

LINKFLOWS

Enabling a web of linked semantic publishing workflows

Cristina-Iulia Bucur

28 March 2018 @ Online Information Quality, Lorentz Workshop

These slides: <https://bit.ly/2Gkd4ud> or <https://goo.gl/jYwifu>

ABOUT LINKFLOWS

- PhD project started in February 2017
- Supervisors: Tobias Kuhn, Davide Ceolin, Lora Aroyo
- Collaborations:
 - Vrije Universiteit Amsterdam
 - IOS Press
 - Netherlands Sound and Vision

WHY LINKFLOWS?

- Scientific publishing: effective means to **share** information and knowledge
- Shift towards the **digital** environment
- New ways to **represent** fine-grained knowledge
- Linked Data: enable the **exchange, reuse** and **linking** of data on the Web

CHALLENGES

CHALLENGES: REPRODUCIBILITY

nature International weekly journal of science

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

Archive > Volume 533 > Issue 7604 > News Feature > Article >

NATURE | NEWS FEATURE

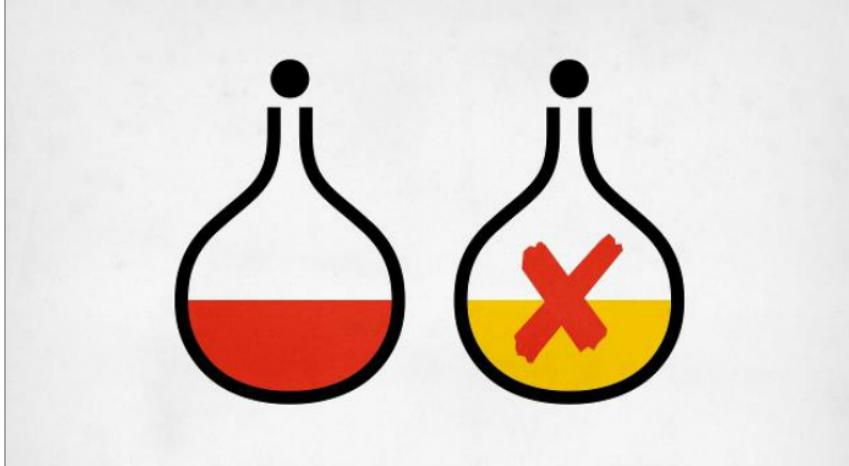
1,500 scientists lift the lid on reproducibility

Survey sheds light on the ‘crisis’ rocking research.

Monya Baker

25 May 2016 | Corrected: 28 July 2016

PDF Rights & Permissions



More than 70% of researchers have tried and failed to reproduce another scientist's experiments, and more than half have failed to reproduce their own experiments. Those are some of the telling figures that emerged from *Nature*'s survey of 1,576 researchers who took a brief online

CHALLENGES: REPRODUCIBILITY

nature International weekly journal of science

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

Archive > Volume 548 > Issue 7668 > Comment > Article

NATURE | COMMENT

A long journey to reproducible results

Gordon J. Lithgow, Monica Driscoll & Patrick Phillips

22 August 2017

Replicating our work took four years and 100,000 worms but brought surprising discoveries, explain Gordon J. Lithgow, Monica Driscoll and Patrick Phillips.

 PDF  Rights & Permissions

Subject terms: Ageing · Research management · Biological techniques · Genetics



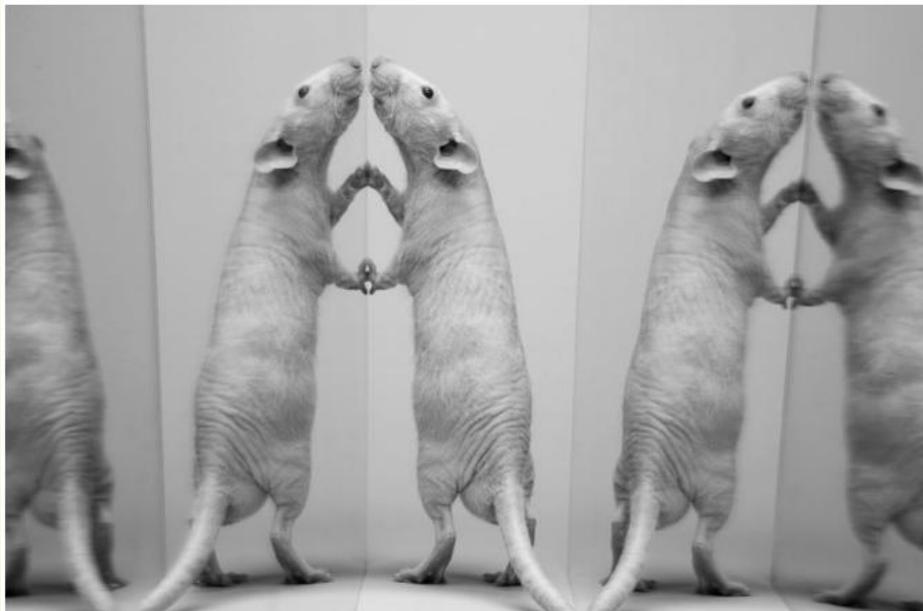
CHALLENGES: REPRODUCIBILITY



EDITORIAL • 21 NOVEMBER 2017 • CORRECTION 15 DECEMBER 2017

Rewarding negative results keeps science on track

Creating a culture of replication takes prizes, grants and magnanimity – as well as publications.



Publishing replication attempts saves other researchers' time. Credit: Grove Pashley/Getty

PDF version

RELATED ARTICLES

Faculty promotion must assess reproducibility:

Go forth and replicate!

