

*Citation Services for  
Institutional Repositories:  
Citebase Search*

**Tim Brody**

**Intelligence, Agents, Multimedia Group  
University of Southampton**





# Content

- The Open Access Literature
- Why Open Access?
- Citation Services for IRs
  - Distributed Archives
  - Citebase Search
- Effect of Open Access



12-18 Months

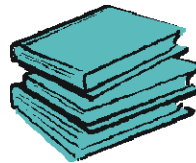
Impact cycle begins:  
Research is done



Researchers write pre-refereeing  
"Pre-Print"



Submitted to Journal



Pre-Print reviewed by Peer Experts - "Peer-Review"



Pre-Print revised by article's Authors

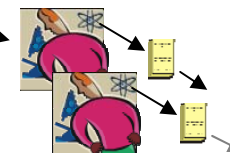
Refereed "Post-Print" Accepted, Certified, Published by Journal



Researchers can access the Post-Print if their university has a subscription to the Journal



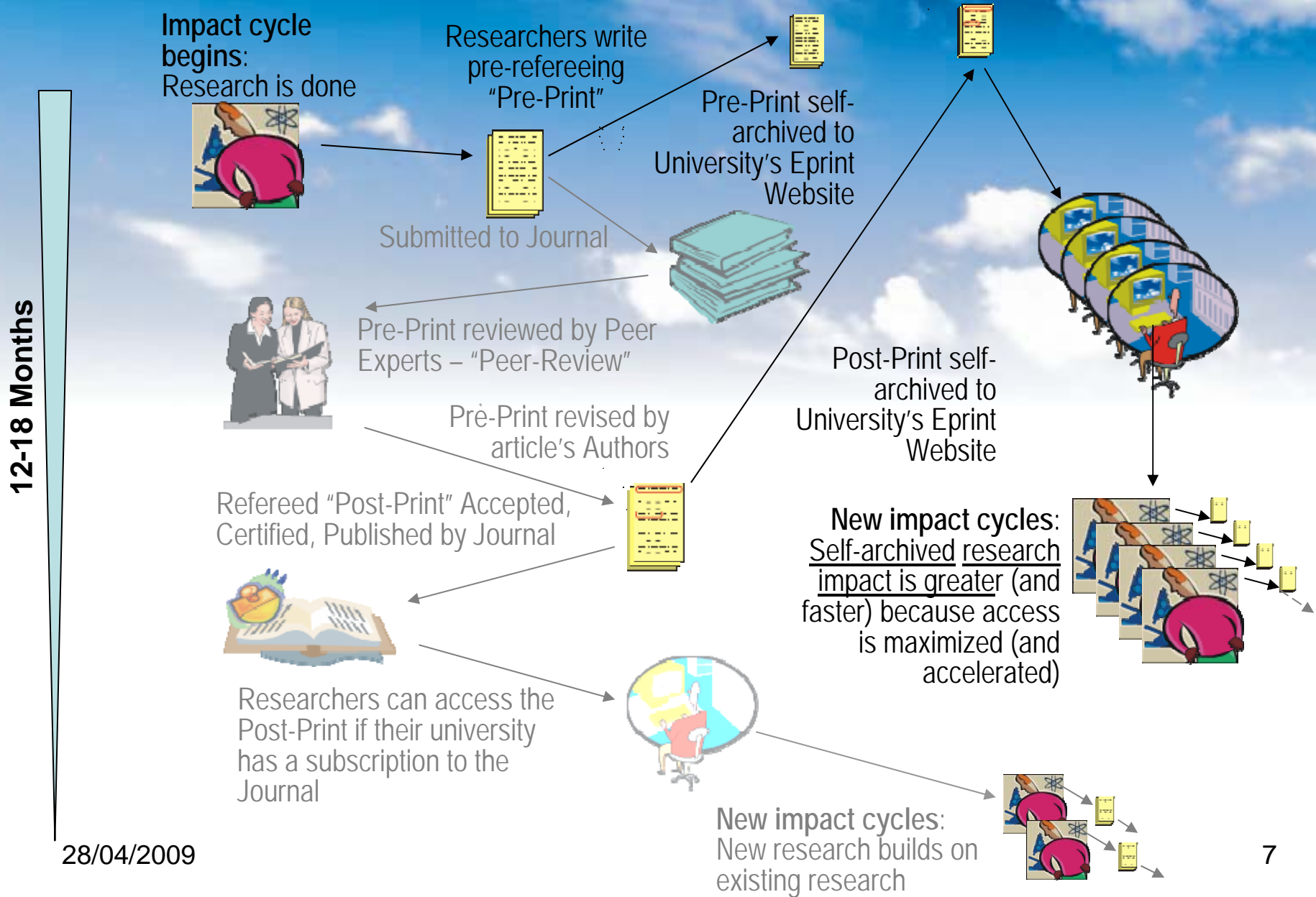
New impact cycles:  
New research builds on existing research



# Open Access Literature

- Research Archives (“author self-archiving”)
  - 340,000 arXiv.org
  - 700,000 citeseer
  - 1,000s in institutional & other repositories
- Open Access Journals
  - BioMed Central/PLOS/ [DOAJ]
- Time-delayed access
  - PubMed Central
  - HighWire Press
- Personal Web pages

# "Skywriting": All research, accessible to all potential users, anywhere, anytime



28/04/2009

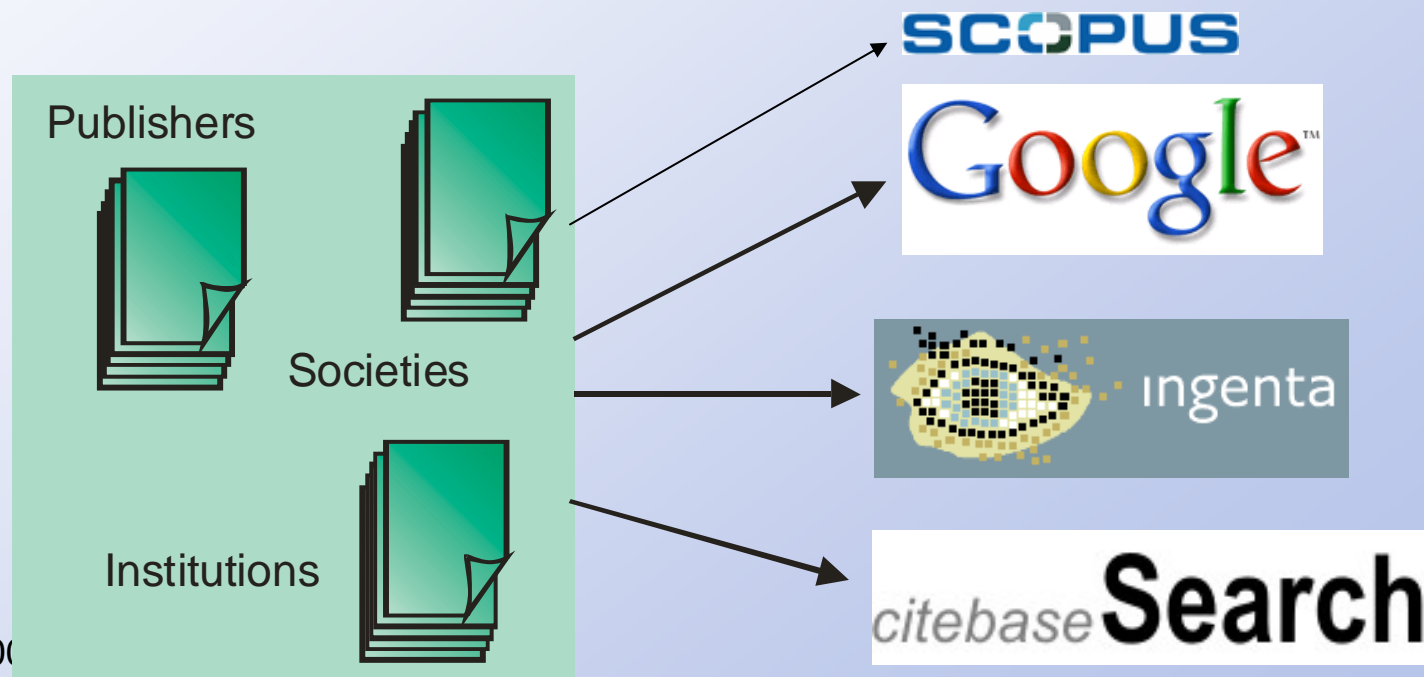
# Why Open Access?

