Scientific Publishing and *Science*

Brooks Hanson, Deputy Editor, Physical Sciences, *Science*

The *Daubert* standard is a rule regarding the admissibility of expert witnesses' testimony in legal proceedings

The Court defined "scientific methodology" as the process of formulating hypotheses and then conducting experiments to prove or falsify the hypothesis, and provided a nondispositive, nonexclusive, "flexible" test for establishing its "validity":

1. Empirical testing: the theory or technique must be falsifiable, refutable, and testable.
2. **Subjected to peer review and publication**.
3. Known or potential error rate.
4. The existence and maintenance of standards and controls concerning its operation.
5. Degree to which the theory and technique is generally accepted by a relevant scientific community.

IPCC Mandate

The IPCC mandate is to assess, on a comprehensive, objective, open and transparent basis, the available scientific information in peer-reviewed literature.

Shelby Amendment/Data Quality Act

Codified importance of peer review in justifying government regulations.

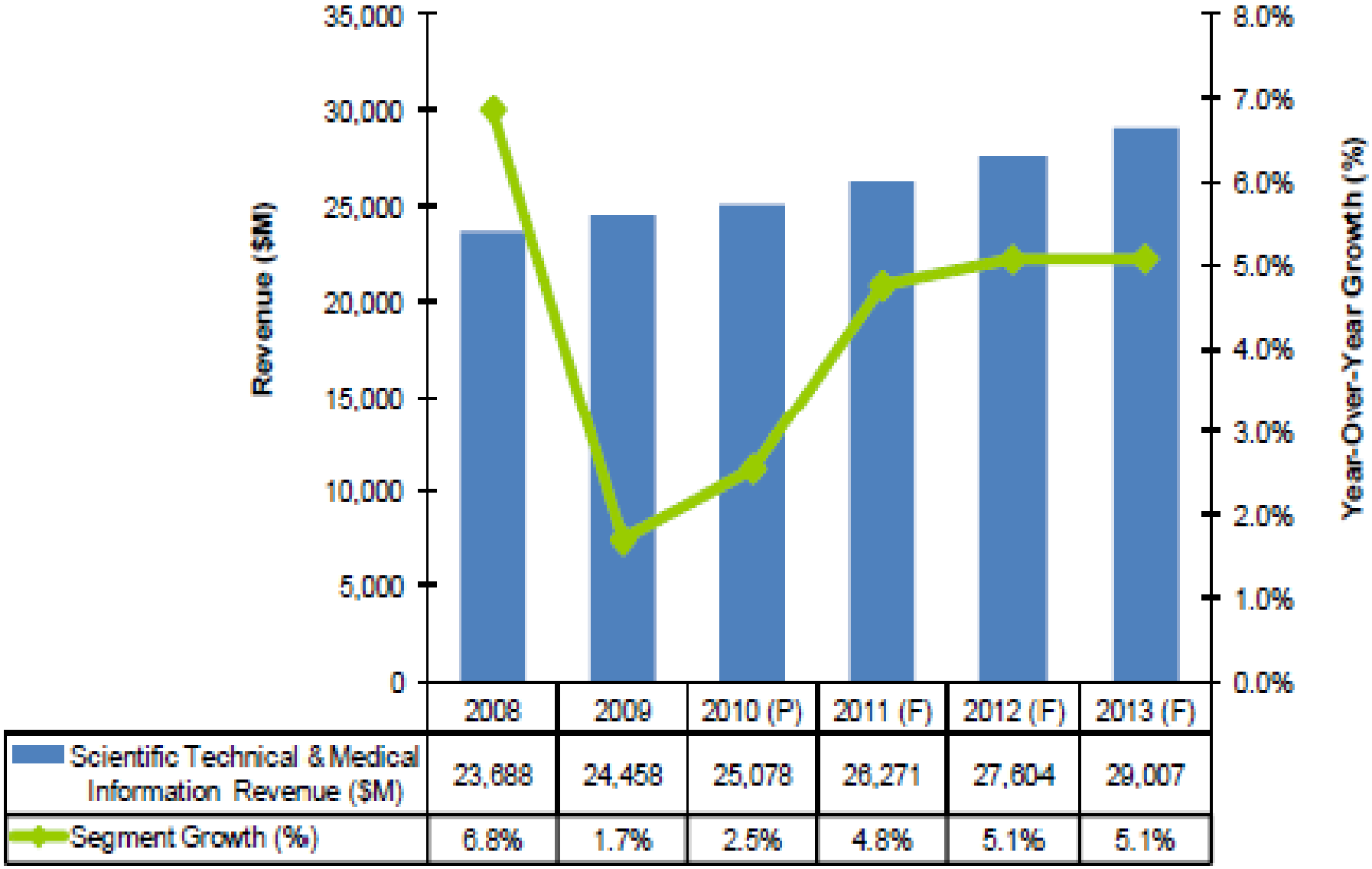
Peer review and scientific publication is a shared societal responsibility among researchers, reviewers, and publishers. It is a fundamental aspect of the integrity and accountability of science, as well as it's advancement.