No publication without confirmation

Receptive to replication

CHALLENGES: REPRODUCIBILITY

nature
biotechnology



Altmetric: 108

Citations: 8

[More detail >](#)

Editorial

Receptive to replication

Do replication studies belong in top-tier journals?

Like many other scientific journals, *Nature Biotechnology* places a strong emphasis on novelty when selecting research for publication. As a result, studies describing replications or confirmations of previously published reports are less likely to be chosen. And studies detailing null or negative findings may not meet stringent editorial requirements for significance and relevance to our broad readership. Why then are we publishing on p. 965 a replication of a report published elsewhere in the literature?

CHALLENGES: REPRODUCIBILITY

The screenshot shows a web page from the journal *nature*. At the top, the word "nature" is written in a large, lowercase, serif font, followed by the subtitle "International weekly journal of science" in a smaller, sans-serif font. Below this is a navigation bar with links: Home, News & Comment, Research, Careers & Jobs, Current Issue, Archive, Audio & Video, Forum, and Log in. A breadcrumb navigation path is visible below the navigation bar: Archive > Volume 541 > Issue 7637 > Editorial > Article. The main title of the article is "Replication studies offer much more than technical details". Below the title, a subtitle reads "They demonstrate the practice of science at its best." The publication date is listed as "18 January 2017". At the bottom of the page, there are two buttons: "PDF" and "Rights & Permissions". The main text of the article begins with: "Purists will tell you that science is about what scientists don't know, which is true but not much of a basis on which to develop new cancer drugs. Hence the importance of knowledge: how crucial this mutation or that cell-surface receptor really is to cancer growth. These are the findings that launch companies and clinical trials — provided, of course, that they have been published in research papers in peer-reviewed journals."

CHALLENGES: REPRODUCIBILITY

The screenshot shows a web browser displaying an article from the journal *Nature*. The header features the *nature* logo and the text "International weekly journal of science". A navigation bar includes links for Home, News & Comment, Research, Careers & Jobs, Current Issue, Archive, Audio & Video, and Forum. Below the navigation is a breadcrumb trail: Archive > Volume 533 > Issue 7604 > Editorial > Article. The main content title is "Reality check on reproducibility". A sub-headline states: "A survey of *Nature* readers revealed a high level of concern about the problem of irreproducible results. Researchers, funders and journals need to work together to make research more reliable." The publication date is listed as 25 May 2016. At the bottom, there are links for PDF and Rights & Permissions. A sidebar contains a statement about a reproducibility crisis.

NATURE | EDITORIAL

Reality check on reproducibility

A survey of *Nature* readers revealed a high level of concern about the problem of irreproducible results. Researchers, funders and journals need to work together to make research more reliable.

25 May 2016

[PDF](#) [Rights & Permissions](#)

Is there a reproducibility crisis in science? Yes, according to the readers of *Nature*. Two-thirds of researchers who responded to a survey by this journal said that current levels of reproducibility are a major problem.

CHALLENGES: REPRODUCIBILITY

nature
International journal of science

Altmetric: 10 Citations: 1 [More detail >>](#)

Correspondence

Research data: Silver lining to irreproducibility

Alyssa Ward , Thomas O. Baldwin & Parker B. Antin

There is room for improvement in how science is done and reported, but something can often be learned from irreproducible experiments. The situation may not be as dire as some headlines imply.

CHALLENGES: REPRODUCIBILITY

nature International weekly journal of science

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

Archive > Volume 542 > Issue 7642 > Comment > Article

NATURE | COMMENT

No publication without confirmation

Jeffrey S. Mogil & Malcolm R. Macleod

22 February 2017

Jeffrey S. Mogil and Malcolm R. Macleod propose a new kind of paper that combines the flexibility of basic research with the rigour of clinical trials.

PDF Rights & Permissions

Subject terms: Publishing · Medical research



Illustration by David Parkins

CHALLENGES: REPRODUCIBILITY

nature International weekly journal of science

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

Archive > Volume 530 > Issue 7588 > Comment > Article

NATURE | COMMENT Share Print

Reproducibility: A tragedy of errors

David B. Allison, Andrew W. Brown, Brandon J. George & Kathryn A. Kalsner

03 February 2016

Mistakes in peer-reviewed papers are easy to find but hard to fix, report David B. Allison and colleagues.

[PDF](#) [Rights & Permissions](#)

Subject terms: Communication · Publishing · Peer review



Illustration by David Parkins

CHALLENGES: REPRODUCIBILITY

The screenshot shows the homepage of the **nature** journal website. The header features the **nature** logo and the tagline "International weekly journal of science". Below the header is a navigation bar with links to Home, News & Comment, Research, Careers & Jobs, Current Issue, Archive, Audio & Video, and a search bar. A breadcrumb navigation path is visible below the navigation bar, showing the hierarchy: Archive > Volume 536 > Issue 7617 > Editorial > Article. The main content area displays an editorial titled "Go forth and replicate!" by **John Bohannon** (published on 24 August 2016). The article discusses the importance of replication studies in science. At the bottom of the article page, there are links for PDF download and rights & permissions.

nature International weekly journal of science

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

Archive > Volume 536 > Issue 7617 > Editorial > Article

NATURE | EDITORIAL

عربي

Go forth and replicate!

To make replication studies more useful, researchers must make more of them, funders must encourage them and journals must publish them.

24 August 2016

PDF Rights & Permissions

No scientist wants to be the first to try to replicate another's promising study: much better to know what happened when others tried it. Long before replication or reproducibility became major talking points, scientists had strategies to get the word out. Gossip was one. Researchers would compare notes at conferences, and a patchy network would be warned about whether a study was worth building on. Or a vague comment might be buried in a related publication. Tell-tale sentences would start "In our hands", "It is unclear why our results differed ..." or "Interestingly, our results did not ...".