- Maximise research impact through maximised access [evidence of 50-250% more citations]
- Efficiency
  - ADS Est. to provide \$250 million benefit to astronomy
- Continuous and comprehensive assessment
- Periphery benefits
  - Institutional management
  - Publicly funded research publicly accessible
  - Developing World Access
  - Easier to identify plagiarism (do a Google search!)



# OA Benefit: Services

- General Web-based full-text search engines
- Research literature specific: A&I, publisher, library, aggregator, journal contents, society ...



# Citebase Search

- “Citebase Search is a semi-autonomous citation index for the free, online research literature. It harvests pre- and post- prints (most author self-archived) from [OAI-PMH](#) compliant archives, parses and links their references and indexes the metadata in a search engine.”

citebase Search - "black hole" energy - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

# citebase Search

[Help](#)  
[\(Impact\) Health-Warning](#)

[Metadata](#) [Citation](#) [OAI Identifier](#)

Author(s) ([explain?](#))

Title/Abstract Keywords

Publication title

Creation Date from  until

Rank matches by:   [explain?](#)

Showing 1 - 10 of 2877 found [ 1-100 in [XML](#), [Refer](#), [BibTeX](#), [RSS](#) ] Query took 0.145 seconds

**[New Dimensions at a Millimeter to a Fermi and Superstrings at a TeV](#) [ [Abstract/Citations](#), [Cached PDF](#) ]**  
1868 [Antoniadis, I.](#); [Arkani-Hamed, N.](#); [Dimopoulos, S.](#) et al (1998-04-24) *In* [Physics Letters B](#) 436 257 (1998)  
... een proposed which does not rely on low energy supersymmetry or technicolor. The gravi ... trong gravity such as the production of black holes. The new dimensions can be probed by events with large missing energy carried off by gravitons escaping into ...

**[The Black Hole in Three Dimensional Space Time](#) [ [Abstract/Citations](#), [Cached PDF](#) ]**  
750 [Bañados, Máximo](#); [Teitelboim, Claudio](#); [Zanelli, Jorge](#) (1992-04-29) *In* [Physical Review Letters](#) 69 1849 (1992)  
... negative cosmological constant, admit a black hole solution. The 2+1 black hole -characterized by mass, angular momentu ... i-de Sitter space appears as a negative energy state separated by a mass gap from the continuous black hole spectrum. Evaluation of the partition f ...

**[The World as a Hologram](#) [ [Abstract/Citations](#), [Cached PDF](#) ]**  
744 [Susskind, L.](#) (1994-09-15) *In* [Journal of Mathematical Physics](#) 36 6377 (1995)  
... Planck scale. The consequences for high energy particle collisions are described. The ... s related to information spreading near black hole horizons. The considerations of this pa ...

**[String Solitons](#) [ [Abstract/Citations](#), [Cached PDF](#) ]**  
490 [Duff, M. J.](#); [Khuri, Ramzi R.](#); [Lu, J. X.](#) (1994-12-21) *In* [Physics Reports](#) 259 213 (1995)

Done

## New Dimensions at a Millimeter to a Fermi and Superstrings at a TeV

Authors: [Antoniadis, I.](#); [Arkani-Hamed, N.](#); [Dimopoulos, S.](#); [Dvali, G.](#)

Recently, a new framework for solving the hierarchy problem has been proposed which does not rely on low energy supersymmetry or technicolor. The gravitational and gauge interactions unite at the electroweak scale, and the observed weakness of gravity at long distances is due the existence of large new spatial dimensions. In this letter, we show that this framework can be embedded in string theory. These models have a perturbative description in the context of type I string theory. The gravitational sector consists of closed strings propagating in the higher-dimensional bulk, while ordinary matter consists of open strings living on D3-branes. This scenario raises the exciting possibility that the LHC and NLC will experimentally study both ordinary aspects of string physics such as the production of narrow Regge-excitations of all standard model particles, as well more exotic phenomena involving strong gravity such as the production of black holes. The new dimensions can be probed by events with large missing energy carried off by gravitons escaping into the bulk. We finally discuss some important issues of model building, such as proton stability, gauge coupling unification and supersymmetry breaking.

Comment: 12 pages, latex

**Full-text available from:** [Cached PDF](#)  
[Linked PDF \(experimental\)](#)  
Phys.Lett. B436 (1998) 257-263  
<http://arxiv.org/abs/hep-ph/9804398>

- [This Article's Citation/Hit History](#)
- [All Articles Cited by this Article \(Reference List\)](#)
- [Top 5 Articles Citing this Article](#)
- [All Articles Citing this Article](#)
- [Top 5 Most Co-cited Articles with this Article](#)
- [All Articles Co-Cited with this Article](#)

**Based on record (harvested at)** oai:arXiv.org:hep-ph/9804398 (2005-09-20)

## This Article's Citation/Hits History (explain?)

> [Top](#) [#Reference List](#) [#Cited By](#) [#Co-Cited](#) [All Citations](#) [All Co-Citations](#)

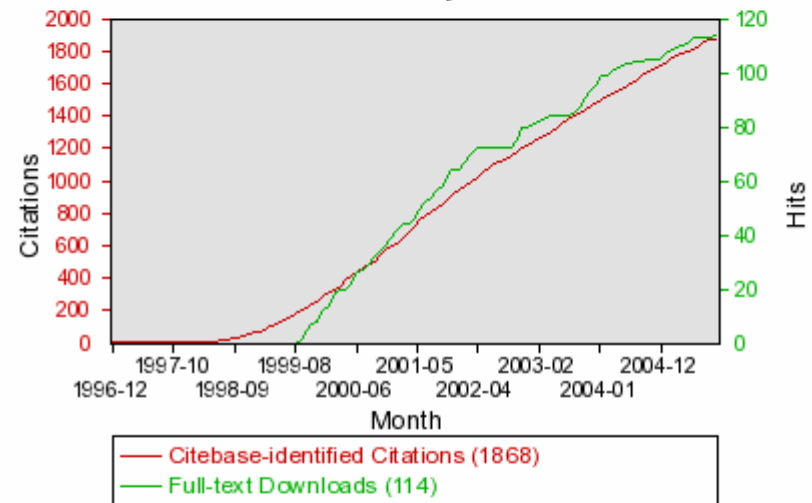
Use the [Correlation Generator](#) to explore the correlation between download impact ("hits") and citation impact.

