

# Physics Journals and the arXiv

What is Myth and What is Reality?

Tim Ingoldsby  
The Evidence on Open Access  
CSE 2009 Meeting, Pittsburgh

# Physics Preprint Distribution pre-Ginsparg

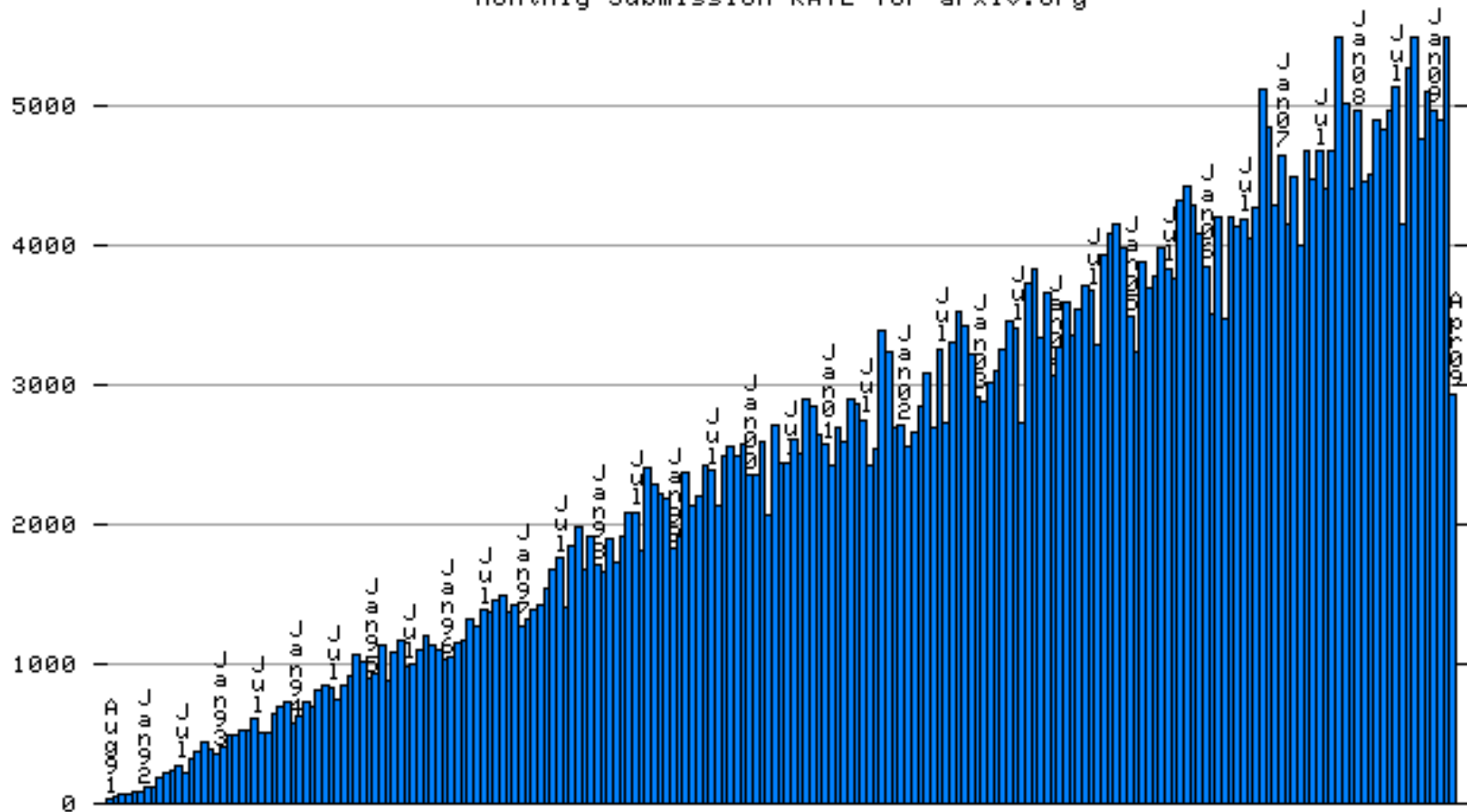
- Paper-based process, mostly confined to theoretical physics
- Distribution for most part confined to major research institutions (or to institutions of research colleagues)
- Long cycle times (postal mail delivery of comments)

# Improvements Made by Paul Ginsparg

- In 1991 Ginsparg implemented an automated deposit ftp server for high energy theoretical physics research pre-prints
  - Rapidly evolved to include (some) other physics, mathematics, and astronomy sub-fields
  - Web interface added in 1994 – 95
  - First received NSF funding in 1996
- No longer confined to just major research institutes – anyone anywhere in the world with Internet connectivity could download any submitted pre-print
- For this work, Ginsparg was awarded a MacArthur Fellowship in 2002
- Deposits surpassed 500,000 in October 2008

# arXiv.org Monthly Submission Rate

Monthly Submission RATE for arxiv.org



First 17.8 years (19 Apr '09 total = 533,598)

# Assertions About the arXiv

“Of course, all *Physical Review Letters* papers are contained in the arXiv”

Jack Sandweiss, Editor, *Physical Review Letters*

“If 14 years of concerted self-archiving in physics have not yet hurt the physics publishers ...”

