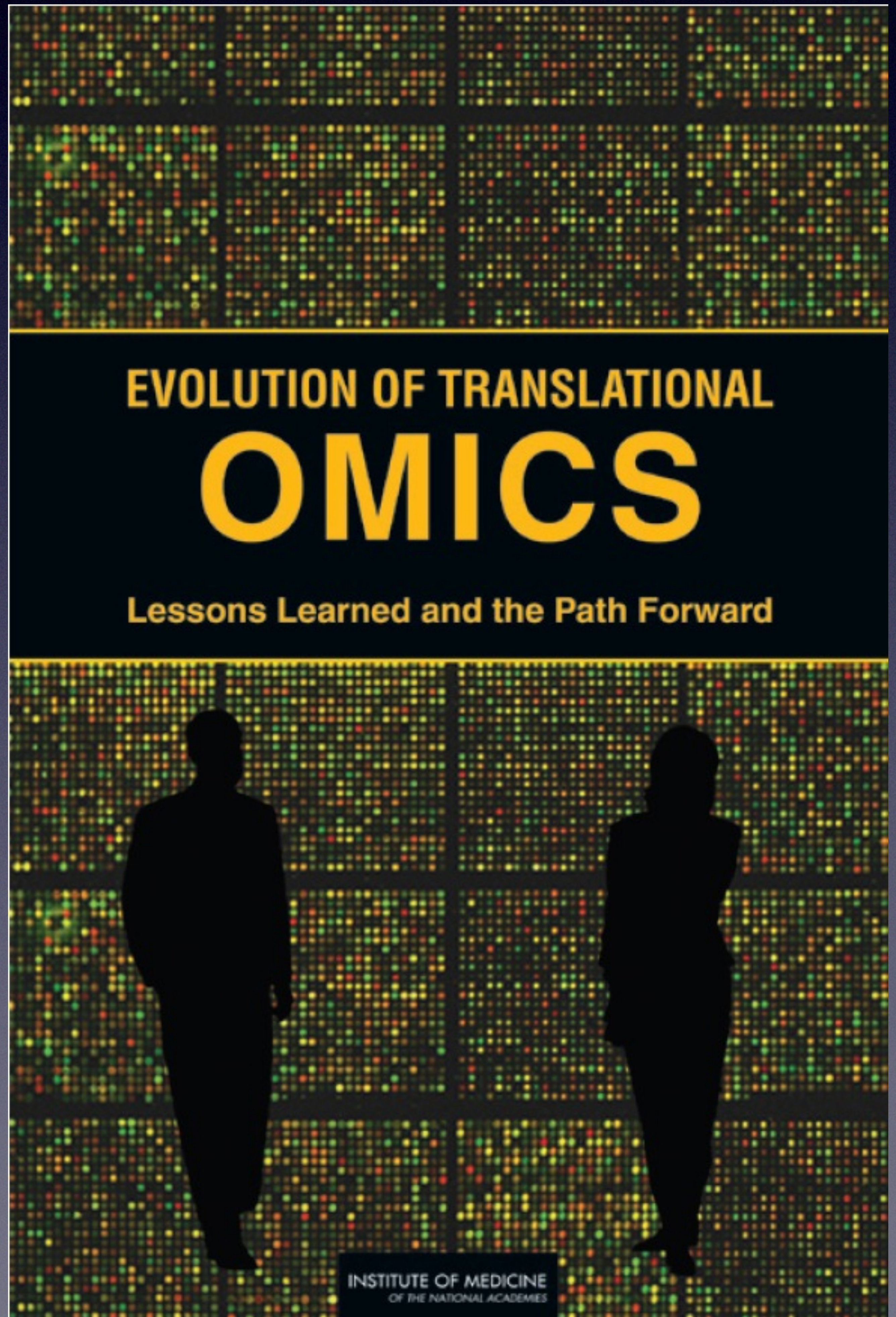
http://www.nsf.gov/nsb/publications/2011/nsb1124.pdf