### Summary

Citebase is currently only an experimental demonstration. Users are cautioned not to use it for academic evaluation yet. Citation coverage and analysis is [incomplete](#) and hit coverage and analysis is both [incomplete](#) and [noisy](#).

Caution!	Citations	Full-text Downloads
To this Article	1868	114
To authors (mean)	25.0000	25.0000

### History



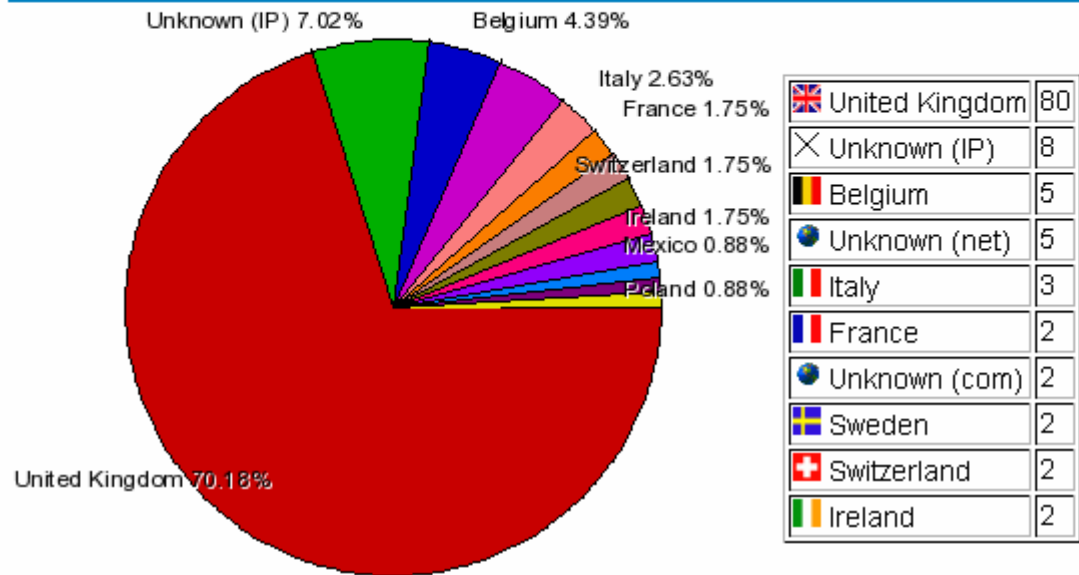
## This Record's Reference List (explain?)

> [Top](#) [#Reference List](#) [#Cited By](#) [#Co-Cited](#) [All Citations](#) [All Co-Citations](#)

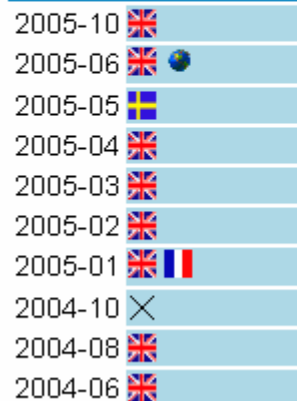
- [eprint](#) [1] N. Arkani-Hamed, S. Dimopoulos and G. Dvali, hep-ph/9803315, to appear in Phys. Lett.B.  
[G/A](#) [2] S. Dimopoulos and H. Georgi, Nucl. Phys. B193 (1981) 150.  
[eprint](#) [3] E. Witten, Nucl. Phys. B471 (1996) 135.  
[eprint](#) [4] I. Antoniadis and M. Quirós, Phys. Lett. B392 (1997) 61.  
[G/A](#) [5] I. Antoniadis, Phys. Lett. B246 (1990) 377  
[eprint](#) I. Antoniadis, C. Munoz and M. Quirós, Nucl. Phys. B397 (1993) 515.  
[G/A](#) [6] C. Bachas, 1995, private communication.  
[eprint](#) [7] J. D. Lykken, [Phys. Rev. D](#)54 (1996) 3693.



