

Metrics for Openness

David Nichols

Department of Computer Science
University of Waikato
New Zealand



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

Outline

- Citation-based metrics
- altmetrics
- Open Access, Open Science
- Metrics to characterise Openness

Citation-based Metrics

- Journal Impact Factors,
- Total citations, avg. Citations
 - half-life, immediacy, EigenFactor
- h-index since 2005
 - Variants: g-index etc.
- Frequently used to characterise
 - Individuals, institutions, publication venues etc.

Keywords (26676)

African American Amino
Acid **Breast Cancer** Cell
Line Clinical Trial Confidence
Interval **Enzyme** Gene
Expression **Genetics**
Indexation Kinetics **North
Carolina Risk Factors**
United States Wild Type



Academic > Organizations > University of North Carolina Chapel Hill

Comparison

University of North Carolina Chapel Hill

[North America]

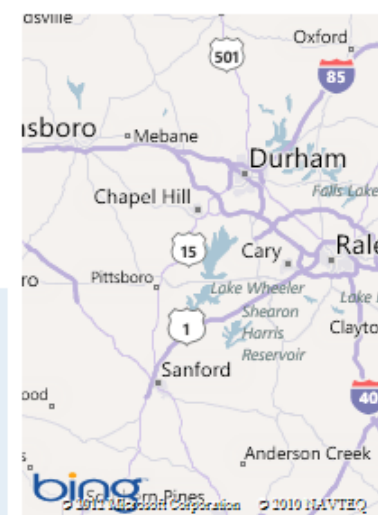
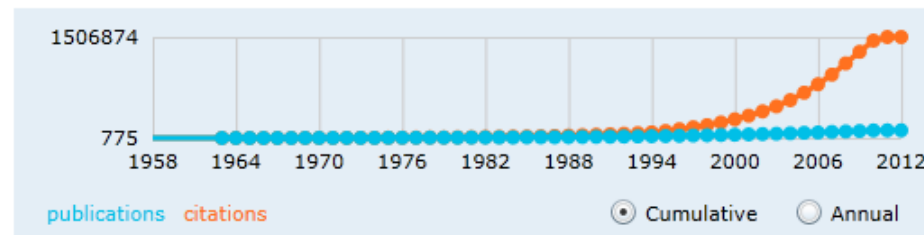
Publications: 126,189 | Citation Count: 1,734,180 | H-Index: 325

Top Areas: **Psychiatry &**

Psychology, Pharmacology, Neuroscience, Diseases, Oncology

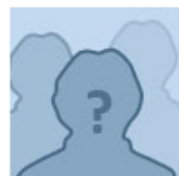
Sub-organization (1): Carolina Population Center

[Homepage](#)



Authors (9809)

All Years



Sidney C. Smith

Publications: 283 | Citations: 13360 | G-Index: 113 | H-Index: 58

Interests: Cardiology, Pharmacology, Plant & Animal Science



Edward D. Salmon

Publications: 220 | Citations: 8101 | G-Index: 81 | H-Index: 50

Interests: Cell Biology, Molecular Biology, Biophysics



Dinesh Manocha

Publications: 407 | Citations: 9721 | G-Index: 85 | H-Index: 47

Interests: Graphics, Artificial Intelligence, Algorithms & Theory

Co-authors (216)

Gary Geisler

Barbara M. Wildemuth

Catherine Plaisant

Ben Shneiderman

Anita H. Komlodi



Conferences (17)

JCDL

Academic > Authors > Gary Marchionini

Embed Subscribe



Gary Marchionini University of North Carolina Chapel Hill

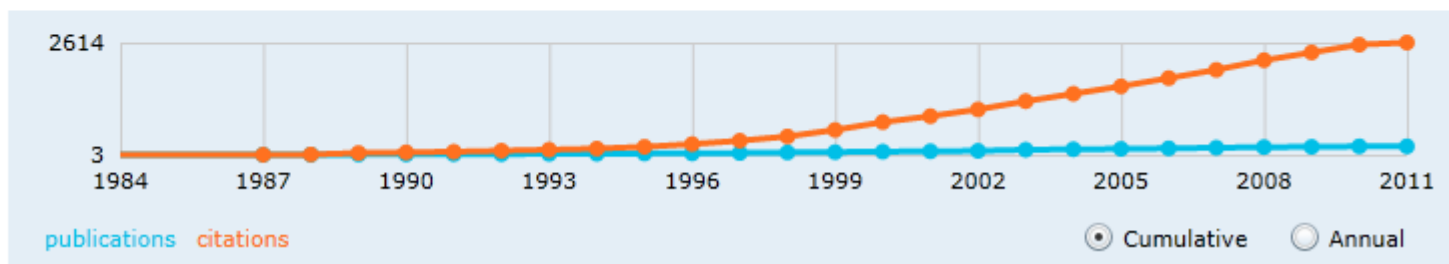
Publications: 228 | Citations: 3414 | G-Index: 54 | H-Index: 26

Interests: Information Retrieval, Human-Computer Interaction, Databases

Collaborated with 216 co-authors from 1987 to 2010; Cited by 3862 authors

Homepage | Bing

Edit



Jeffrey Pomerantz

Associate Professor, University of North Carolina at Chapel Hill

Digital Libraries - Digital Reference - Assessment - Evaluation - Pedagogy

Verified email at unc.edu

[Homepage](#)

Citation indices

	All	Since 2007
Citations	618	455
h-index	17	13
i10-index	19	16

Citations to my articles



<http://academic.research.microsoft.com/Author/1826681/gary-marchionini>

<http://scholar.google.co.nz/citations?user=hBKGymIAAAAJ&hl=en>

University of North Carolina Chapel Hill

Compare

Duke University

University of North Carolina Chapel Hill

North America

Publications: 8,628 | Citation Count: 152,932 | H-Index: 149

Top Areas: Psychiatry & Psychology, Pharmacology, Neuroscience, Diseases, Oncology