Data, and long-term data sets are increasingly important in peer reviewed papers.

Scientific, Technical, and Medical (STM) Publishers

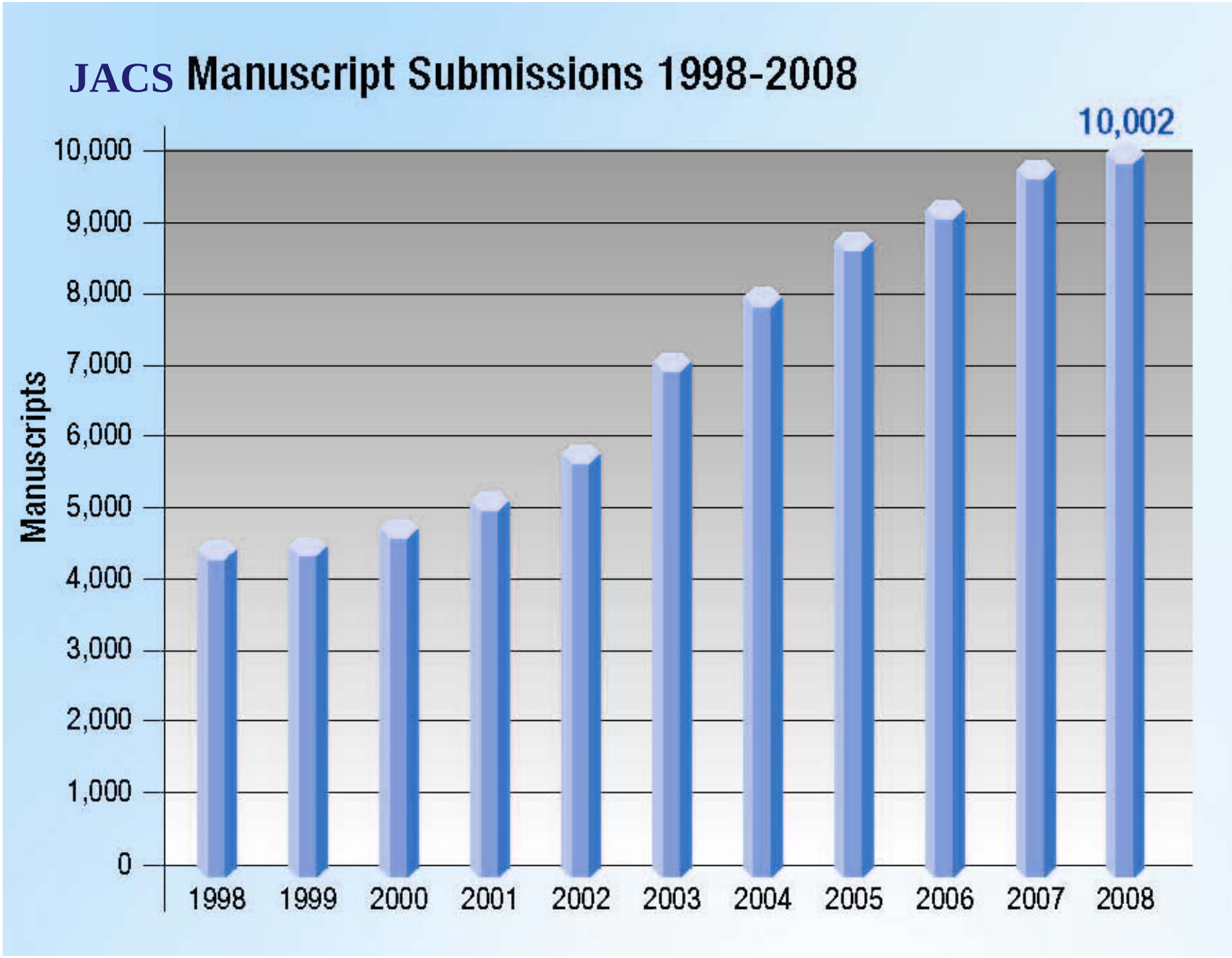
- 24,500 active scholarly peer reviewed titles
- Published by over 2000 journal publishers
- Publishing 1.5 million peer reviewed articles per year.