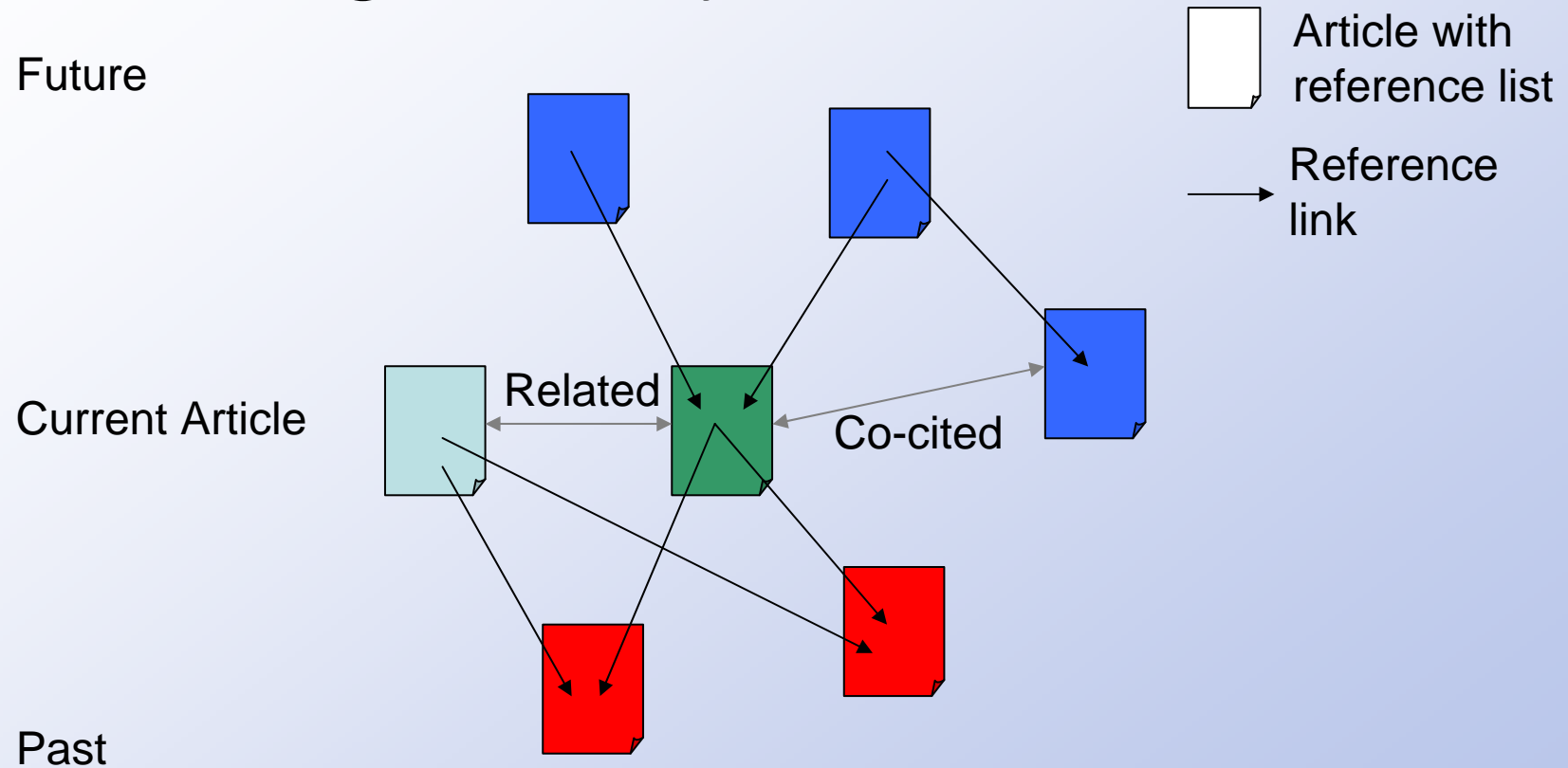
### Top 10 Countries



### All Downloads over Time



# Citebase Search: Navigation by Citation Links



## This Record's Reference List (explain?)

> Top #Reference List #Cited By #Co-Cited All Citations All Co-Citations

- [eprint](#) [1] N. Arkani-Hamed, S. Dimopoulos and G. Dvali, hep-ph/9803315, to appear in Phys
- [G/A](#) [2] S. Dimopoulos and H. Georgi, Nucl. Phys. B193 (1981) 150.
- [eprint](#) [3] E. Witten, Nucl. Phys. B471 (1996) 135.
- [eprint](#) [4] I. Antoniadis and M. Quirós, Phys. Lett. B392 (1997) 61.
- [G/A](#) [5] I. Antoniadis, Phys. Lett. B246 (1990) 377
- [eprint](#) [6] I. Antoniadis, C. Munoz and M. Quirós, Nucl. Phys. B397 (1993) 515.
- [G/A](#) [7] J. D. Lykken, [Phys. Rev. D](#)54 (1996) 3693.
- [eprint](#) [8] P. Horava and E. Witten, Nucl. Phys. B460 (1996) 506 and B475 (1996) 94.
- [eprint](#) [9] J. Polchinski and E. Witten, Nucl. Phys. B460 (1996) 525.
- [G/A](#) [10] N. Arkani-Hamed, S. Dimopoulos and G. Dvali, in preparation.
- [eprint](#) [11] M.R. Garousi and R.C. Myers, Nucl. Phys. B475 (1996) 193
- [eprint](#) [12] A. Hashimoto and I.R. Klebanov, Phys. Lett. B381 (1996) 437 .
- [eprint](#) [13] I. Antoniadis and K. Benakli, Phys. Lett. B326 (1994) 69.
- [eprint](#) [14] C.D. Carone and H. Murayama, [Phys. Rev. D](#)52 (1995) 484.
- [G/A](#) [15] L.E. Ibanez and G.G. Ross, Nucl. Phys. B368 (1992) 3.
- [eprint](#) [16] K.R. Dienes, E. Dudas, T. Gherghetta, hep-ph/9803466.
- [eprint](#) [17] S. Dimopoulos and G.F. Giudice, Phys. Lett. B379 (1996) 105
- [eprint](#) [18] I. Antoniadis, S. Dimopoulos and G. Dvali, Nucl. Phys. B516 (1998) 70.

## Top 5 Records Citing this Record (explain?)

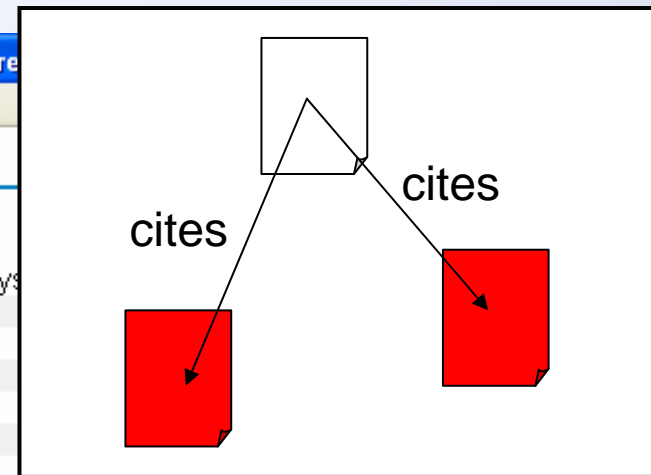
> Top #Reference List #Cited By #Co-Cited All Citations All Co-Citations

### [A Large Mass Hierarchy from a Small Extra Dimension](#) [ [Abstract/Citations](#), [Cached PDF](#) ]

2548 [Randall, Lisa](#); [Sundrum, Raman](#) (1999-05-03) In [Physical Review Letters](#) 83 3370 (1999)

We propose a new higher-dimensional mechanism for solving the Hierarchy Problem. The Weak scale is generated from a large scale of order the Planck scale through an exponential hierarchy. However, this exponential arises not from gauge interactions but from the background metric (which is a slice ... Comment: 9 pages, LaTeX

### [An Alternative to Compactification](#) [ [Abstract/Citations](#), [Cached PDF](#) ]



## Top 5 Most Co-cited Records with this Record (explain?)

> Top #Reference List #Cited By #Co-Cited All Citations All Co-Citations

### [The Hierarchy Problem and New Dimensions at a Millimeter](#) [ [Abstract/Citations](#), [Cached PDF](#) ]

1714 [Arkani-Hamed, Nima; Dimopoulos, Savas; Dvali, Gia](#) (1998-03-11) *In* [Physics Letters B](#) 429 263-275 (1999)

We propose a new framework for solving the hierarchy problem which does not rely on either supersymmetry or technicolor. In this framework, the gravitational and gauge interactions become united at the weak scale, which is the distance scale in nature. The observed w ... Comment: 16 pages, latex, no figures

### [A Large Mass Hierarchy from a Small Extra Dimension](#) [ [Abstract/Citations](#), [Cached PDF](#) ]

1055 [Randall, Lisa; Sundrum, Raman](#) (1999-05-03) *In* [Physical Review Letters](#) 83 3370 (1999)

We propose a new higher-dimensional mechanism for solving the Hierarchy Problem. The Weak scale is generated from a large scale of order the Planck scale through an exponential hierarchy. However, this exponential arises not from gauge interactions but from the background metric (which is a slice ... Comment: 9 pages, LaTeX

### [Phenomenology, Astrophysics and Cosmology of Theories with Sub-Millimeter Dimensions and TeV Scale Quantum Gravity](#)

[ [Abstract/Citations](#), [Cached PDF](#) ]

835 [Arkani-Hamed, Nima; Dimopoulos, Savas; Dvali, Gia](#) (1998-07-11) *In* [Physical Review D](#) 59 086004 (1999)