CHALLENGES: REPRODUCIBILITY

The screenshot shows a web page from the journal *nature*. At the top, the *nature* logo is displayed with the subtitle "International weekly journal of science". Below the logo is a navigation bar with links to Home, News & Comment, Research, Careers & Jobs, Current Issue, Archive, Audio & Video, and Forum. A breadcrumb navigation path is shown below the navigation bar: Archive > Volume 523 > Issue 7561 > Research Highlights: Social Selection > Article. The main title of the article is "Collaborate and listen to reproduce research". A sub-headline reads: "Better communication between labs may resolve many reproducibility problems, according to report." The author's name, Chris Woolston, is listed. The publication date is 16 July 2015, with a correction made on 17 July 2015. Below the article title are two buttons: "PDF" and "Rights & Permissions". The main text of the article discusses the challenges faced by cell-biology labs in reproducing research results and highlights a report from the American Society for Cell Biology (ASCB) that suggests improved communication and data quality standards could help. It quotes Arturo Casadevall, a microbiologist and immunologist at Johns Hopkins Bloomberg School of Public Health.

nature International weekly journal of science

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

Archive > Volume 523 > Issue 7561 > Research Highlights: Social Selection > Article

NATURE | RESEARCH HIGHLIGHTS: SOCIAL SELECTION

Collaborate and listen to reproduce research

Better communication between labs may resolve many reproducibility problems, according to report.

Chris Woolston

16 July 2015 | Corrected: 17 July 2015

PDF Rights & Permissions

Cell-biology labs often struggle to reproduce the research results of other groups. But a 15 July report suggests that many of those troubles would vanish if scientists reached out to the original experimenters. The report, released by the American Society for Cell Biology (ASCB), includes survey results from hundreds of ASCB members and calls for changes in scientific culture to make results easier to confirm. Besides better communication, it urges scientists to adopt more-uniform standards within their fields and to focus more on data quality rather than on publishing in high-impact journals. Arturo Casadevall, a microbiologist and immunologist at Johns Hopkins Bloomberg School of Public Health in Baltimore, Maryland, tweeted:

CHALLENGES: REPRODUCIBILITY

nature International weekly journal of science

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

Archive > Volume 521 > Issue 7552 > News Feature > Article

NATURE | NEWS FEATURE

Reproducibility crisis: Blame it on the antibodies

Antibodies are the workhorses of biological experiments, but they are littering the field with false findings. A few evangelists are pushing for change.

Monya Baker

19 May 2015

PDF Rights & Permissions



Illustration by Nik Spencer/Nature

CHALLENGES: REPRODUCIBILITY

nature International weekly journal of science

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

News & Comment > News > 2018 > March > Article

NATURE | NEWS

Over half of psychology studies fail reproducibility test

Largest replication study to date casts doubt on many published positive results.

Monya Baker

27 August 2015

 Rights & Permissions

Don't trust everything you read in the psychology literature. In fact, two thirds of it should probably be distrusted.

In the biggest project of its kind, Brian Nosek, a social psychologist and head of the Center for Open Science in Charlottesville, Virginia, and 269 co-authors repeated work reported in 98 original papers from three psychology journals, to see if they independently came up with the same results.



Brian Nosek's team set out to replicate scores of

CHALLENGES: REPRODUCIBILITY

nature International weekly journal of science

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

News & Comment > News > 2018 > March > Article

NATURE | NEWS

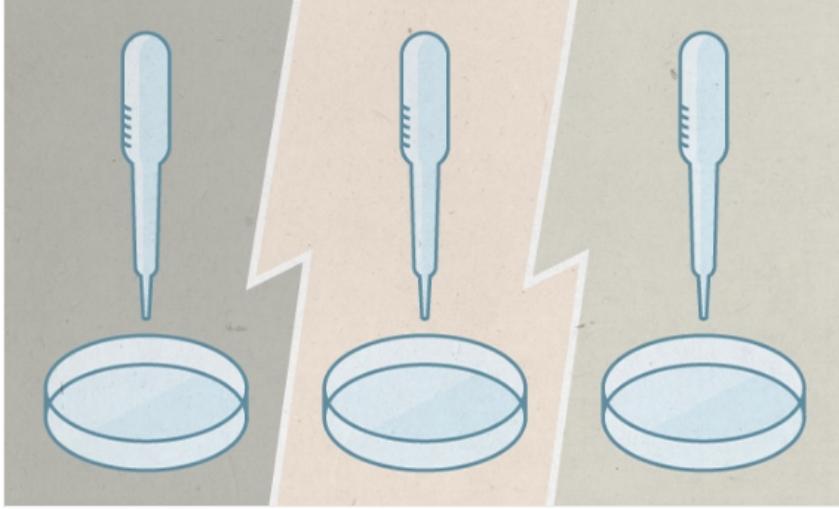
Irreproducible biology research costs put at \$28 billion per year

Study calculates cost of flawed biomedical research in the United States.

Monya Baker

09 June 2015

Rights & Permissions



A graphic illustration showing three test tubes standing upright in petri dishes. The background is divided into three vertical sections of different colors: grey, orange, and beige. Each section contains one test tube and its corresponding petri dish. The test tubes have blue caps and are positioned vertically. The overall image serves as a visual metaphor for the concept of irreproducibility in scientific research.

Robust research: Institutions must do their part for reproducibility

C. Glenn Begley, Alastair M. Buchan & Ulrich Dirnagl

01 September 2015

Tie funding to verified good institutional practice, and robust science will shoot up the agenda, say C. Glenn Begley, Alastair M. Buchan and Ulrich Dirnagl.