NAS Data Sharing Report

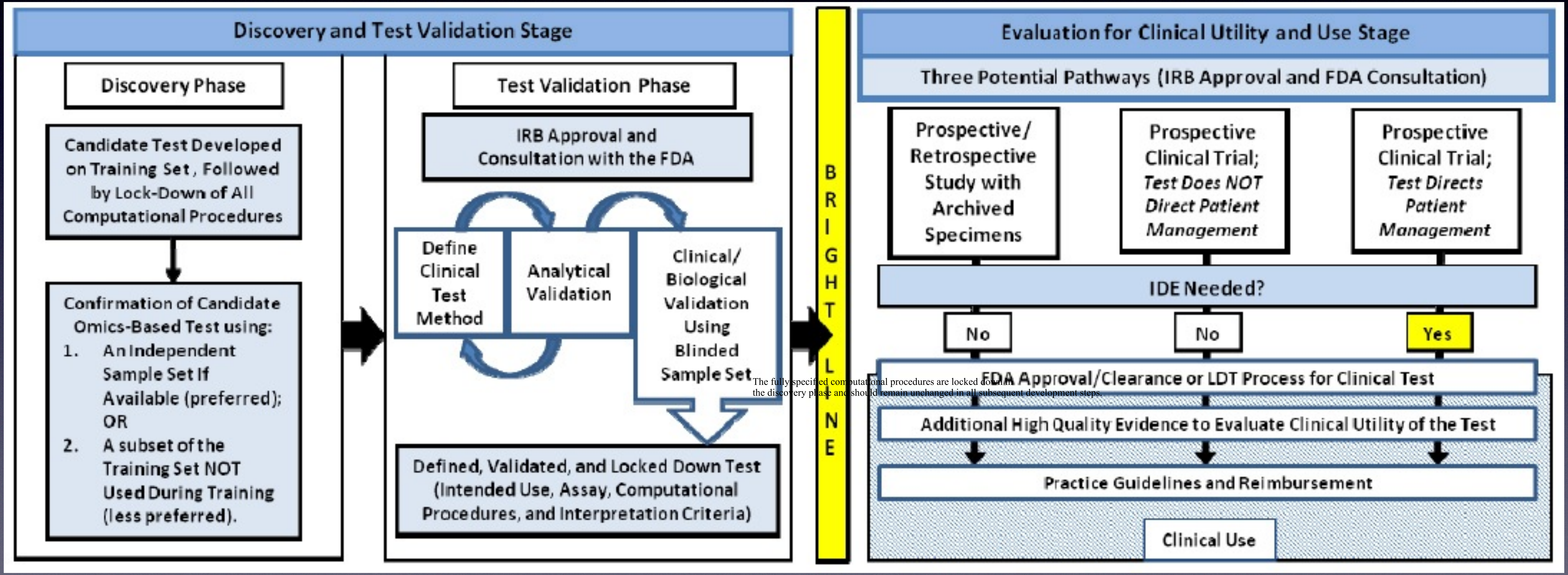
- Sharing Publication-Related Data and Materials: Responsibilities of Authorship in the Life Sciences, (2003)
- "Principle 1. Authors should include in their publications the data, algorithms, or other information that is central or integral to the publication—that is, whatever is necessary to support the major claims of the paper and would enable one skilled in the art to verify or replicate the claims."

IOM "Evolution of Translational Omics: Lessons Learned and the Path Forward"

- March 23 2012, Institute of Medicine releases report,
- Recommends new standards for omics-based tests, including a fixed version of the software, expressly for verification purposes.



IOM Report: Figure S-1



"The fully specified computational procedures are locked down in the discovery phase and should remain unchanged in all subsequent development steps."

2. Statistical Inference in “Big Data” Settings

Traditional Sources of Error

Statistics has a long history of identifying and controlling for error in empirical experiments, e.g.:

- sampling errors,
- selection bias,
- model specification errors, functional form misspecification,
- omitted variable bias,
- …