We recently proposed a solution to the hierarchy problem not relying on low-energy supersymmetry or technicolor. Instead, the problem is nullified by bringing quantum gravity down to the TeV scale. This is accomplished by the presence of  $n \geq 2$  new dimensions of sub-millimeter size, with the SM fie ... Comment: 51 pages, latex

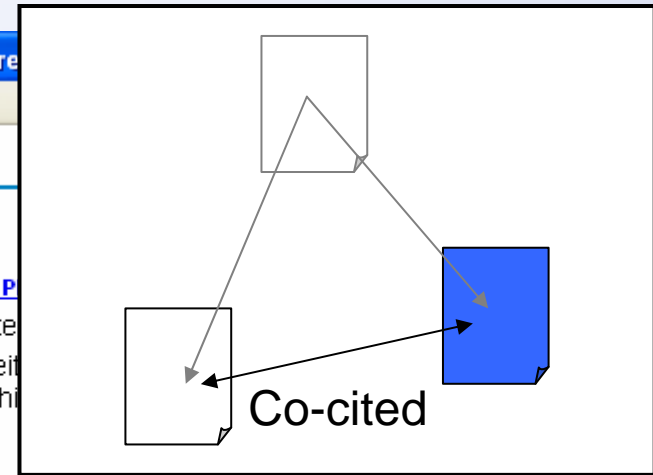
### [An Alternative to Compactification](#) [ [Abstract/Citations](#), [Cached PDF](#) ]

795 [Randall, Lisa; Sundrum, Raman](#) (1999-06-08) *In* [Physical Review Letters](#) 83 4690 (1999)

Conventional wisdom states that Newton's force law implies only four non-compact dimensions. We demonstrate that this is not necessarily true in the presence of a non-factorizable background geometry. The specific example we study is a single 3-brane embedded in five dimensions. We show that even ... Comment: LaTeX, 9 pages

### [Heterotic and Type I String Dynamics from Eleven Dimensions](#) [ [Abstract/Citations](#), [Cached PDF](#) ]

402 [Horava, Petr; Witten, Edward](#) (1995-10-29) *In* [Nuclear Physics B](#) 460 506 (1996)



citebase **Search** [Help](#)  
([Impact](#)) [Health-Warning](#)[Metadata](#) [Citation](#) [OAI Identifier](#)Identifier ([explain?](#)) Rank matches by   [explain?](#)     Showing 1 - 10 of 4253 found [ 1-100 in [XML](#), [Refer](#), [BibTeX](#), [RSS](#) ]

Query took 0.536 seconds

**[A Large Mass Hierarchy from a Small Extra Dimension](#)** [ [Abstract/Citations](#), [Cached PDF](#) ]2548 [Randall, Lisa](#); [Sundrum, Raman](#) (1999-05-03) *In* [Physical Review Letters](#) 83 3370 (1999)

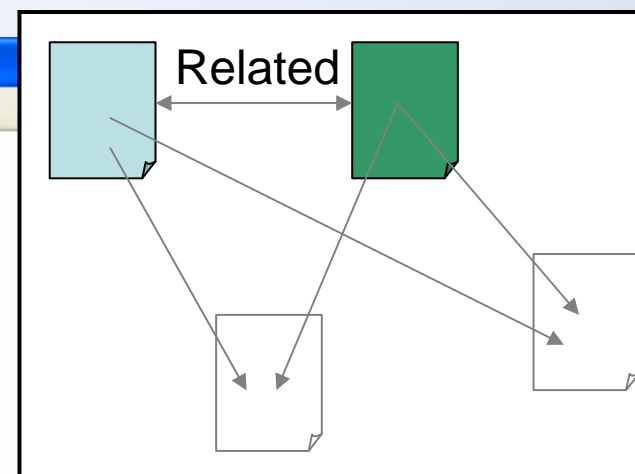
We propose a new higher-dimensional mechanism for solving the Hierarchy Problem. The Weak scale is generated from a large scale of order the Planck scale through an exponential hierarchy. However, this exponential arises not from gauge interactions but from the background metric (which is a slice ... Comment: 9 pages, LaTeX

**[An Alternative to Compactification](#)** [ [Abstract/Citations](#), [Cached PDF](#) ]2229 [Randall, Lisa](#); [Sundrum, Raman](#) (1999-06-08) *In* [Physical Review Letters](#) 83 4690 (1999)

Conventional wisdom states that Newton's force law implies only four non-compact dimensions. We demonstrate that this is not necessarily true in the presence of a non-factorizable background geometry. The specific example we study is a single 3-brane embedded in five dimensions. We show that even ... Comment: LaTeX, 9 pages

**[Heterotic and Type I String Dynamics from Eleven Dimensions](#)** [ [Abstract/Citations](#), [Cached PDF](#) ]1438 [Horava, Petr](#); [Witten, Edward](#) (1995-10-29) *In* [Nuclear Physics B](#) 460 506 (1996)

We propose that the ten-dimensional  $E_8 \times E_8$  heterotic string is related to an eleven-dimensional theory on the orbifold  $\{bf R\}^{10} \times \{bf S\}^1 / \{bf Z\}_2$  in the same way that the Type IIA string in ten dimensions is related to  $\{bf R\}^{10} \times \{bf S\}^1$ . This in particular determines the strong coupling behavi ... Comment: 27 pages, 2 postscript figures, harvmac, epsf, a reference added