PDF



Rights & Permissions

Subject terms: Research management · Institutions · Lab life

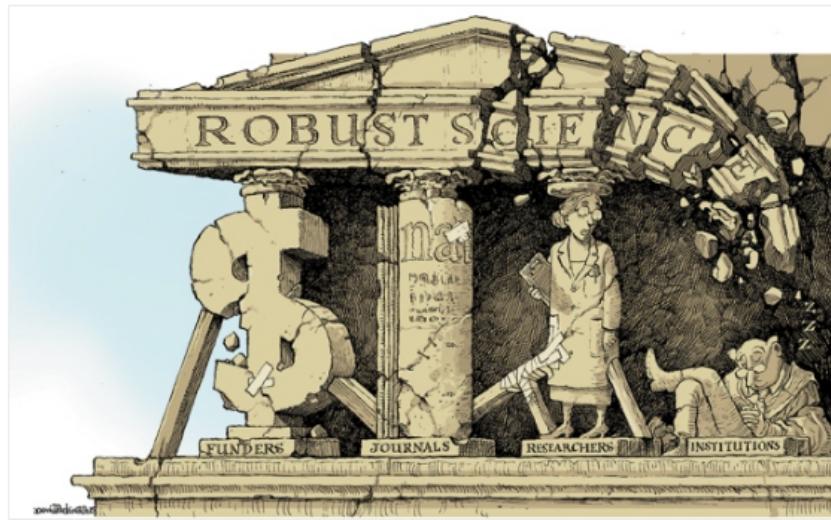


Illustration by David Parkins

CHALLENGES: REPRODUCIBILITY

The screenshot shows a news article from the journal **nature**. The header includes the title "nature International weekly journal of science" and a navigation bar with links to Home, News & Comment, Research, Careers & Jobs, Current Issue, Archive, Audio & Video, and Forum. Below the header is a breadcrumb navigation path: News & Comment > News > 2018 > March > Article. The main headline reads "Dutch agency launches first grants programme dedicated to replication". A sub-headline below it states "Three-year pilot devotes €3 million to verifying other studies." The author is listed as "Monya Baker". The publication date is "20 July 2016 | Updated: 20 July 2016". A "Rights & Permissions" section is indicated by a button. The main text discusses the launch of a national fund for replication studies by the Netherlands Organisation for Scientific Research (NWO), providing a pot of €3 million over three years. It quotes Brian Nosek, executive director of the Center for Open Science, who notes that this is a tiny fraction of NWO's budget but an important step.

nature International weekly journal of science

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

News & Comment > News > 2018 > March > Article

NATURE | NEWS

Dutch agency launches first grants programme dedicated to replication

Three-year pilot devotes €3 million to verifying other studies.

Monya Baker

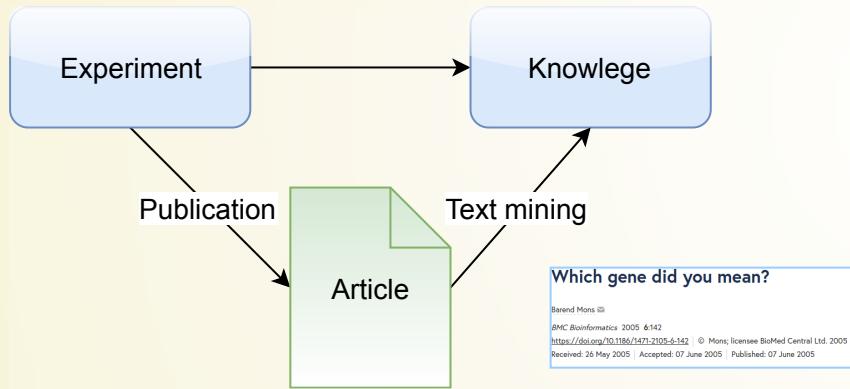
20 July 2016 | Updated: 20 July 2016

Rights & Permissions

The Netherlands has launched what researchers say is the world's first national fund dedicated to replication studies: a pot of €3 million (US\$3.3 million) over the next 3 years for Dutch scientists to test whether they can reproduce important research results in social and medical sciences.

The pilot programme was announced on 19 July by the Netherlands Organisation for Scientific Research (NWO), the country's largest research-funding agency. It marks a tiny fraction of the agency's €700-million annual budget, but is an important step, says Brian Nosek, executive director of the Center for Open Science in Charlottesville, Virginia. "If my calculations are correct, this is an increase of infinity per cent of federal funding dedicated to replication studies," he says.

CHALLENGES: "KNOWLEDGE BURYING"