63

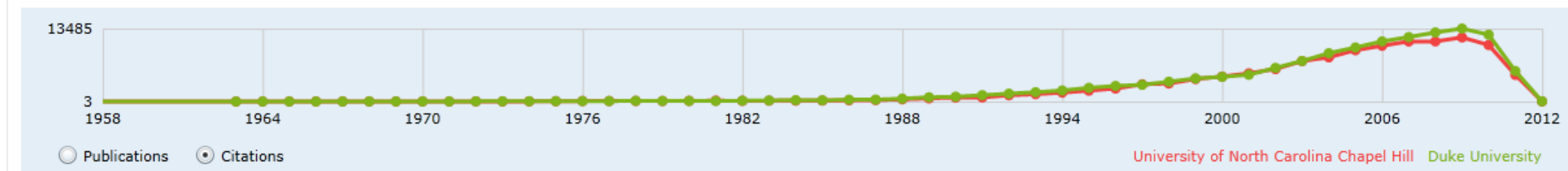
Duke University

North America

Publications: 8,812 | Citation Count: 164,954 | H-Index: 161

Top Areas: Neuroscience, Psychiatry & Psychology, Pharmacology, Cardiology, Oncology

55



Academic Ranking of World Universities in Computer Science - 2011

World Rank	Institution	Country	Total Score	Score on HiCi
35	University of North Carolina at Chapel Hill		47.9	0.0
36	Duke University		47.4	35.4

Rank	Title	Country	Academic	Employer	Citations	Score
51-100	Duke University	United States	24.9	43.3	39.7	
51-100	University of North Carolina, Chapel Hill	United States	17.8	24.0	57.3	

<http://academic.research.microsoft.com/Comparison?entitytype=7&id1=719&id2=814&topdomainid=2&subdomainid=0>

<http://www.shanghairanking.com/SubjectCS2011.html>

<http://www.topuniversities.com/university-rankings/world-university-rankings/2011/subject-rankings/engineering/computer-science>

- Citations are not enough
- Blogs, twitter, Mendeley, CiteULike etc.
- Article-level metrics

10.1371/journal.pbio.0050082

Eisen (2007) Environmental Shotgun Sequencing:
Its Potential and Challenges for Studying the
Hidden World of Microbes *Plos Biol.*





Gudmundur Thorisson

@gthorisson

United Kingdom

Academic researcher based in Leicester, UK and Reykjavik, Iceland. Interested in applications of Internet identity and Linked Data in research data sharing.

<http://gthorisson.name>

[JSON](#) | [XML](#) | [BibTeX](#) | [RIS](#) | [CSV](#)

[Identifiers](#)

[Metrics](#)

[Journal Articles \(14\)](#)

Altmetrics How often works by this author have been read, bookmarked or blogged by others

citeulike

118

MENDELEY

155

altmetric

10

Citations How often works by this author have been cited by others

PubMed Central

997

Microsoft Academic Search

1468

Scopus

4380

crossref

2659

an **altmetrics** project.

- We start to confuse the “kind of use we can track” with use and “citation impact” with impact
 - Priem (2012)
- Also, we confuse the “act of publishing” with effective *dissemination*
- We measure the *publication count*, *venue quality* and *citations* but not *other attributes* of publications



Open Access

“By ‘open access’ to this literature, we mean *its free availability on the public internet*, permitting any users to *read*, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself.”

<http://www.soros.org/openaccess/read>

- Research Works Act, Elsevier boycott



Fake Elsevier @FakeElsevier

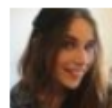
3h

We've increased the freq. with which we say the words "Open Access". This is what you wanted, right? Everything can go back to normal now.

Results for #icanhazpdf

Tweets · All ▾

[Refine results](#)



physilology Michelle

[#ICanHazPDF?](#) tandfonline.com/doi/abs/10.120...

8 hours ago



LizNeeley LizNeeley

[#icanhazpdf](#) please? Kasperson 1986, "Six Propositions on Public Participation and Their Relevance for Risk Communication" Risk Analysis

9 hours ago



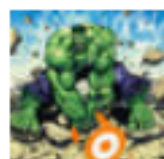
rmounce Ross Mounce

[#icanhazpdf](#) dx.doi.org/10.1063/1.3399... AIP paper. no access :(to ross.mounce @ gmail.com pls, thx!

kzelnio Kevin Zelnio

"I never thought much about it until I couldn't get a pdf copy of my own publication." Voices of [#OpenAccess](#)

1 Jan



OpenAccessHulk OpenAccessHulk

OA HULK SMASH THOUGHTLESS IR MGRS! ALSO OPAQUE IR SOFTWARE! melissaterras.blogspot.com/2011/11/on-thu...

2 Nov

Home	Research	Education	News	Comme
------	----------	-----------	------	-------

Search all BMJ news articles  From 1840  Jan 

Rapid responses are electronic letters to the editor and new ones are published each day

NEWS

Scientists are urged to oppose new US legislation that will put studies behind a pay wall

BMJ 2012; 344 doi: 10.1136/bmj.e452 (Published 17 January 2012)
Cite this as: BMJ 2012;344:e452

 Health informatics  Internet  Sociology

Article	Related content	Article metrics
---------	-----------------	-----------------

Access to the full text of this article requires a subscription or payment.
Please log in or subscribe below.