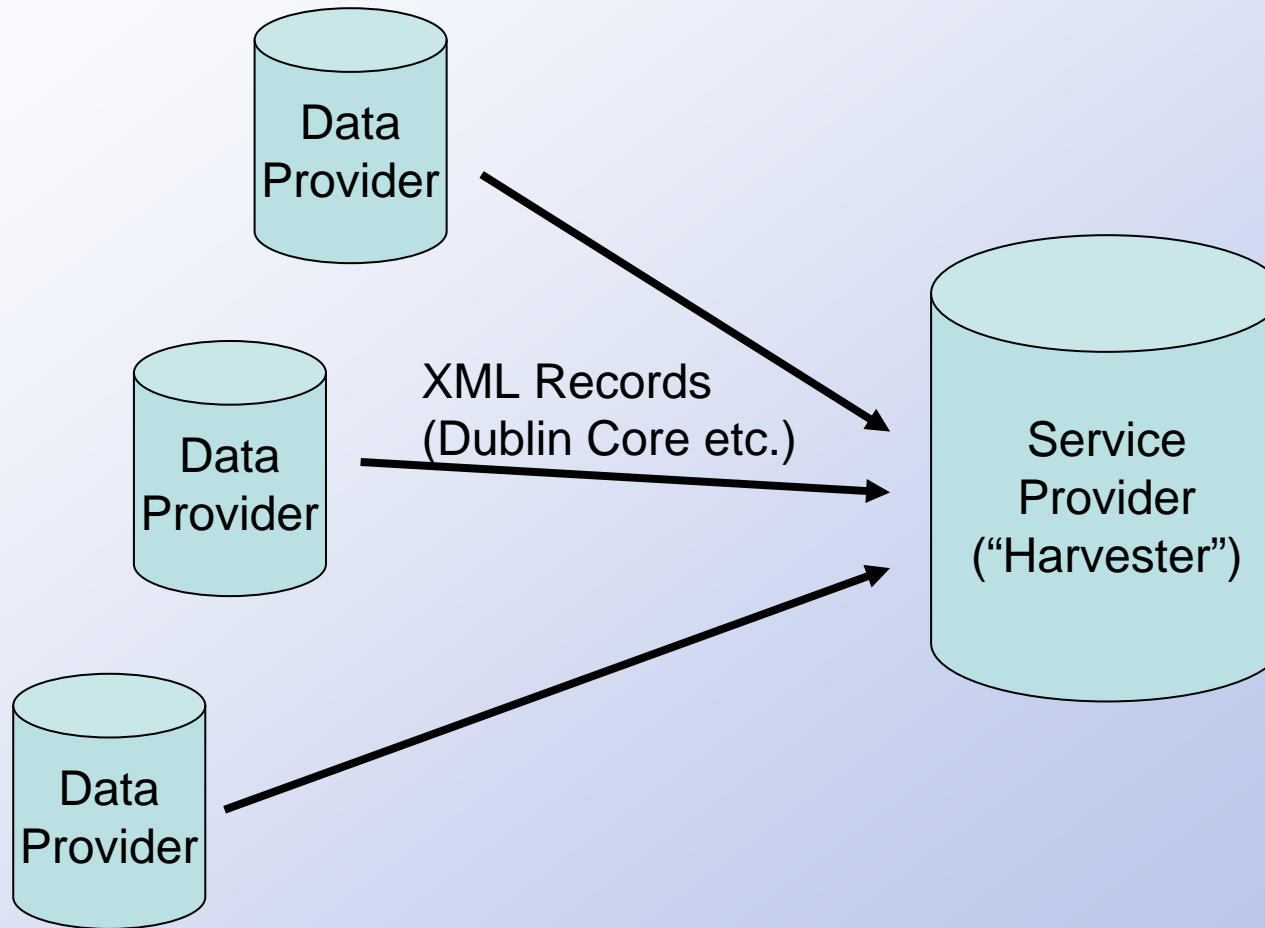
**[Large N Field Theories, String Theory and Gravity](#)** [ [Abstract/Citations](#), [Cached PDF](#) ]1284 [Aharony, O](#); [Gubser, S. S](#); [Maldacena, J. et al](#) (1999-05-14) *In* [Physics Reports](#) 323 183 (2000)



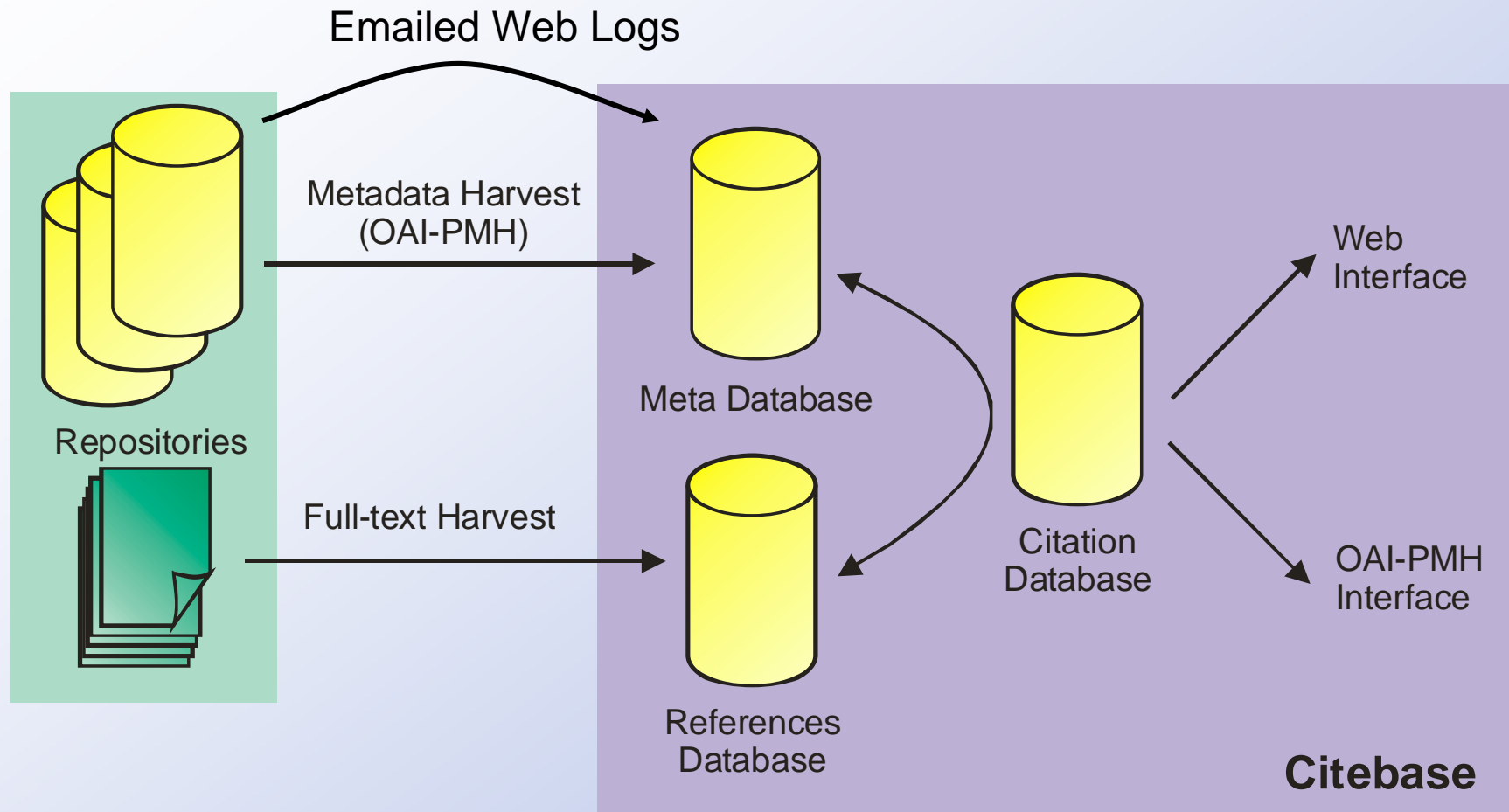
# Citebase Search: Usage

- 8000 users per-day (“visits”)
  - 60,000 hits
- 350,000 full-text records
- 12 million references
  - Of which 2.7 million linked to full-text
- 5 million Web download hits ([uk.arXiv.org](http://uk.arXiv.org))

# OAI-based Infrastructure



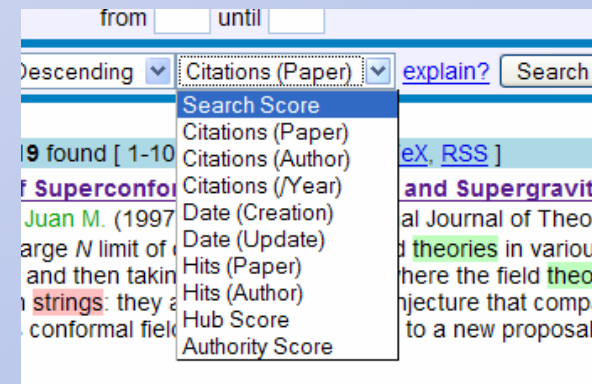
# Citebase Search



# Future Citation Ranking

- Citation- and Web- usage by:
  - Article
  - Authors
  - Journal
- Fan-in/Fan-out type rank
  - Hub/Authority (Google PageRank)
- Co-cites
  - Latent semantic analysis
- [All by time-trends]