More than 40% of information lost from publishing to mining

RIP (Rest in Paper) knowledge

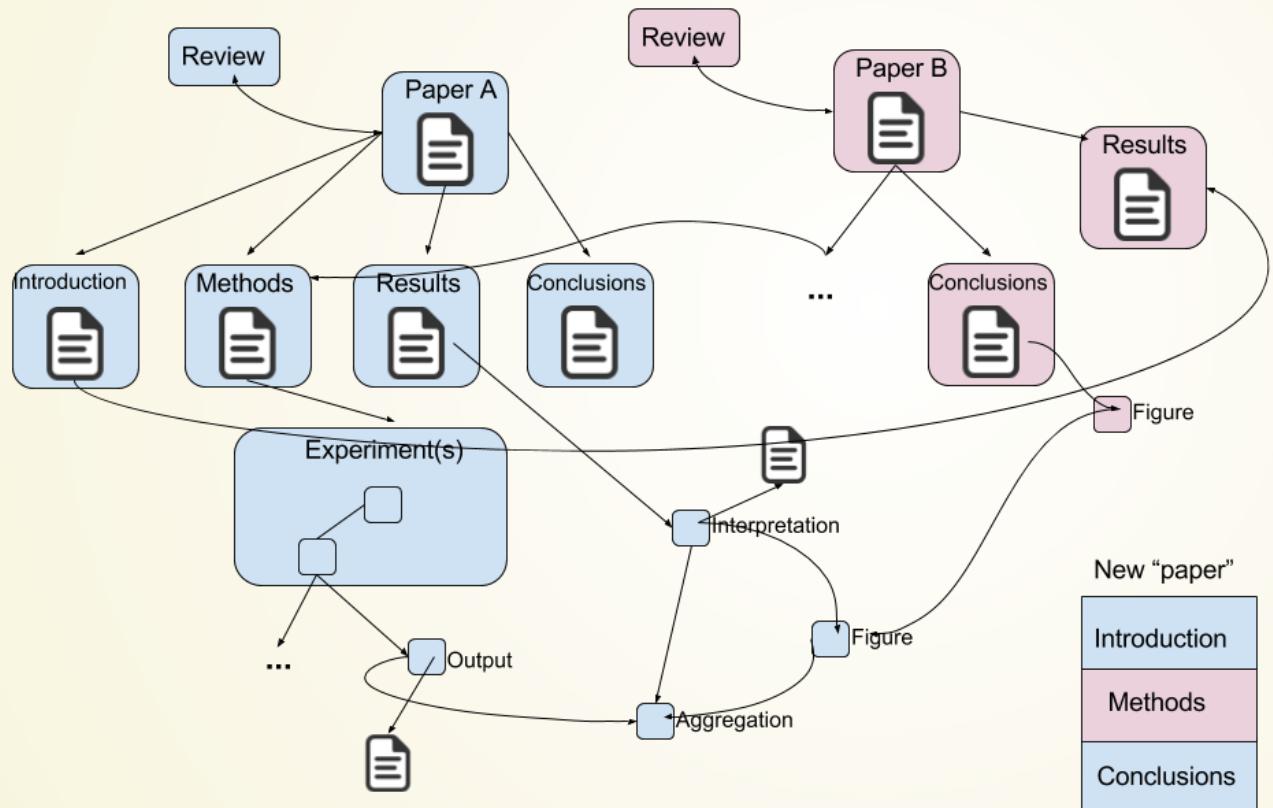
Need to preserve information about the processes

CHALLENGES: DIGITAL PUBLISHING

- Semantic Web, Linked Data
- Understandable content for both humans and machines
- **Digital artifacts**: text, datasets, code, figures, slides, spreadsheets, multimedia objects, methods, protocols, results and reviews, annotations, etc.
- **Scientific workflows**: how digital artifacts are consumed and produced