<http://www.flickr.com/photos/dullhunk/6875231621/in/pool-1928892@N23/>

<http://www.citeulike.org/group/13803>

<http://www.flickr.com/groups/open-access-irony-award/>

Missing Metrics

- We don't have metrics that characterise authors' behaviour with respect to Open Access issues
- So we can't easily compare, chide or celebrate authors' OAness



Villavelius Jan Velterop

Unfortunately it's 'publish or perish' and not 'share or be shunned' in the scientific ego-system. #SOPA #RWA

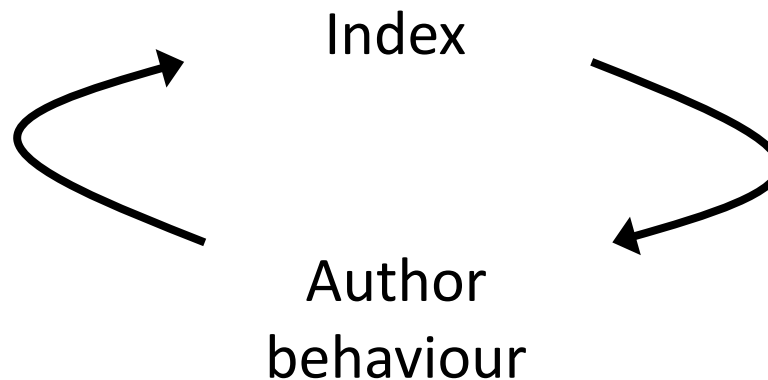
- Let's fix that...

- “The University of North Carolina is dedicated to the *transmission* and advancement of knowledge and understanding”
 - Chapter 6, Section 600, Clause I, *The Code Of The Board of Governors, The University of North Carolina*
 - <http://www.northcarolina.edu/policy/index.php?tag=100.1>
- OA is part of “transmission”
- Not being measured because of a lack of a culture, and metrics, of measuring Openness

“What makes it complicated for students is that they don't understand why we would write articles, give them away, and be okay with it showing up on the Internet with a \$35 price tag in the first place. Why make it so hard to get scholarly articles when sharing your research findings *is the whole point?*”

Openness Index

- Of the items *you* have published how many are free for anyone to read?
 - Practically: if someone searches for the title of your paper in Google do they find a copy they can read?



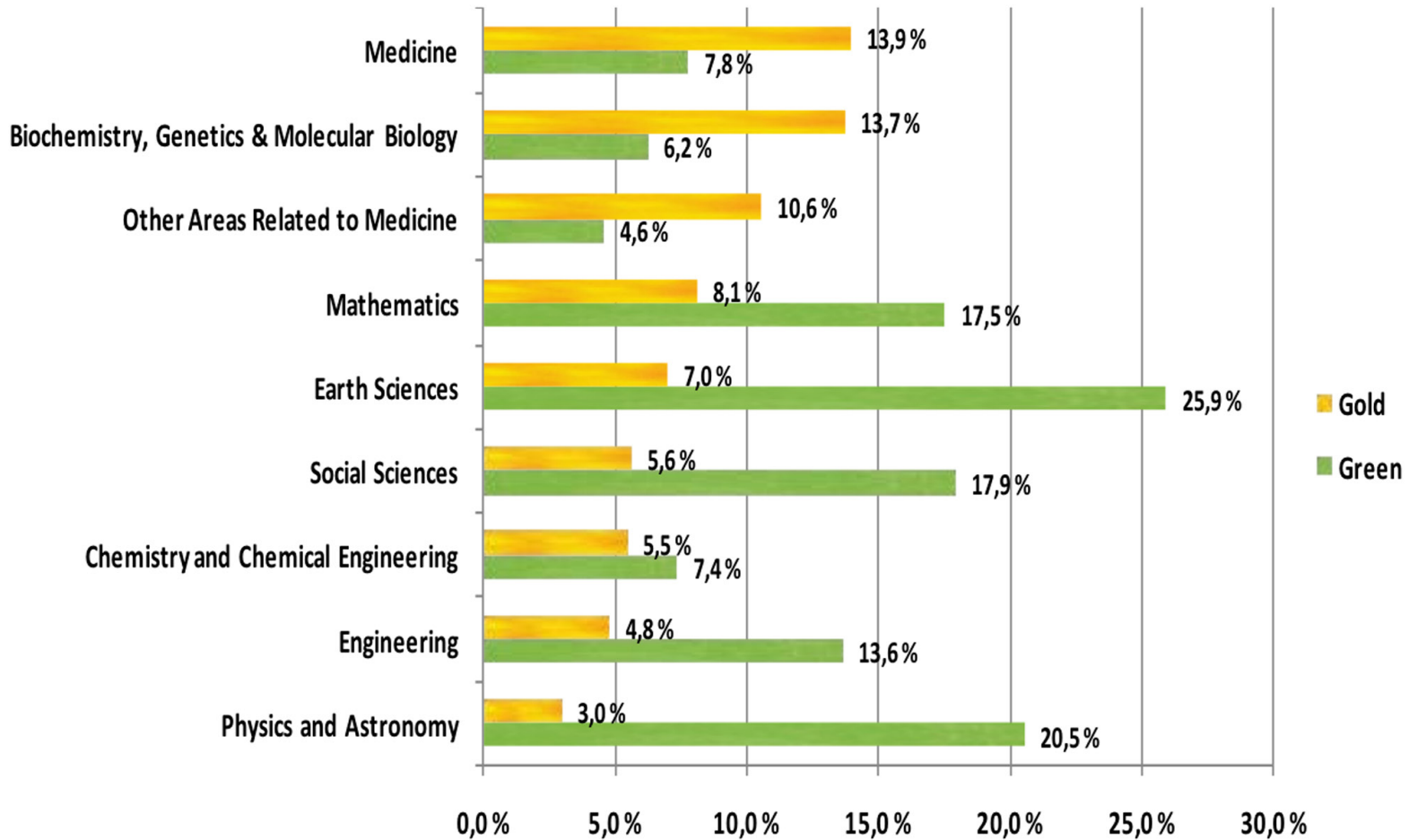
How to measure Openness?