28/04/2009



# Citation Indices for OA

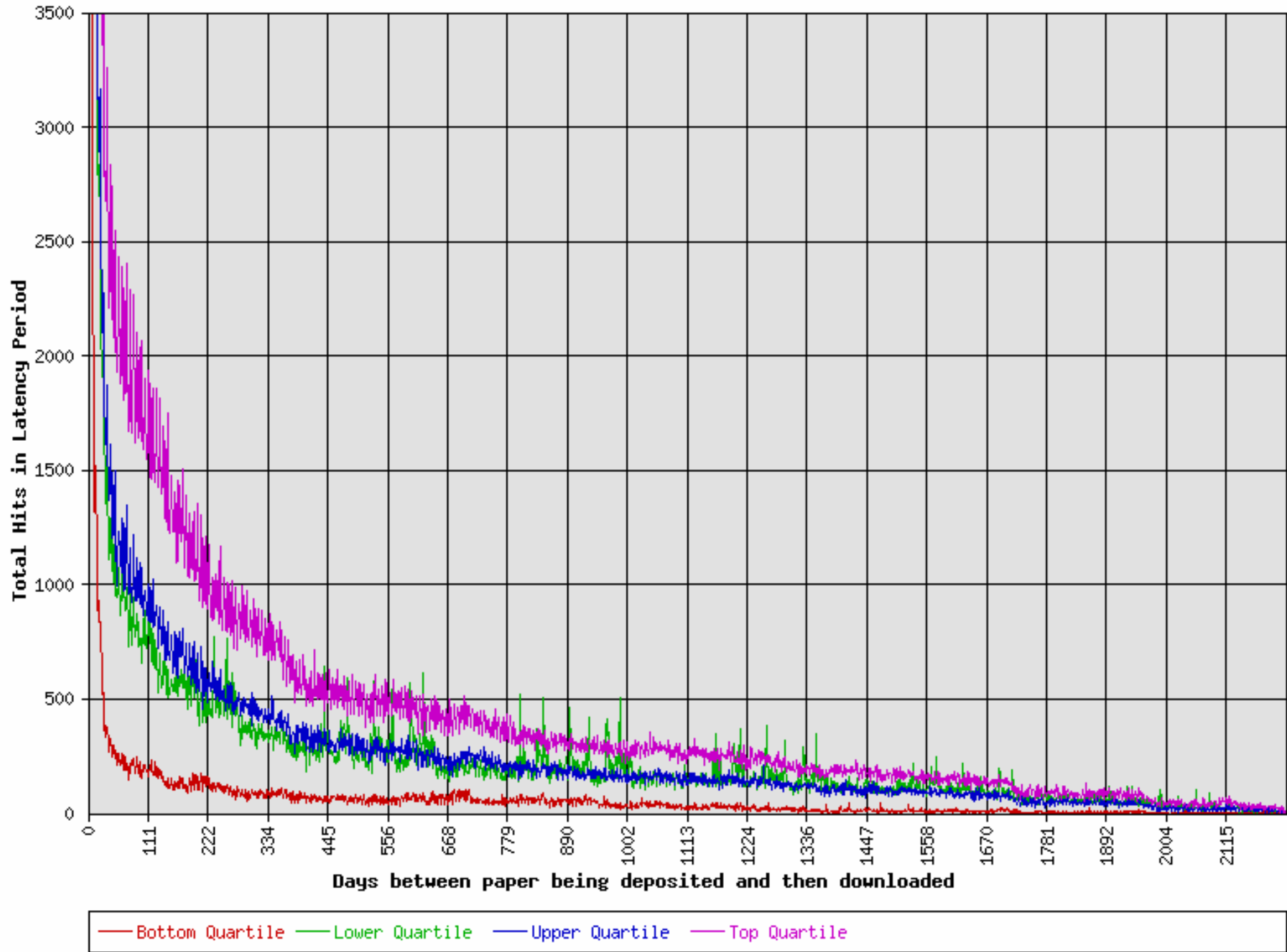
- Free to access:
  - Google Scholar
  - Citeseer
  - Astrophysics Data Service (NASA ADS)
- Subscription:
  - Scopus
  - ISI WCI



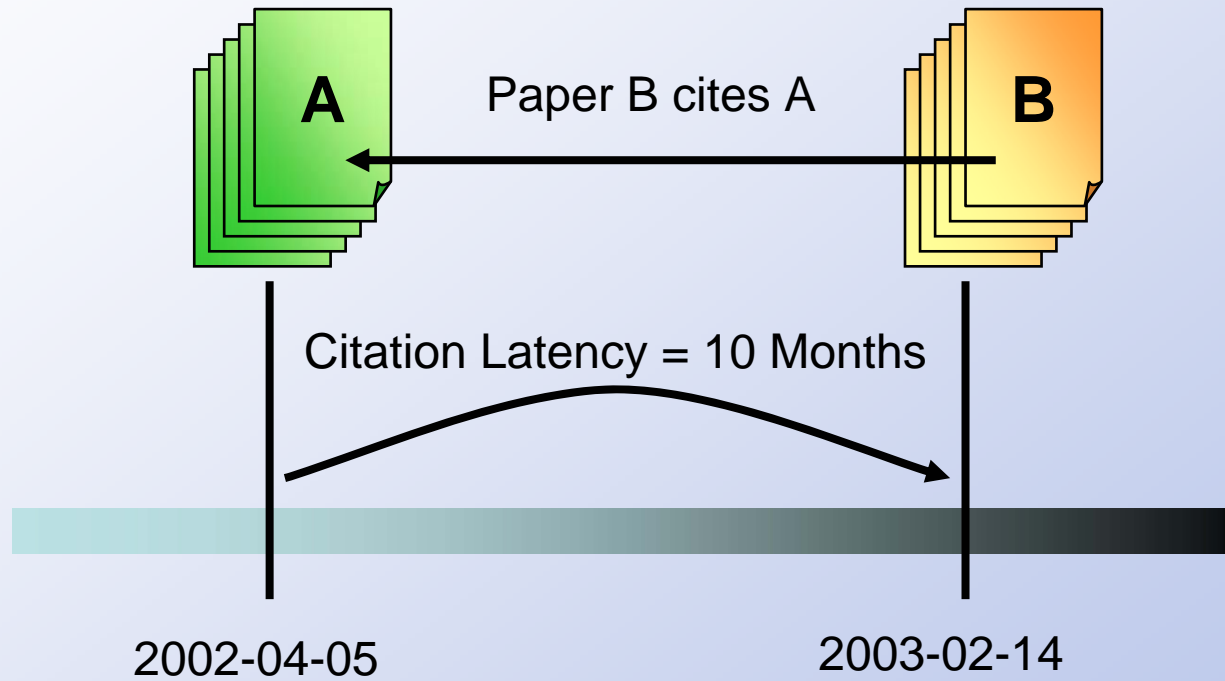
# Analysis: Effect of Open Access

1. Correlation of Citation Impact with Web Impact (downloads)
  - (Based on arXiv.org)
2. Effect of Open Access on citation behaviour
3. OA Impact Advantage