Stevan Harnad, American Scientist OA Forum

# Research Process

- Select sufficient number of articles from all AIP journals and selected APS journals to produce reliable results.
  - Articles should come from issues published at least 12 months ago to enable all arXiv.org processes to have played out.
  - About 2,800 AIP articles and 2,100 APS articles were studied
- Search arXiv.org to determine if preprint of each article exists.
  - Also used SPIRES (HEP Physics Literature Database) to improve results for HEP
- Analyze results and issue findings.

# Research Process Timeline

- Original process carried out in Summer, 2008
  - Initial searches performed by summer intern
  - All unsuccessful searches repeated by a trained physicist
- Research completed by end of April 2009
- Process design and analysis supported by AIP Statistical Research Center's Roman Czujko (Director)

# APS Journal Content in arXiv.org

Journal Title	Articles in arXiv
<b><i>Physical Review Letters</i></b> Short, important papers from all branches of physics	<b>55.0%</b>
<b><i>Physical Review B</i></b> Condensed Matter and Materials Physics	<b>40.0%</b>
<b><i>Physical Review D</i></b> Particles, Fields, Gravitation, and Cosmology	<b>97.1%</b>

**n = 0.05**

# AIP Journal Content in arXiv.org

<b>Journal Title</b>	<b>Content in arXiv</b>
<i>Applied Physics Letters</i>	4.64%
<i>Journal of Applied Physics</i>	2.43%
<i>Chaos</i>	17.16%
<i>Journal of Chemical Physics</i>	7.18%
<i>Journal of Mathematical Physics</i>	53.57%
<i>Physics of Fluids</i>	3.94%
<i>Physics of Plasmas</i>	2.67%
<i>Review of Scientific Instruments</i>	4.23%

**n = 0.10**

# PRL Content by Sub-Field

<b>Sub-Field Description</b>	<b>Content in arXiv</b>	<b>n</b>
General Physics	82.8%	0.09
Gravitation & Astrophysics	90.9%	0.21
Elementary Particles & Fields	100.0%	0.15
Nuclear Physics	66.7%	0.19
Atomic, Molecular, & Optical	20.8%	0.13
Nonlinear & Fluid Dynamics	32.3%	0.12
Plasmas & Beams	10.5%	0.15
Condensed Matter	49.3%	0.05
Soft Matter (Biological Physics)	42.9%	0.11

# American PhDs by Work Subfield

<b>Physics Sub-Field</b>	<b>AIP Survey</b>	<b>NSF Data</b>
Condensed Matter	21.5%	22.0%
Atomic, Molecular, Optical	10.4%	13.4%
Chemical Physics	8.2%	--
General Physics	14.6%	32.3%
Astronomy & Astrophysics	11.6%	11.0%
Soft Matter, Medical Physics	11.0%	--
Elementary Particles & Fields	5.0%	11.3%
Nuclear Physics	4.6%	5.6%
Plasmas, Beams, & Fluids	8.2%	4.5%
Acoustics	4.9%	--

# Research Conclusions

- The percentage of research articles deposited to arXiv.org varies greatly by physics sub-field
  - Lowest: plasmas & beams; atomic, molecular & optical
  - Highest: elementary particles & fields; gravitation & astrophysics
- Only for a narrow range of sub-fields, representing at most 15% - 20% of physics, can it be said that the arXiv provides comprehensive coverage

# Additional Observations

- Physicists are not particularly diligent about updating their arXiv.org deposits to indicate that a manuscript has been published (less than one in three do so for AIP journals)
  - A small portion of arXiv content is not deposited until after the paper has been published in a peer-reviewed journal
- Journal editors change titles of manuscripts more frequently than might be expected (making it more difficult to identify which articles are deposited to arXiv.org)

# Areas for Further Research

- Studies of additional physics journals to determine % content in arXiv.org
  - *Nature* and *Science* (physics content), Nature specialist journals, other AIP Member Society journals, IoP journals, Elsevier journals, etc.
- Global research on Ph.D.s awarded in each physics sub-field

# For Additional Information

[tingoldsby@aip.org](mailto:tingoldsby@aip.org)

[www.aip.org](http://www.aip.org)

[scitation.aip.org](http://scitation.aip.org)