- Grain size:
 - Person, Group, Dept, School, Institution, System
 - OA mandates (e.g. Harvard) will be evaluated at the institution/sub-institution level
 - Journal/Conference/etc, Publisher
 - Sub-discipline, Discipline
 - Funder
 - NSF Data Management Plans
 - State/Province, Country

Everything

“Björk et al. (2009) studied the annual volume and open access availability of scientific journal publishing. They estimated that after one year **11.3%** of the scientific output in 2006 could be found in specific or institutional repositories or on the home pages of the authors.”

Cross-Discipline



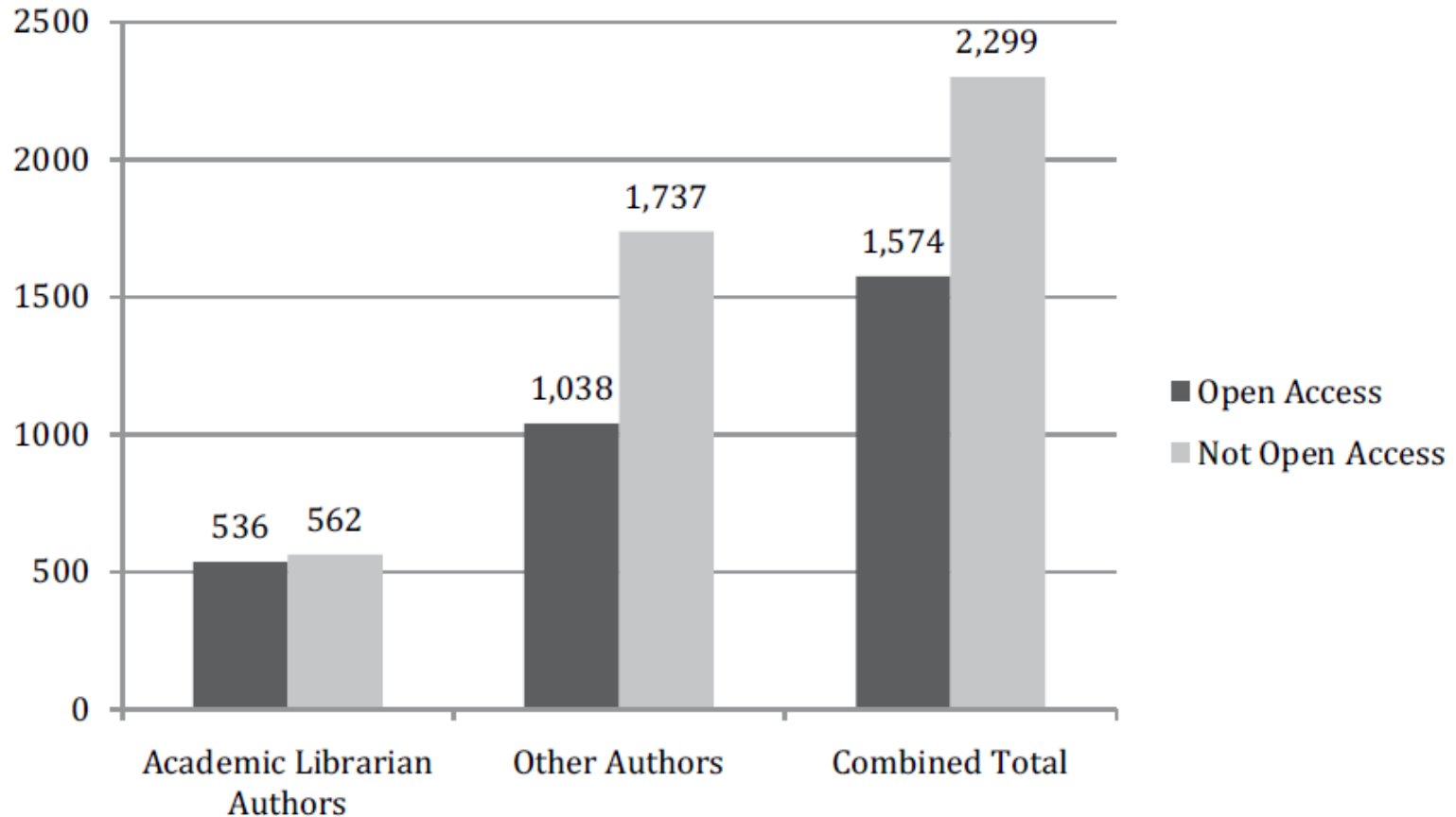
Cross-Discipline

TA = Toll Access

Subject	Total articles	Percentage OA	Percentage TA
Ecology	628	34.39	65.61
Economics	966	54.45	45.44
Sociology	925	24.32	75.68

Single Discipline - LIS

Open Access Comparison



“academic librarians ... OA choices compare favorably to the other authors in LIS journals”

Mercer (C&RL 2011)