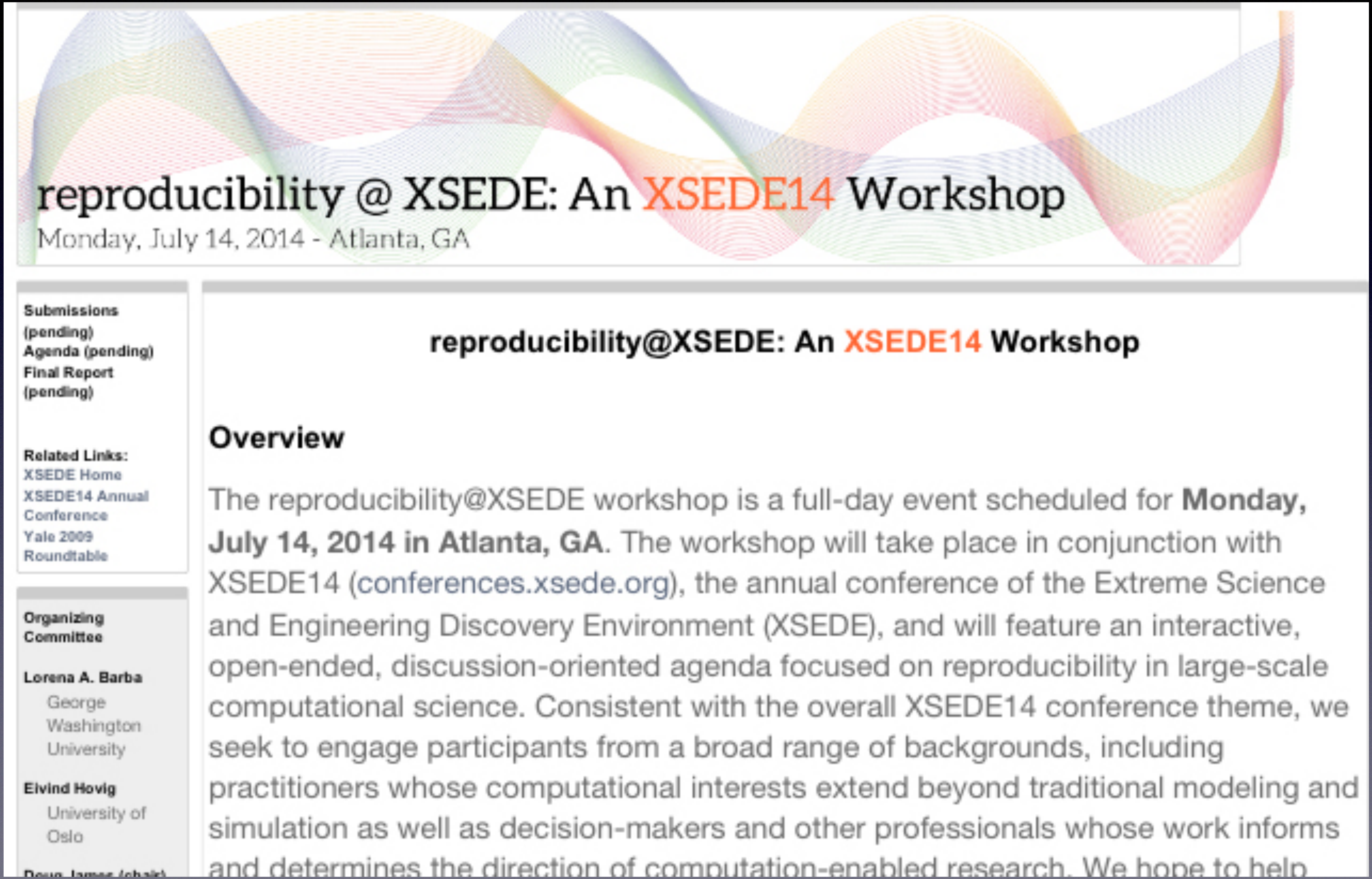
ICERM Reporting Approach

- Omitted information regarding the implementation of the experiment (e.g. parameters, parameter estimation) as source of irreproducibility.
- Introduces uncertainty when interpreting or re-estimating results.
- Addressed in ICERM workshop report.

Reproducibility at Scale

Scale Issues: both for large datasets and compute time.