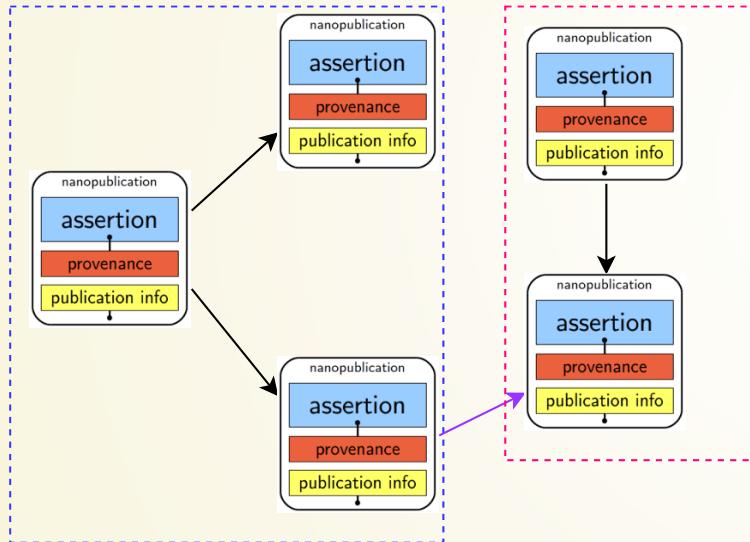
CHALLENGES: DIGITAL PUBLISHING

- Semantic Web, Linked Data



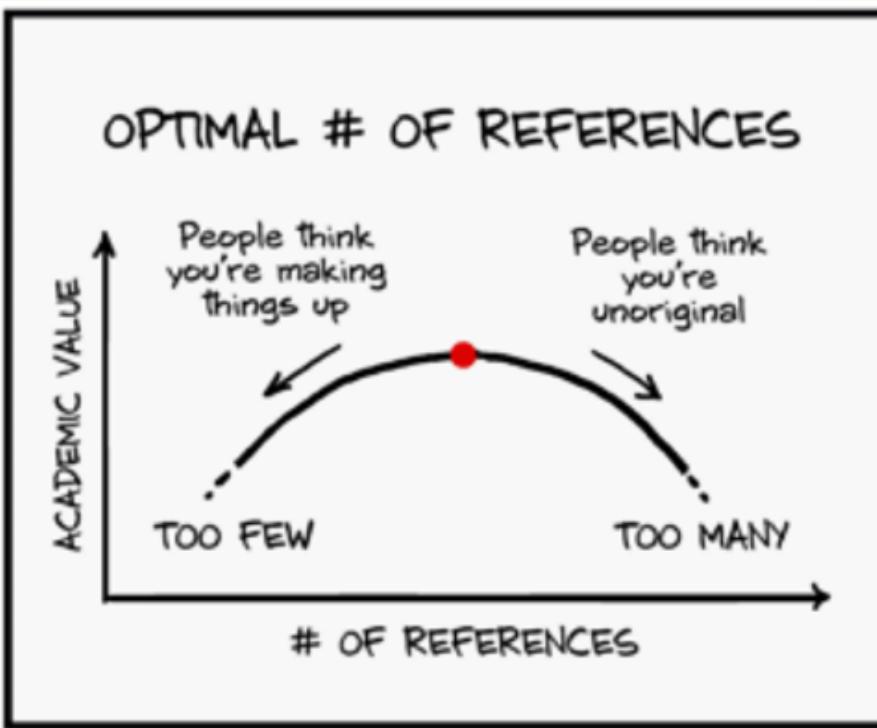
CHALLENGES: DIGITAL PUBLISHING

- Semantic Web, Linked Data



Fine-grained interconnected parts of knowledge; e.g. nanopublications

CHALLENGES: QUALITY ASSESSMENT



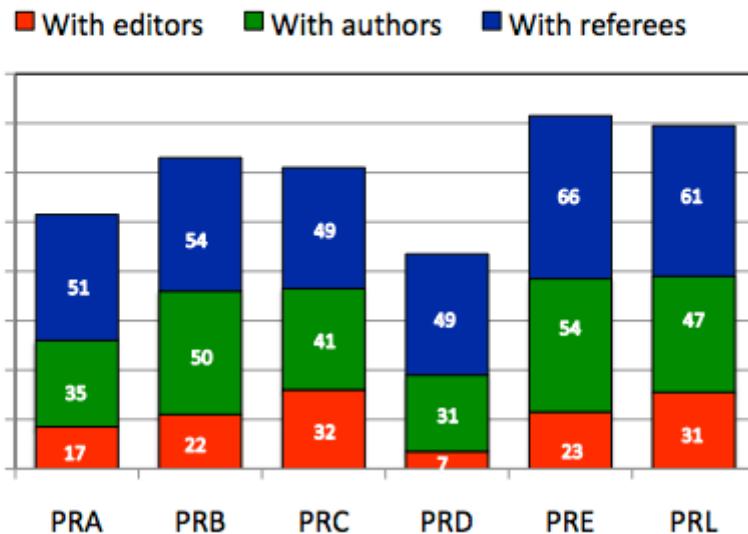
CHALLENGES: QUALITY ASSESSMENT

- We need a new way of assessing the quality of scientific publications
- Debatable and bias-able indicators; e.g. [Journal Impact Factor](#)
 - assessing assessment
 - "not-so-deep" impact
 - science metrics not so scientific

CHALLENGES: DECENTRALIZATION



Time consumption during the review process (2008)



*Editors are from Mars,
Referees are from Venus, and
Authors are from Earth.*

No control of a central authority anymore; e.g. [publishing house](#)

Decentralization of content; e.g. [academic peer-to-peer networks](#)

CHALLENGES

- Reproducibility
- "Knowledge Burying"
- Digital publishing
- Quality assessment
- Decentralization of content

FOCUS

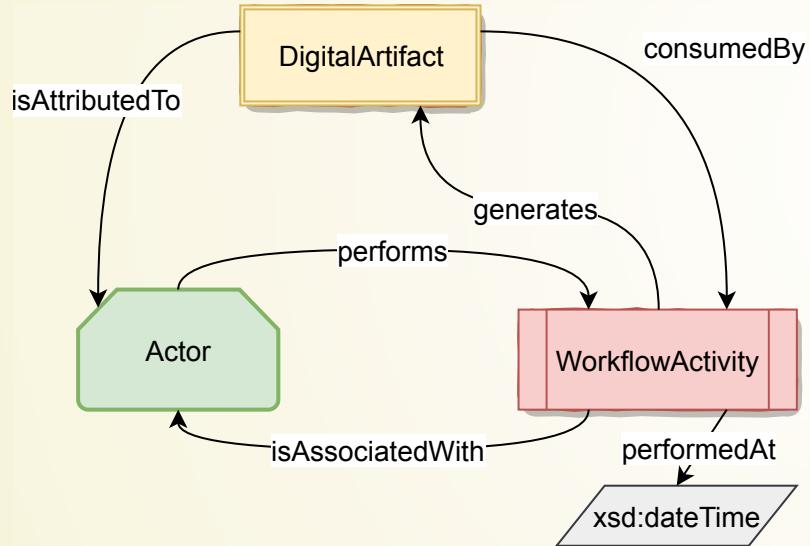
How can *scientific workflows* that produce and consume digital artifacts be **assessed, linked** and **decentrally executed** across platforms, such that **individual steps** of a single workflow can be **distributed**?

APPROACH

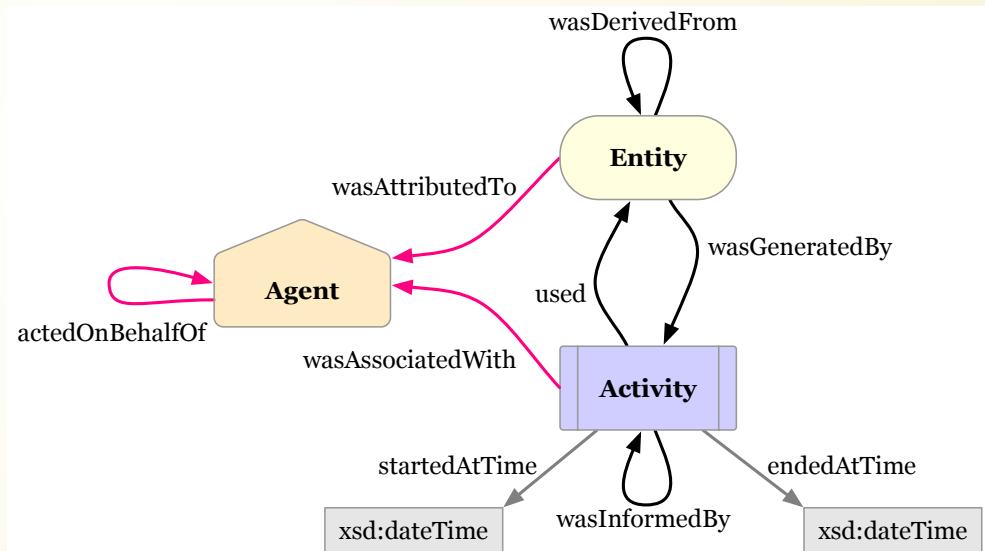
- Develop a **simple model** for linked workflows → **linkflows**
- Choose **sample articles** from IOS Press and Netherlands Sound and Vision
- **Instantiate model** with sample articles
- Use **Linked Data Notifications (LDNs)** for **decentralized execution**