Journals

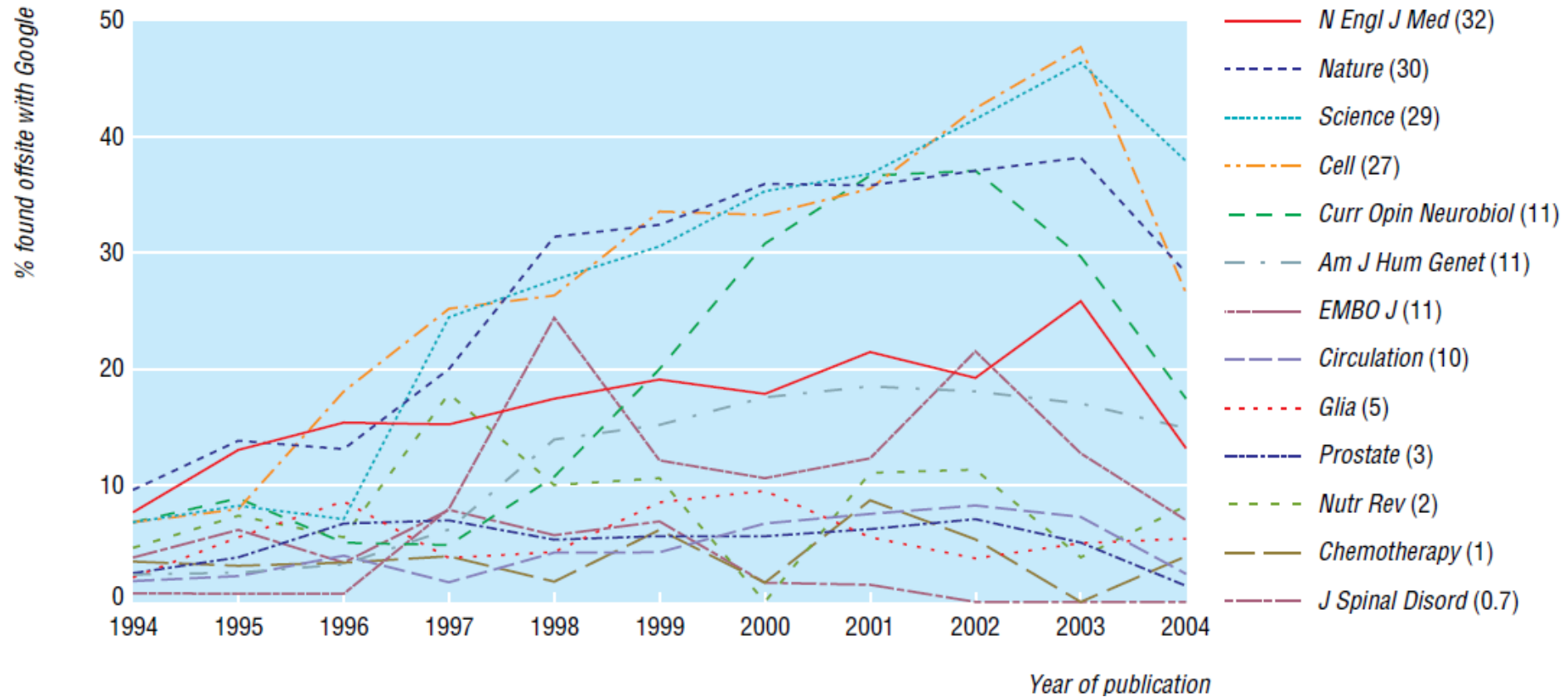


Fig 1 Subscription based journal articles locatable with Google at non-journal websites, with approximate impact factors for 2002 in parentheses. No articles were found in Medline for *J Spinal Disord* from 2002-3

“Locatable with Google” – not necessarily legal copies

Wren (BMJ 2005)

Institution

- University of Helsinki
- 7771 journal articles from 2007-8
- 5% sample

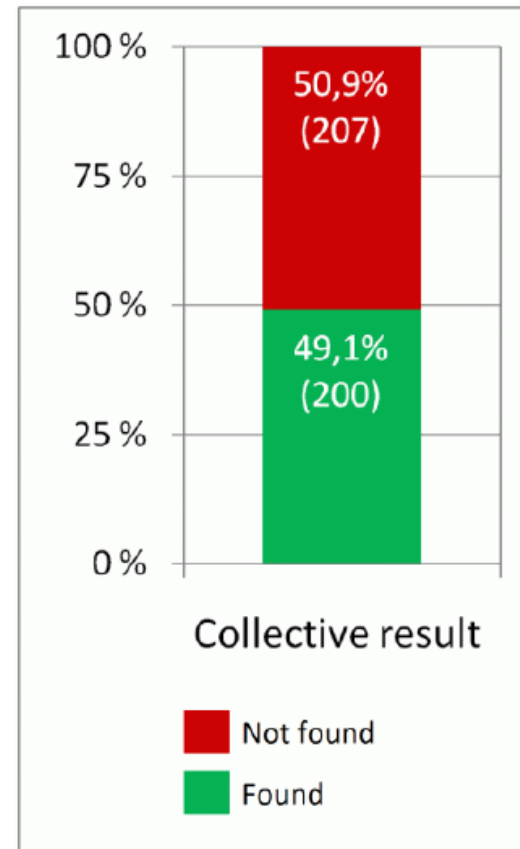


Figure 1: Collective Open Access availability.