Source: STM Association

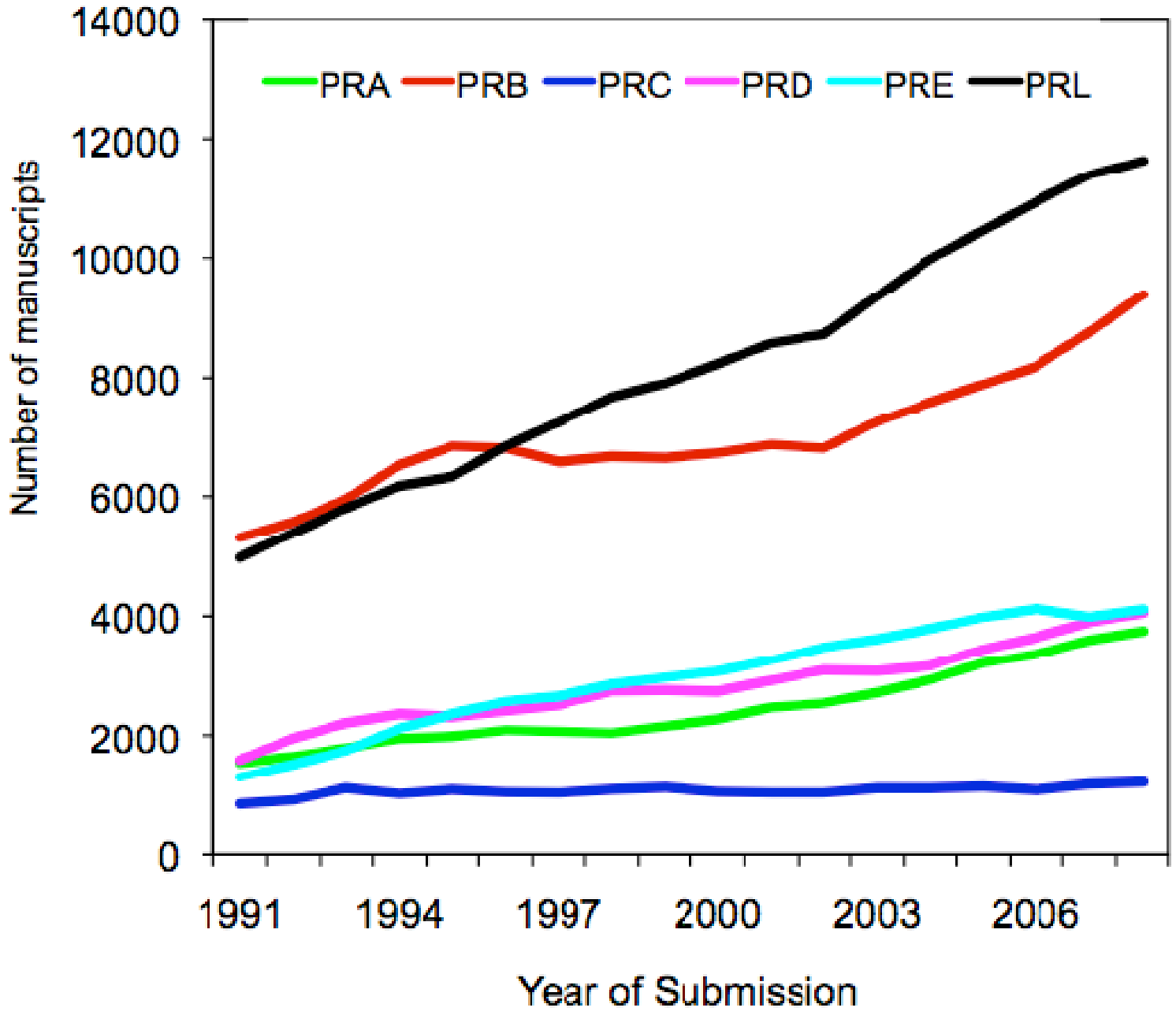


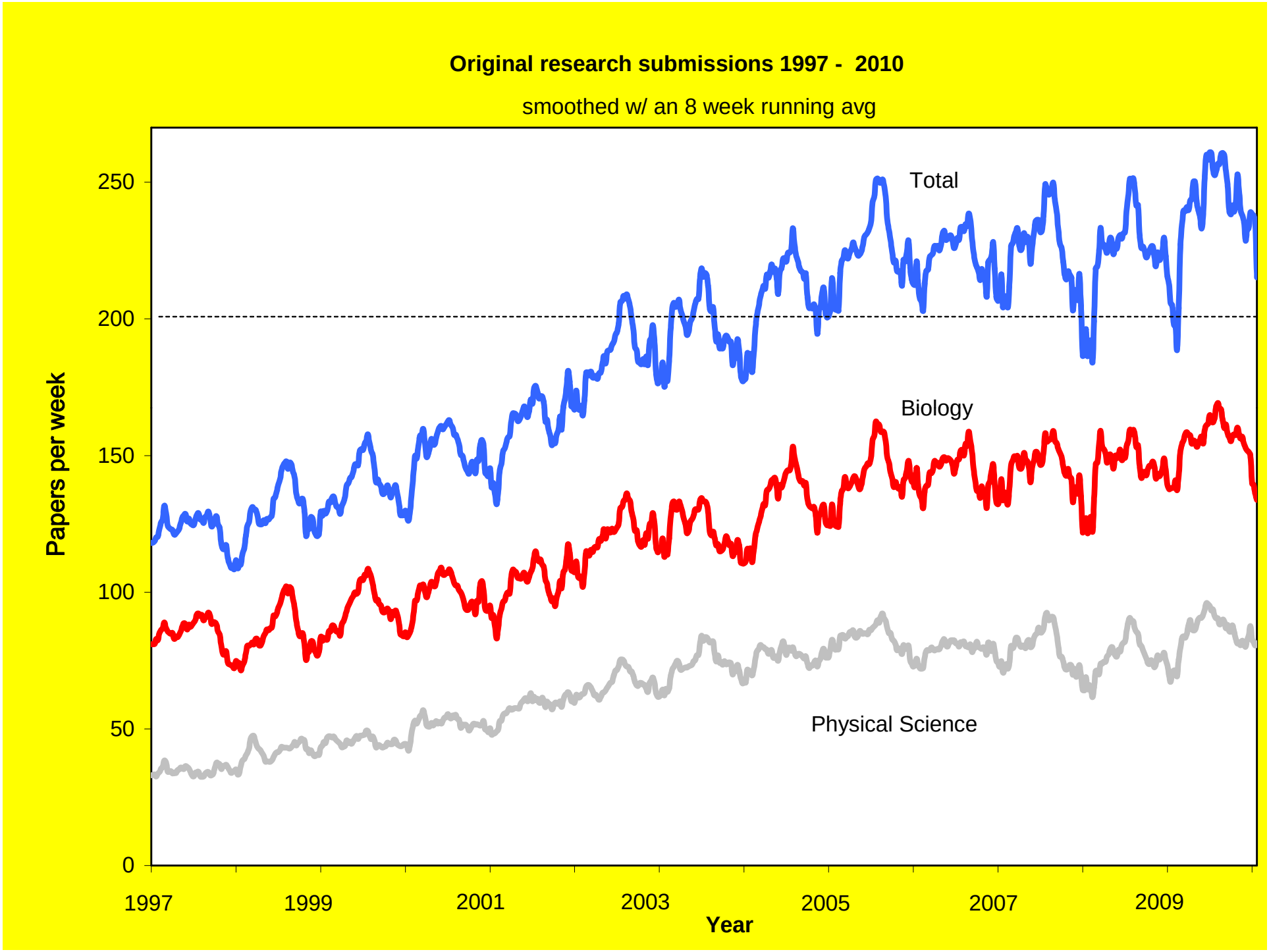
	2008	2009	2010 (P)	2011 (F)	2012 (F)	2013 (F)
Scientific Technical & Medical Information Revenue ($M)	23,688	24,458	25,078	26,271	27,604	29,007
Segment Growth (%)	6.8%	1.7%	2.5%	4.8%	5.1%	5.1%