- data produced by code - making the code available permit data regeneration, but may involve prohibitive runtimes.
- partial data precomputation - what standards? partial results checking and testing. data synopses, simulated data.
- do some research areas pose extra difficulties?


reproducibility @ XSEDE: An XSEDE14 Workshop
Monday, July 14, 2014 - Atlanta, GA
Submissions (pending)
Agenda (pending)
Final Report (pending)
Related Links:
XSEDE Home
XSEDE14 Annual Conference
Yale 2009 Roundtable
Organizing Committee
Lorena A. Barba
George Washington University
Eivind Hovig
University of Oslo
reproducibility@XSEDE: An XSEDE14 Workshop
Overview
The reproducibility@XSEDE workshop is a full-day event scheduled for Monday, July 14, 2014 in Atlanta, GA. The workshop will take place in conjunction with XSEDE14 (conferences.xsede.org), the annual conference of the Extreme Science and Engineering Discovery Environment (XSEDE), and will feature an interactive, open-ended, discussion-oriented agenda focused on reproducibility in large-scale computational science. Consistent with the overall XSEDE14 conference theme, we seek to engage participants from a broad range of backgrounds, including practitioners whose computational interests extend beyond traditional modeling and simulation as well as decision-makers and other professionals whose work informs and determines the direction of computation-enabled research. We hope to help

Stability and Robustness of Results

- Computational reproducibility addresses whether fixed codes/data can replicate findings, permitting the reconciliation of differences in independent efforts.
- does not directly address whether these findings improve our understanding of the world.
- we might expect that for such findings, repeated independent replications yield results that are "close." Possible sources of variation (B. Yu, 2013):
    - *Stability:* "reasonable" perturbations in the underlying data.
    - *Robustness:* perturbations in methods (due to changes in the parametrization, model, or model assumptions).

Some "New" Sources of Error

- Statistical: frequentist paradigm, p-values, multiplicity, power (dataset size improves the reliability results!), outlier classification, hypothesis testing.
- Computational: traversing data, tools, scaling algorithms (dataset size is a challenge!).
- Reconciling conflicting results,
- Duplication of efforts,
- Standing on the shoulders of giants: software perspective.

3. Rethinking Inference and Scientific Knowledge Production

We need:

- Standards for reproducibility of big data findings:

  1. data access, software access (what form? what documentation? collaboration with proprietary software holders).

  2. data generation vs data analysis.

  3. robust methods, producing stable results, emphasis on reliability and reproducibility.

Example: Google Flu Trends results: worked at first, but what happened?

4. Some Efforts to Address these Challenges

Openness and Science Policy

- Science Policy must support scientific ends: Reliability and accuracy of the scientific record.
- Facilitate Reproducibility - the ability to regenerate published results (data and code availability, alongside results).
- Need infrastructure to facilitate (1):
    1. deposit/curation of versioned data and code,
    2. link to published article,
    3. permanence of link.

Constructing Policy

- "Open Data" is not well-defined. Scope: Share data and code that permit others in the field to replicate published results. (traditionally done by the publication alone). Corollary: maximizes data reuse.
- Data and code availability at the time of publication.
- Public access. "With many eyeballs, all bugs are shallow."
- Need infrastructure/software tools to facilitate (2):
    1. data/code suitable for sharing, created *during the research process*.

2013: Open Science in DC

- Feb 22: Executive Memorandum directing federal funding agencies to develop plans for public access to data and publications.
- May 9: Executive Order directing federal agencies to make their data publicly available.

Tools for Computational Science

- 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   IPython Notebook
  Galaxy   GenePattern   Paper Mâché
  Sumatra   Taverna   Pegasus