LINKFLOWS: MODEL

"New" model (v.0.1 alpha)



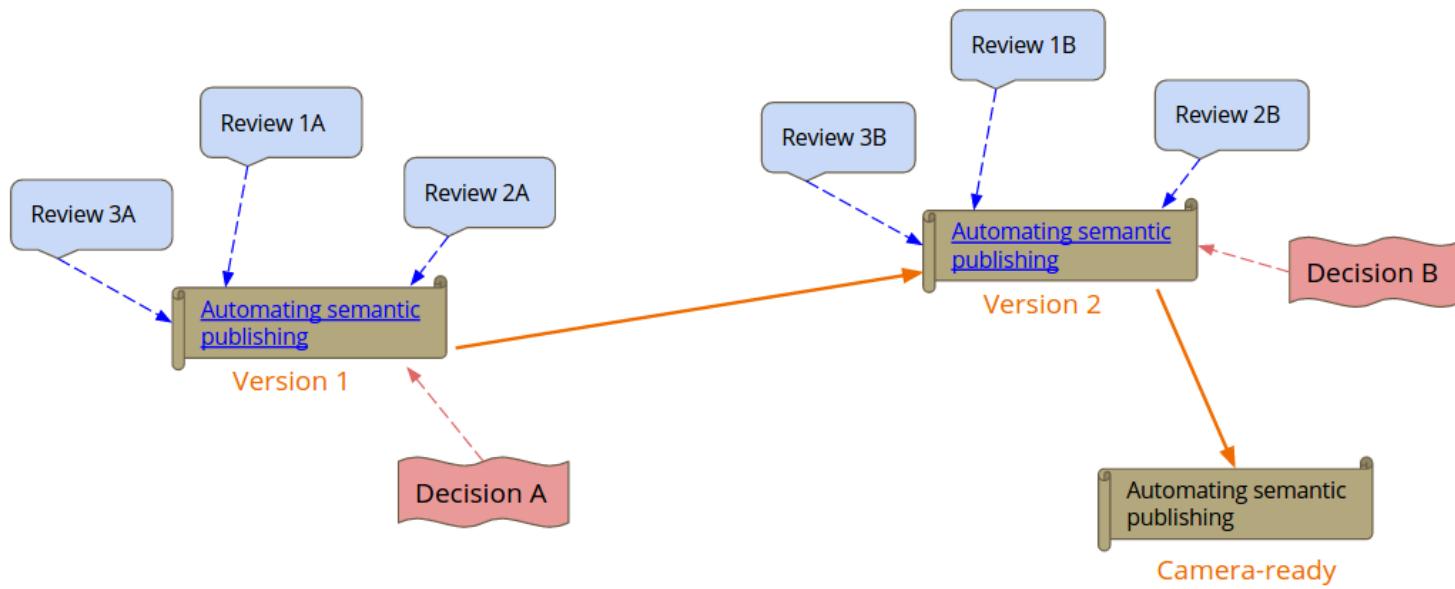
PROV-O



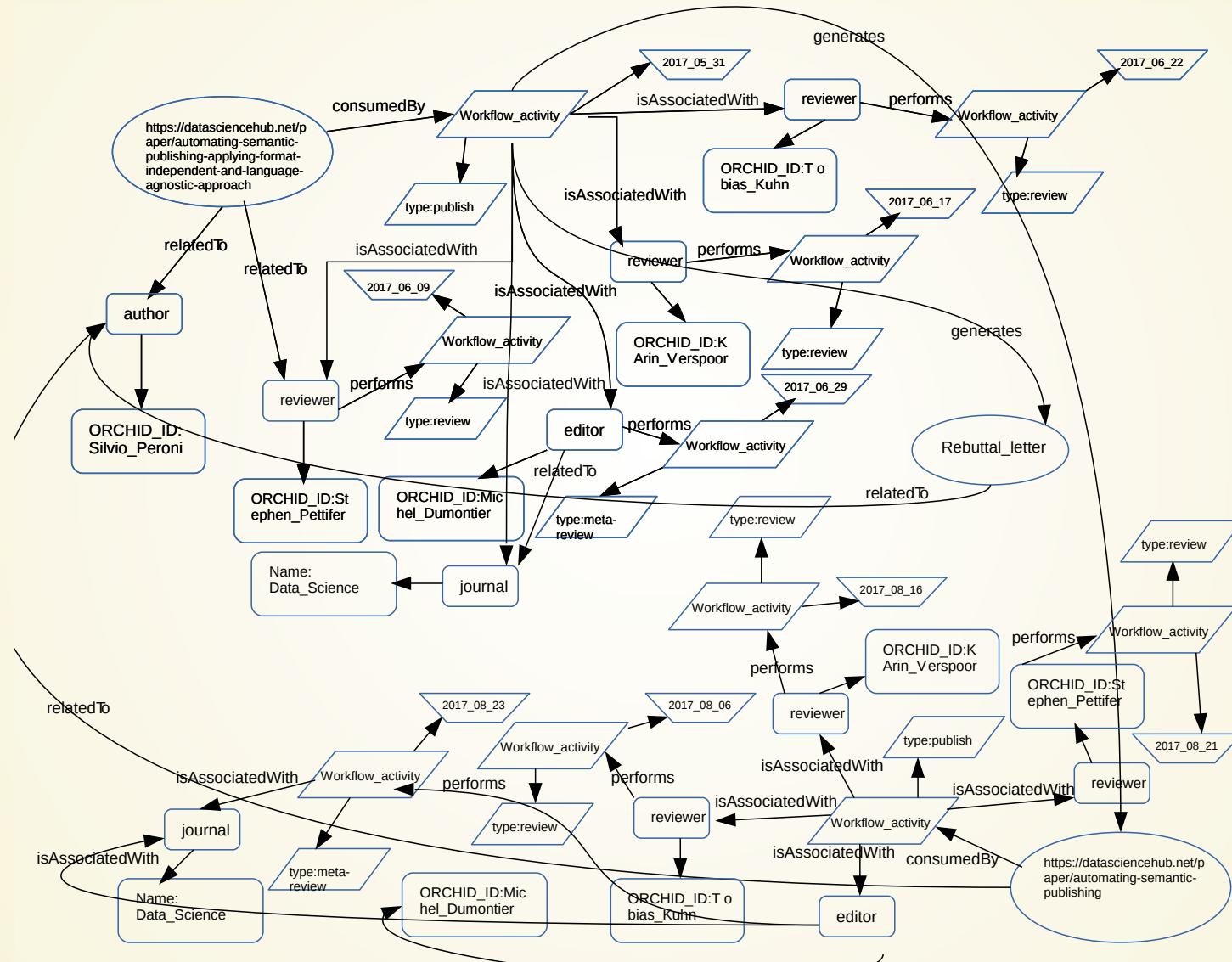
LINKFLOWS: USE CASES

- **IOS Press**: publishing house based in Amsterdam
 - Data Science Journal
 - Journal of Alzheimer's Disease Reports
- **Netherlands Sound and Vision**, the Netherlands cultural archive, has over 70% of the Dutch audio-visual heritage
 - VIEW Journal of European Television History and Culture
 - Tijdschrift voor Mediageschiedenis (Magazine for Media History)
 - Mind of the Universe videos

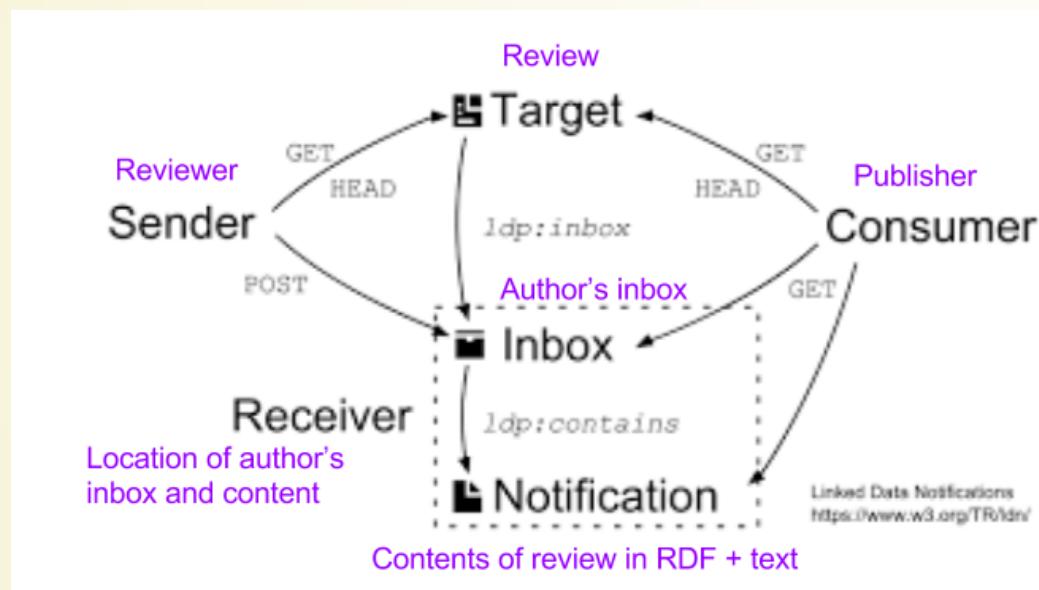
Paper workflow example



Model instantiation with example paper (v.0.1 alpha)



REVIEWING WORKFLOW USING LDNS



1. submit article/snippet
2. invite reviewer
3. handshake with reviewer
4. submit review
5. use review for publishing

Example content of a reviewing notification

HTTP/1.1 200 OK

Content-Type: text/turtle

```
@prefix schema: <http://schema.org/> .  
[ a schema:ReviewAction;  
  schema:agent [  
    a schema:Person;  
    schema:name "Stephen Pettifer"  
  ];  
  schema:object <https://datasciencehub.net/system/files/ds-paper-511.zip>  
  schema:result [  
    a schema:Review;  
    schema:reviewBody "This is an interesting and somewhat controversial..."  
  ]  
] .
```

LINKFLOWS: LINKED WORKFLOWS

- Represent scientific knowledge as **nodes in a network**
- Approach **compatible** with "traditional" PDFs, SPAR-represented articles, nanopublications, etc.
- **Decentralized peer-review process**
- New measure for **quality assessment**; e.g. network analysis