Source: Outsell Research


JACS Manuscript Submissions 1998-2008
10,002
Manuscripts
10,000
9,000
8,000
7,000
6,000
5,000
4,000
3,000
2,000
1,000
0
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008

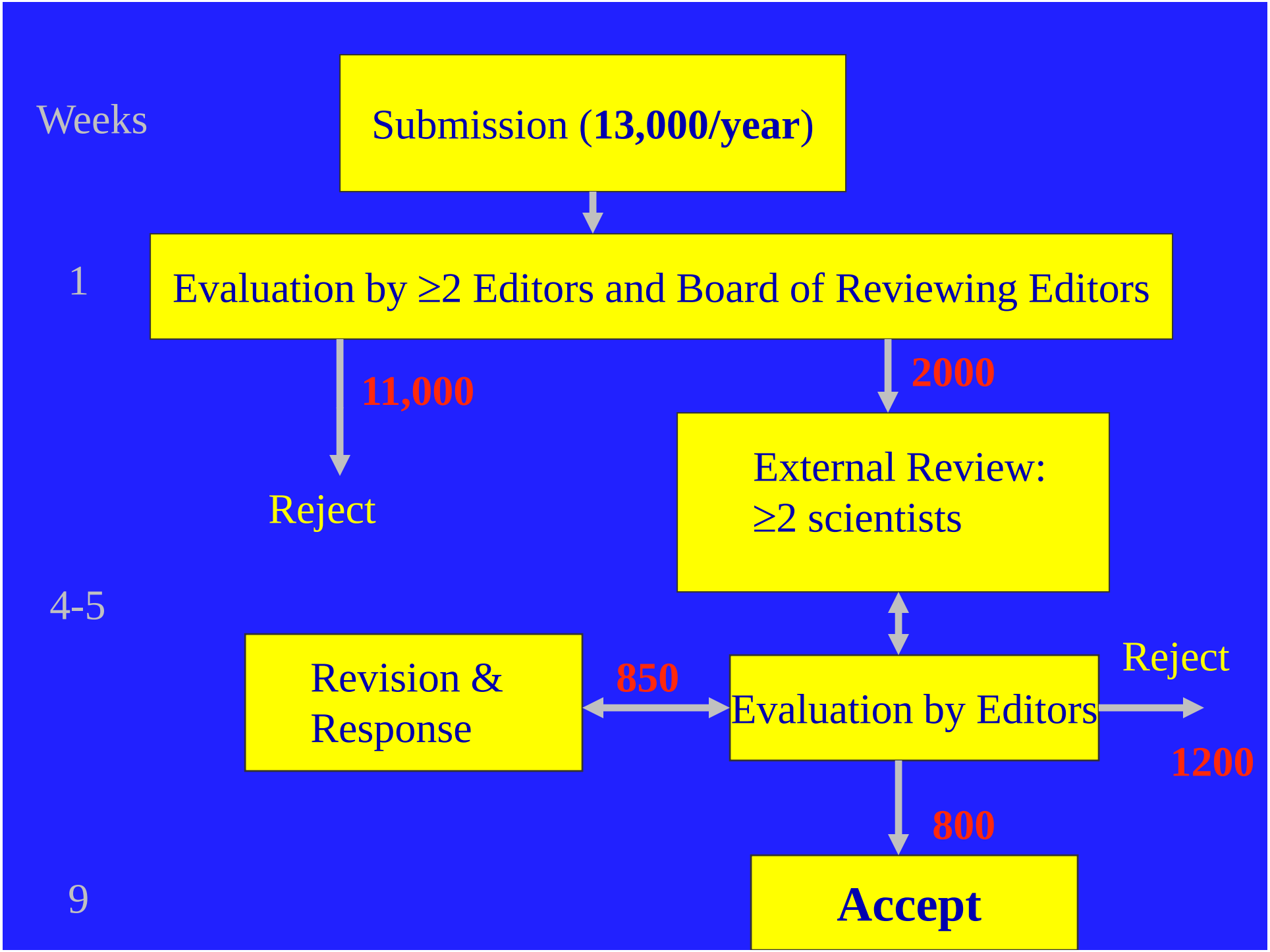
Submissions to APS journals




Original research submissions 1997 - 2010
smoothed w/ an 8 week running avg
Papers per week
250
200
150
100
50
0
1997
1999
2001
2003
2005
2007
2009
Year
Total
Biology
Physical Science

Science's Mission Statement

Science seeks to publish those papers that are most influential in their fields and that will significantly advance scientific understanding. Selected papers should present novel and broadly important data, syntheses, or concepts. They should merit the recognition by the scientific community and general public provided by publication in *Science*, beyond that provided by specialty journals.


Weeks
Submission (13,000/year)
1
Evaluation by ≥2 Editors and Board of Reviewing Editors
11,000
Reject
2000
External Review:
≥2 scientists
4-5
Revision &
Response
850
Evaluation by Editors
Reject
1200
800
9
Accept

Major Evolution in Policy/Practice

- Electronic submission only.
- All co-authors notified upon manuscript submission, to check authorship.
- Detailed authorship and conflict-of-interest disclosure before acceptance by all authors.
- All figures checked at revision for inappropriate adjustments
- No MTA's or restrictions on data.
- No unpublished data allowed. All references/data must be available at the time of publication.

Authorship

All authors have agreed to be so listed and have seen and approved the manuscript, its content, and its submission to Science. **We will verify this requirement with all co-authors upon submission. Non-compliance may lead to rejection.** Any further change in authorship must be approved in writing by all authors. Before acceptance, each author will be required to indicate their role in the research and verify that they meet our criteria for authorship, and the senior author from each lab must assert that they have seen and checked the reported raw data.

Notice to all authors upon submission

Dear Dr. XXXXX

You are listed as a coauthor on the above manuscript, which has recently been submitted to Science. According to Science policy, all authors must have seen and approved the submission of their manuscript. If you have seen the manuscript and approved its submission, no action is necessary.

If you have not read this paper and do not approve its submission to Science, please let us know as soon as possible. Please refer to the manuscript number listed above in any correspondence (you can just reply to this message).