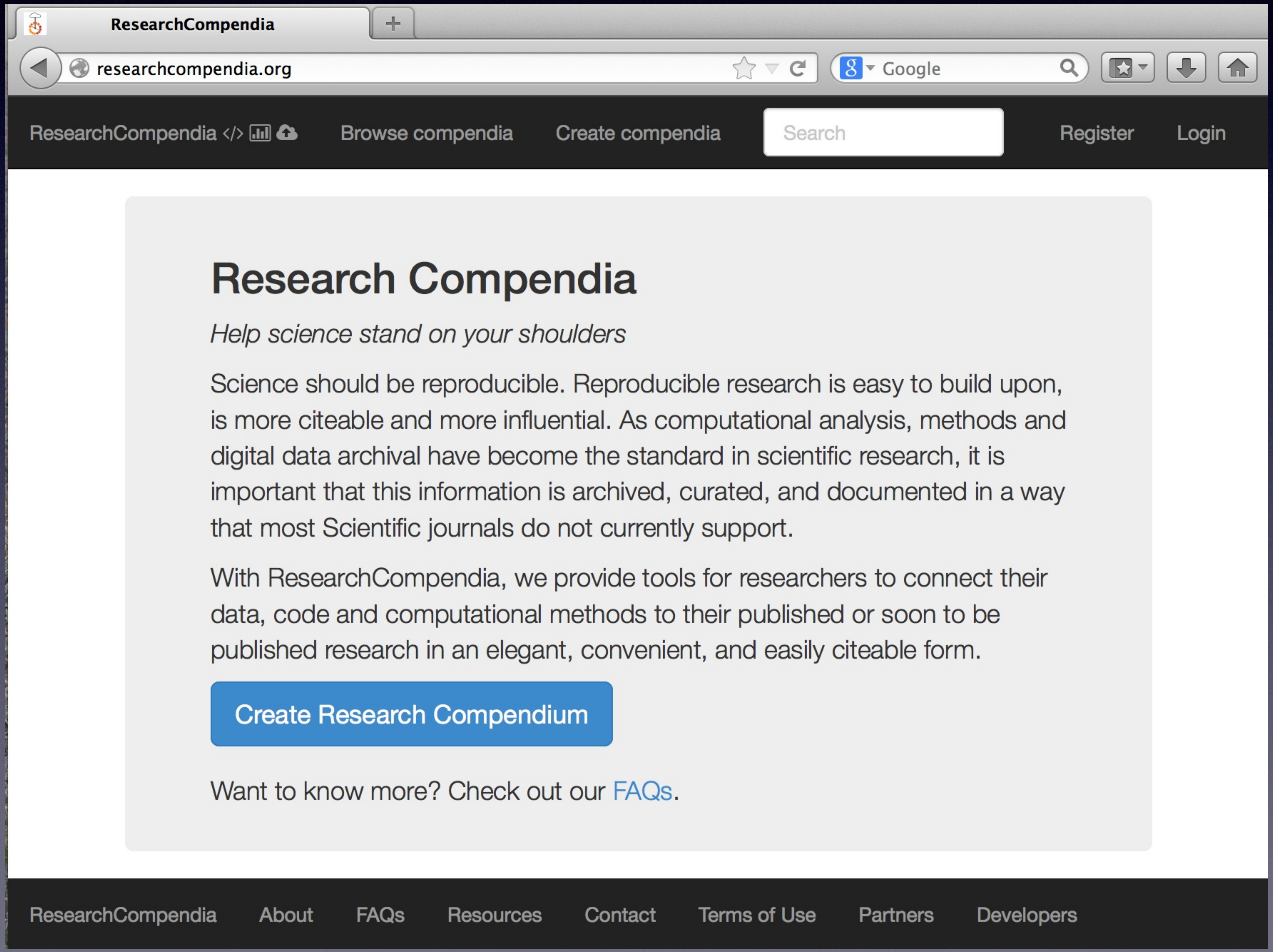
- Embedded Publishing:

  Verifiable Computational Research   Sweave   knitR
  Collage Authoring Environment   SHARE