Sub-Institution

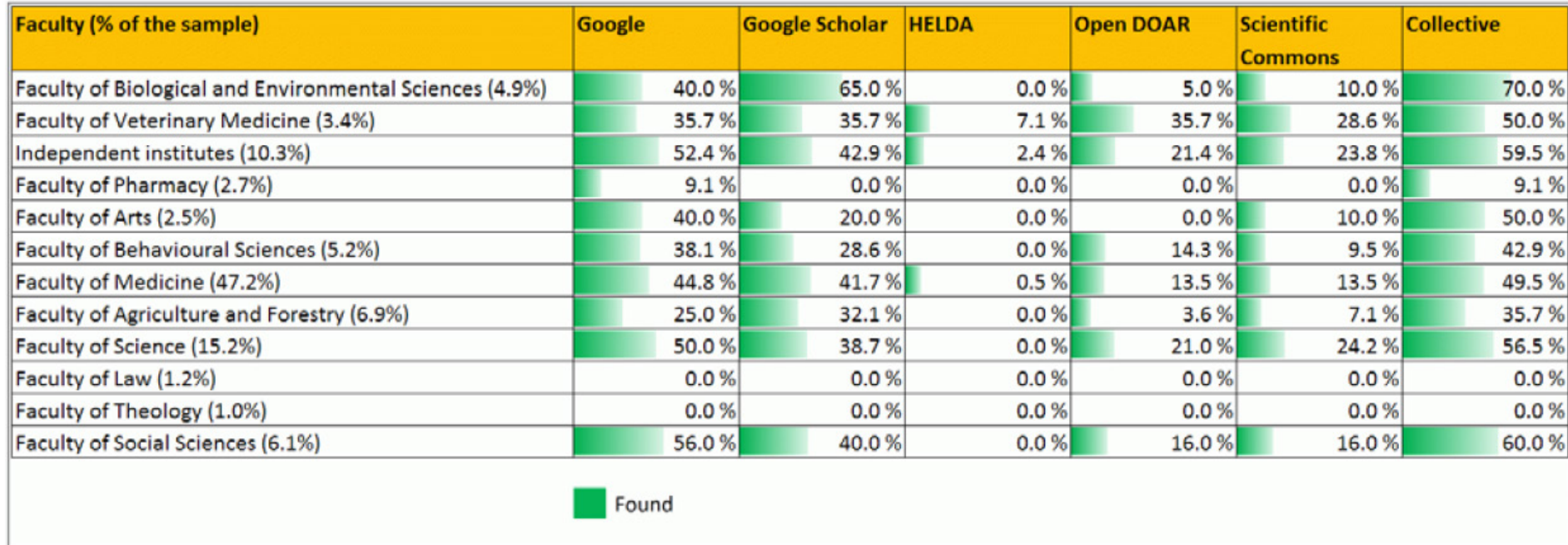


Figure 3: Open Access availability by faculty.

*The sum is more than 100% because of the co-authored articles across faculties. (Figure 3)

Faculties of the University of Helsinki

Repository

Repository Statistics

QUT ePrints

Search

- ☐ Advanced
- ☐ QUT Theses
- ☐ Affiliation

Browse

- ☐ Person
- ☐ Subjects
- ☐ Year

About

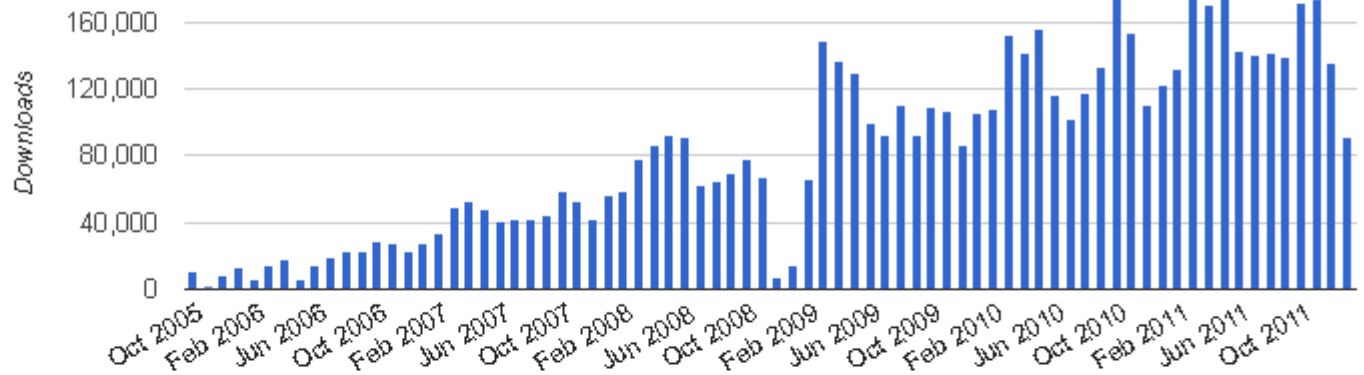
- ☐ Contact us
- ☐ Copyright matters
- ☐ Deposit Guide