Latency of Downloads to Articles by Citation Impact Quartile



# Citation Latency





# OA Advantage

- (University of Quebec)
- ISI on CD-ROM, records linked to arXiv.org equivalents
  - In arXiv.org = OA
  - Not in arXiv.org = Not OA

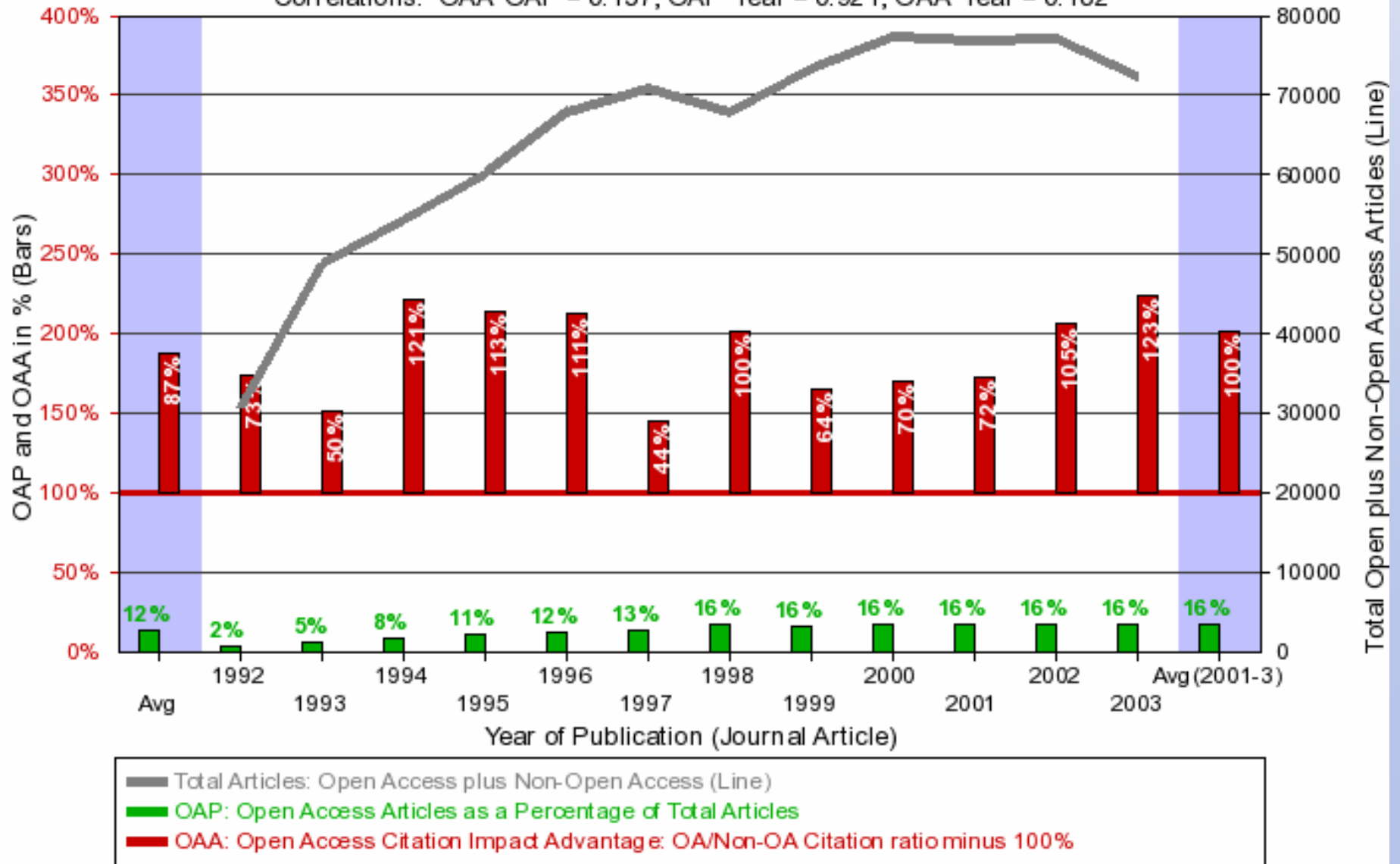


# Same-Journal Control, No Sample-Size Equalizing, incl. Self-Citations

OA vs. Non-OA Citation Impact Advantage (Physics)

On average 134.0 journals/year (incl. self-citations) - Fri Oct 7 11:30:32 2005

Correlations:  $OAA * OAP = 0.157$ ,  $OAP * Year = 0.921$ ,  $OAA * Year = 0.182$



# OA Advantage Analysis

- Open Access (OA)
  - + Competitive advantage vs. subs-only
- arXiv Advantage (AA)
  - + Primary resource for physicists
- Quality Bias (QB)
  - + Author self-selection (publish anything, self-archive the best ...)
- Early Advantage
  - + Cites to pre-print + (later) cites to post-print

# Conclusions

- High impact papers are read more (and this can be measured online)
  - Web downloads may be an pre-indicator of impact
- Faster access leads to reducing Citation Latencies
  - Hence faster research cycles, higher impact, and more productivity

# Summary

- The Web makes Open Access research literature possible, and hence more effective scholarship
- Services compete without holding the literature hostage
- OAI allows repositories to concentrate on getting and storing the literature
- Citebase Search provides citation navigation for OAI archive(s)
  - Or anyone else who wants to provide a similar service

# The Last Slide

- Tim Brody [tdb01r@ecs.soton.ac.uk](mailto:tdb01r@ecs.soton.ac.uk)
  - <http://opcit.eprints.org/> (papers & presentations)
- Citebase Search
  - <http://www.citebase.org/>
- EPrints.org
  - (advocacy, answers & software)
  - <http://www.eprints.org/>

I am a doctoral student in the Intelligence, Agents, Multimedia Group at the University of Southampton working with digital library systems: Citebase Search, E-Prints UK, TARDIS & OAI.

Prof. Stevan Harnad <[harnad@ecs.soton.ac.uk](mailto:harnad@ecs.soton.ac.uk)>

# The Research Literature

- The grey literature
  - Technical reports
  - Monographs
  - Presentations
- Royalty literature
  - Books
- Refereed journal corpus

# The Refereed Journal Literature

- Written without the expectation of royalties
- Akin to 'Advertising' for authors and their work (=maximise use & uptake)
- Reviewed for free by peers
- Est. 20,000 Peer-reviewed Journals
  - B.L. archives 60,000 serials
- Est. 2,000,000 Articles Annually

# Open Access: A Definition

- Immediate, free access via the Web to the author give-away **peer-reviewed** literature
  1. Author posted pre-prints, post-prints (“e-prints”) linked to journal version
  2. Open Access publishing