Research Compendia

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

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



ResearchCompendia

researchcompendia.org/compendia/423/

ResearchCompendia Browse compendia Create compendia Search Register Login

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

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

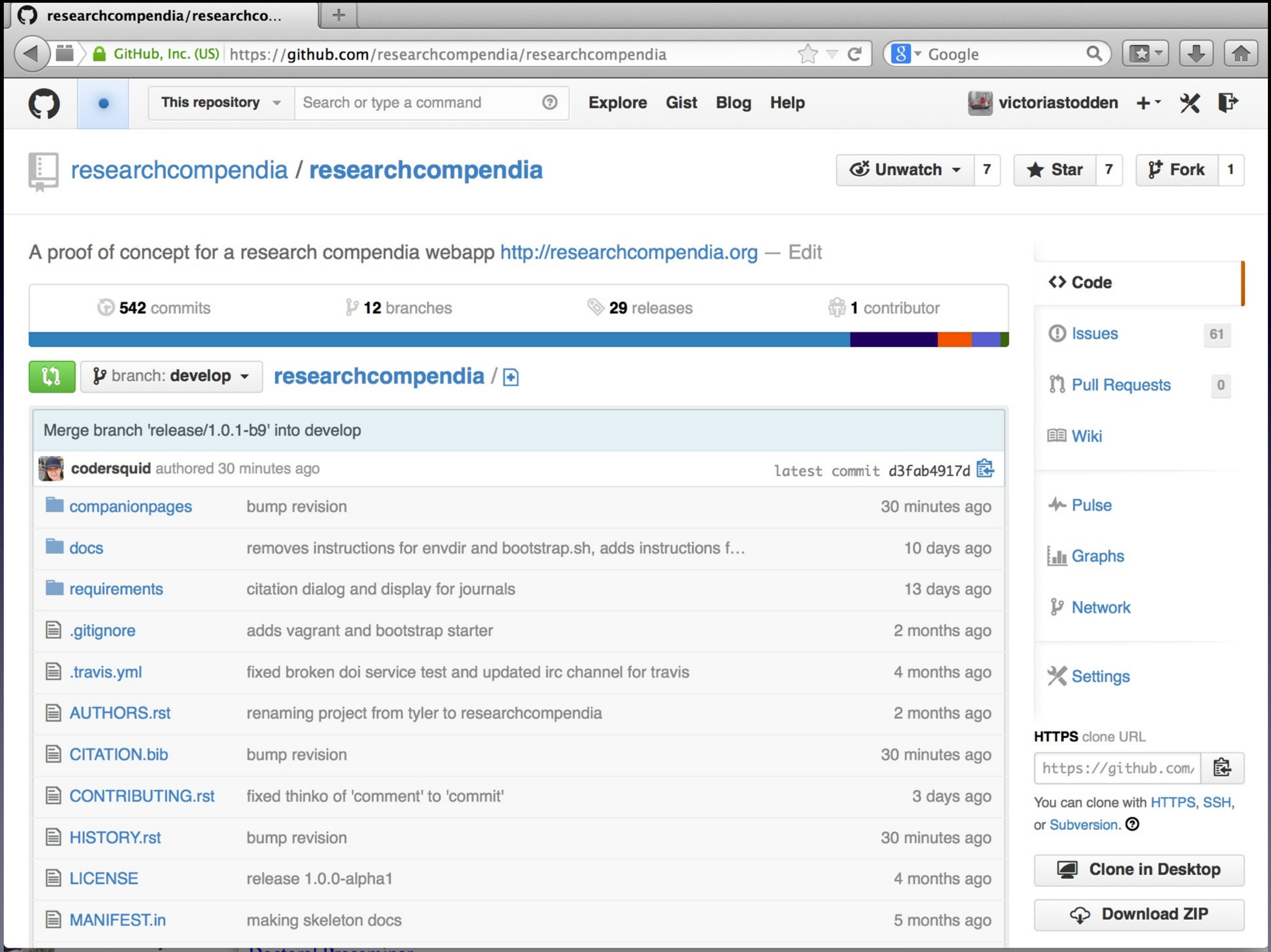
Code and Data Abstract

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 About FAQs Resources Contact Terms of Use Partners Developers


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

"Reproducible Research" is Grassroots

- reproducibility@XSEDE: An XSEDE14 Workshop
- AMP 2011 "Reproducible Research: Tools and Strategies for Scientific Computing"
- Open Science Framework / Reproducibility Project in Psychology
- AMP / ICIAM 2011 "Community Forum on Reproducible Research Policies"
- SIAM Geosciences 2011 "Reproducible and Open Source Software in the Geosciences"
- ENAR International Biometric Society 2011: Panel on Reproducible Research
- AAAS 2011: "The Digitization of Science: Reproducibility and Interdisciplinary Knowledge Transfer"
- SIAM CSE 2011: "Verifiable, Reproducible Computational Science"
- Yale Law School 2009: Roundtable on Data and Code Sharing in the Computational Sciences
- ACM SIGMOD conferences
- NSF/OCI report on Grand Challenge Communities (Dec, 2010)
- IOM "Review of Omics-based Tests for Predicting Patient Outcomes in Clinical Trials"

References

"Toward Reproducible Computational Research: An Empirical Analysis of Data and Code Policy Adoption by Journals," PLoS ONE, June 2013

"Reproducible Research," guest editor for Computing in Science and Engineering, July/August 2012.

"Reproducible Research: Tools and Strategies for Scientific Computing," July 2011.

"Enabling Reproducible Research: Open Licensing for Scientific Innovation," 2009.

available at http://www.stodden.net