Some representative responses:

- I had indicated to Dr. XXXXX that it was not necessary to include me as an author as all we did was provide a reagent.
- I cannot approve of this paper as I have not read it and as I did not know of its existence. Could you send me a copy of this text?
- Please remove my name from the list of authors.
- I do not approve of the submission of manuscript number 11xxxxx. I had serious issues with a prior version, which were not addressed in the present version.
- I have transmitted some concerns about Manuscript Number 11xxxx to my lawyer.

Evolution of data archiving policies

- Pre-1990, limited. Genbank started in 1982. Protein crystal structures in 1971; expanded 1998.
- 1993--AGU began encouraging Supporting material.
- 1995--Journals begin to go online. Most are online by 2000.
- 2000--Most journals are hosting some online supplements.
- 1990-2010—Data archiving and sharing policies implemented and strengthened.
- NISO standards meeting for Supplementary Material (2010).

Science
AAAS

Data must be available—in SOM or archived.

Data and materials availability All data necessary to understand, assess, and extend the conclusions of the manuscript must be available to any reader of Science. After publication, all reasonable requests for materials must be fulfilled. Any restrictions on the availability of data or materials, including fees and original data obtained from other sources (Materials Transfer Agreements), must be disclosed to the editors upon submission. Fossils or other rare specimens must be deposited in a public museum or repository and available for research.

No unpublished data or reference

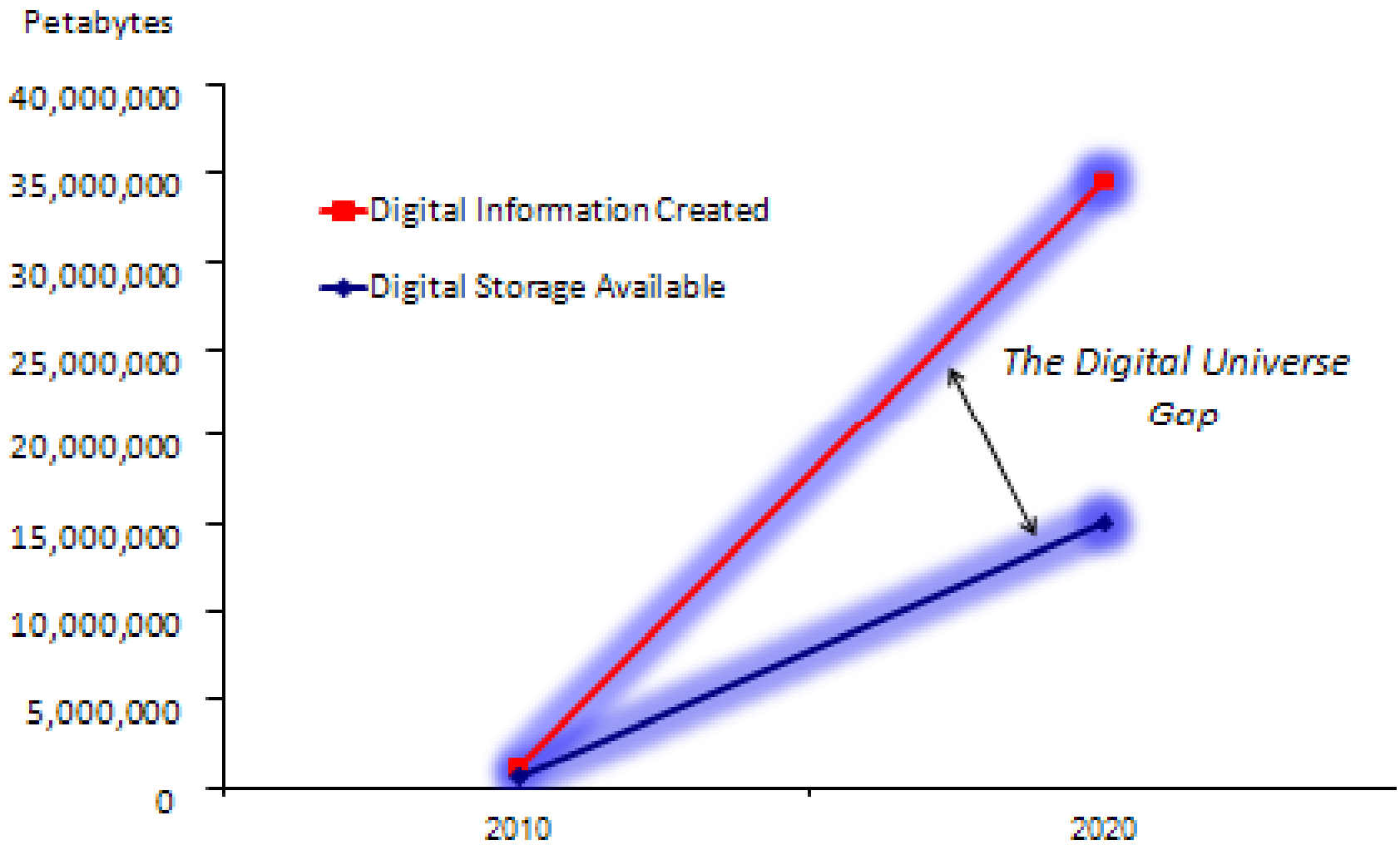
Unpublished data and personal communications Citations to unpublished data and personal communications cannot be used to support claims in a published paper. Papers will be held for publication until all "in press" citations are published.