Statistics

- ☐ Repository Statistics

QUT Users

- ☐ Deposit items



Key Figures

Collection

35,265 Works

55.2% Full-text

48.7% Open-access

Total Usage

6,273,655 Downloads

97.5% External

Yesterday

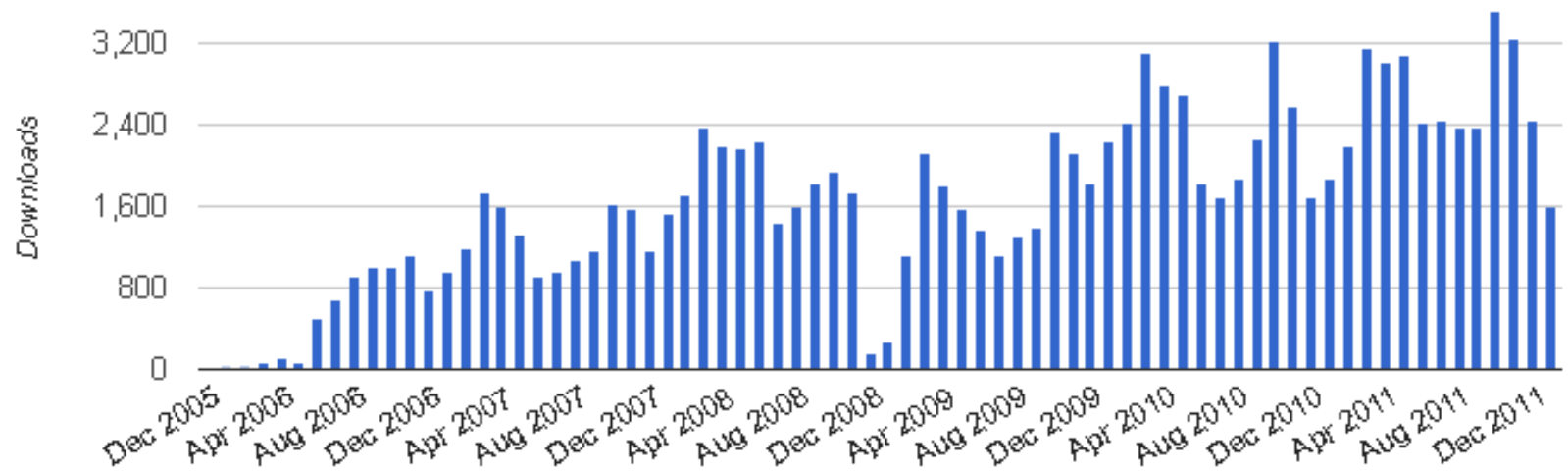
3,691 Downloads

4 Deposits

Repository - Individual

Statistics Overview

Skitmore, Martin



Key Figures

Collection

214 Works

98.1% Full-text ?

94.9% Open-access ?

Total Usage

123,148 Downloads ?

98.9% External ?

Last 28 Days

2,134 Downloads

<http://eprints.qut.edu.au/statistics/creators/Skitmore,%20Martin/>

Individual Level

- Missing?
- Just like the h-index for citations
- Each author gets an Openness Index defined across their publications

Openness Index

- = Openly Accessible Publications / Total Publications
- Example:
 - Time period: 2011
 - Publication type:
 - conference papers (C), journal papers (J)

$$OI_{2011}^{C,J} = 0.1$$

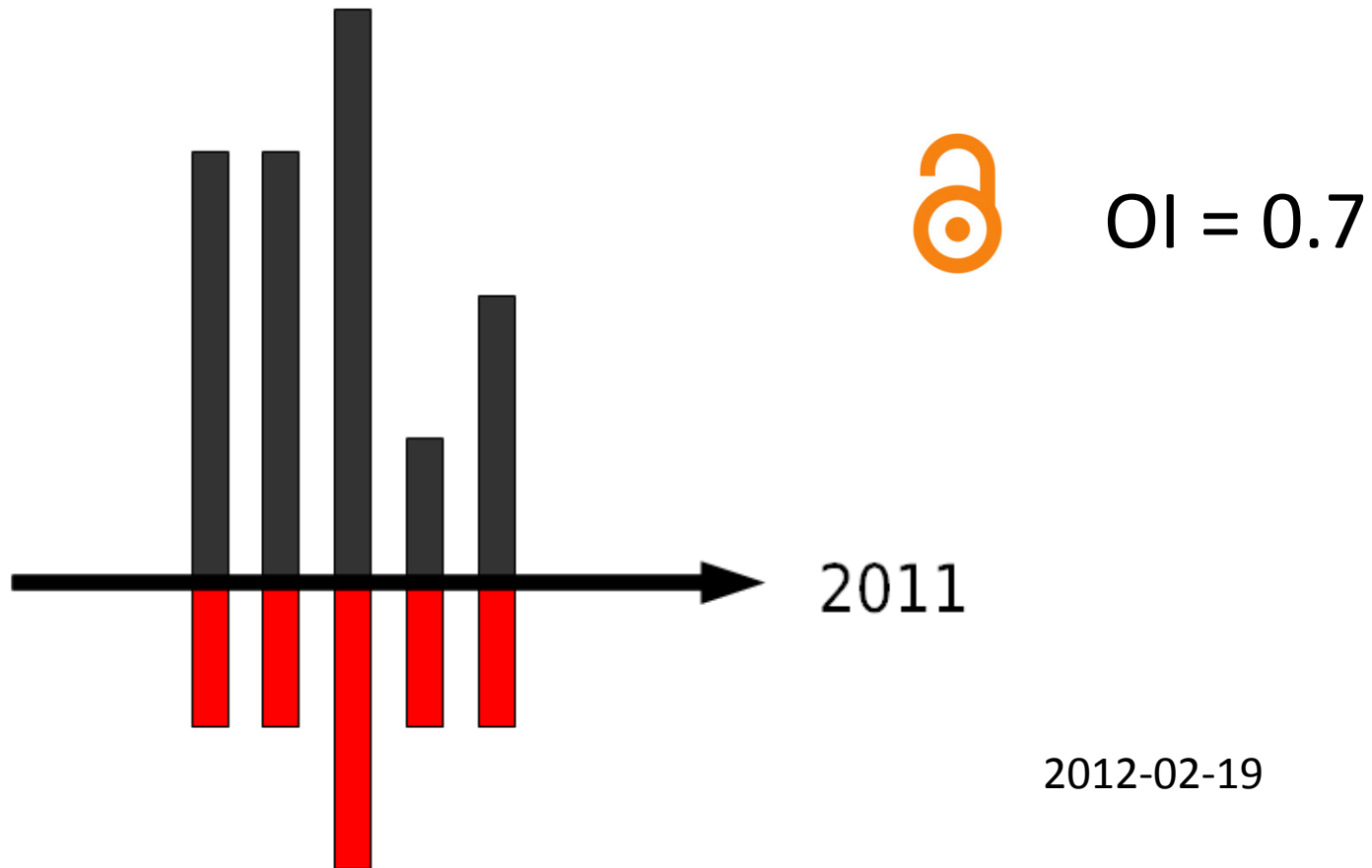
iSchool Dean *Openness Index* 2011

- 1 Georgia Tech, Illinois, IU-SIC, Syracuse,
Toronto, UCD, UC-Irvine, UCL, UMD,
UMich, UNC, UW
- 0.5 Tampere, UBC
- 0.25 CMU
- 0.2 Humboldt
- 0 IU-SLIS, PSU, RSLIS, Sheffield, UK, UMBC

No data: Drexel, FSU, Pitt, Rutgers, SMU,
UC-Berkeley, UCLA, UNT, UT-Austin

No evaluation: Nanjing, Wuhan

Visualising Individual Openness



What to measure?