Major Problems

- Some papers now have huge SOM files (terabytes) that have no reliable archive.
- Journal supplements in PDF format generally, though some other formats used. Very difficult to search.
- Few metadata, if any. Supplements are poorly organized.
- Much data in some labs are discarded because of amount and non-availability of archives and difficulty of transfer.
- Increasing burden on peer review.
- Funding/support of databases is still limited and curation is expensive
- Much private data (particularly in the social sciences/health sciences)

Figure 5: The Emerging Gap

Information Creation > Storage Available



Source: IDC Digital Universe Study, sponsored by EMC, May 2010

Nature **435**, 1010-1011 (23 June 2005) | doi: 10.1038/4351010a

Special Report: Databases in peril

Zeeya Merali and Jim Giles

Nature contacted 89 databases operating in 2000, and more than half said they are now struggling financially. Seven databases have folded.

Nature 462, 258-259 (2009) | doi:10.1038/462258b

Plant genetics database at risk as funds run dry

National Science Foundation to cut support for Arabidopsis resource.

Alison Abbott

The world's most valued plant database faces extinction because its funding is being phased out by the US National Science Foundation (NSF), and no alternative source is on the horizon.

What needs to be done

- We must meet our responsibilities. NSF is requiring data curation plans as part of grant proposals.
- Databases and data infrastructure need long-term support.
- Long-term data collection efforts need long-term support.
- Communities/societies must develop metadata and archiving standards. Journals look to these for guidance.
- Journals must also standardize, set standards, and enforce policies.

PARLIAMENT SCIENCE AND TECHNOLOGY COMMITTEE

Select Committee Announcement 28 July 2011

The Science and Technology Committee today concludes that in order to allow others to repeat and build on experiments, researchers should aim for the gold standard of making their data fully disclosed and made publicly available.

ENSURING THE INTEGRITY, ACCESSIBILITY, AND STEWARDSHIP OF RESEARCH DATA IN THE DIGITAL AGE

Committee on Ensuring the Utility and Integrity of Research Data in a Digital Age

Committee on Science, Engineering, and Public Policy

THE NATIONAL ACADEMIES PRESS

Science
AAAS

Plan when you start a research project!

- Authorship, credit, and responsibility—in writing.
- Data archiving and availability
- Material availability
- Patents
- Related work by all authors
- Plan for publication once you have results.