- Just proportion?
 - Should 1 closed access paper get the same value as 10 closed access papers?
 - Do pre-peer review papers count as anything?
 - Wiley-Blackwell only allow pre-review versions to be archived in an IR
- OA-gratis v OA-libre
 - Same value?
- Publication OA status
 - Equal value for an OA journal v self-archive?
- Publication type
 - Is C + J a reasonable middle ground?
 - Book chapters , books, ...
- Location
 - Archive-friendly v (temporary) faculty web space

OA-libre v OA-gratis

- OA-gratis
 - Free to read
- OA-libre (CC-BY)
 - OA-gratis
 - +
 - Free to text-mine
 - Free to remix
 - etc.

Preservation

- “Knowing that faculty Web sites are deleted after they leave the university and that the maintenance of departmental servers varies over time, some faculty interviewed in 2006 expressed concern about the preservation of their ‘legacy’”
 - Covey (2009)
- Simple OI doesn't take into account the location
 - personal web space as good as IR?

Legacy-Friendly OI

- Yes: open access journal, publisher DL (e.g. HICSS at IEEE), IR
- No: personal web/FTP space
- ? Research group web space

Effective Openness Index

- After taking account of existing copyright agreements
- What proportion of the research items you *could* have made open – have you *actually* made open?
 - Difficult to automate
 - Individual author addenda, special issue one-off copyright arrangements

Open Acce\$\$ Cost Index

- How much does it cost to read *your* research?
- For all the items that are not open
 - Sum the cost to access them
 - Simple: individual independent items
 - Complex: bulk deals, joining societies etc
- At the institution level – what is UNC's OACI?
 - State legislature “pays twice”
 - What would it cost v what does it actually cost the public?

9162

I N A S S E M B L Y

January 31, 2012

Introduced by M. of A. HEVESI -- read once and referred to the Committee
on Governmental Operations

AN ACT to establish the taxpayer access to publicly funded research act

THE PEOPLE OF THE STATE OF NEW YORK, REPRESENTED IN SENATE AND ASSEMBLY,
DO ENACT AS FOLLOWS:

- 1 Section 1. Short title. This act shall be known and may be cited as
2 the "taxpayer access to publicly funded research act".
3 S 2. (a) As used in this act, the term:
4 (i) "Agency" shall mean any subdivision of the state, including, but
5 not limited to any agency, commission or authority;
6 (ii) "Direct research" shall mean research resulting directly from
7 state agency funding; and
8 (iii) "Policy" shall mean the public access policy established pursuant
9 to subdivision (b) of this section.



Purchase this Article:

NuggetMine: intelligent groupware for opportunistically sharing information nuggets

Jeremy Goecks, Dan Cosley

Step 1 Sign in or create a free Web account

Sign in with your Web account
Web Account

Password

[sign in](#)

[Need sign in help?](#)

-or-


Create a free Web account

Email Address

[continue](#)

[cancel](#)

Step 2 Pricing and access depends on your membership or subscriptions with ACM.

☐ Download the article as  PDF



ACM Member

\$10.00

ACM Student Member

\$5.00

Non Member

\$15.00

Join ACM and subscribe to the Digital Library to gain access to all materials

- | | |
|---|----------|
| <input type="radio"/> Professional Membership with a Digital Library subscription | \$198.00 |
| <input type="radio"/> Student Membership with a Digital Library subscription | \$42.00 |

Join a sponsoring SIG and access this article and all other materials they sponsor

- | | |
|--|---------|
| <input type="radio"/> SIGART Membership | \$35.00 |
| <input type="radio"/> SIGCHI (includes print interactions) | \$55.00 |

Research. Rent. Read.



Available for
rent through
DeepDyve

[Read this article](#)

View-only access for
24 hours

\$2.99

[DeepDyve](#) is the largest
online rental service for
scientific, technical and
medical research.

The Rental Market for Research

deepdyve BETA

Learn More Sign Up Login

Browse Search

MILLIONS OF ARTICLES AT YOUR FINGERTIPS
RENT PREMIER RESEARCH ARTICLES AND SAVE UP TO 90%

Try it today risk-free for 14 days! [Learn more](#)

Research
Search millions of articles from thousands of journals

Rent
Rent premium articles and save up to 90%

Read
Read the entire article from the DeepDyve site

Sign Up for a 14-Day Free Trial

The digital library as place

Jeffrey Pomerantz; Gary Marchionini

Journal of Documentation, Volume 63 (4): 29
Emerald Publishing – Jul 31, 2007

[More Like This](#)

\$3.99

Tools-at-hand and learning in multi-session, collaborative search

Capra, Robert; Marchionini, Gary; Velasco-Martin, Javier; Muller, Katrina

Association for Computing Machinery — Apr 10, 2010

[More Like This](#)

\$2.99

Open Science

- “Open science is the idea that scientific knowledge of all kinds should be openly shared as early as is practical in the discovery process.”
 - Nielsen (2011)
- Data
- Code

The (in)famous Jones quote



“Even if WMO [World Meteorological Organization] agrees, I will still not pass on the data. We have 25 or so years invested in the work. **Why should I make the data available to you, when your aim is to try and find something wrong with it?**” 21 Feb 2005

Professor Phil Jones

Director of the Climatic Research Unit (University of East Anglia, UK) to “amateur climate researcher” (and sceptic) Warwick Hughes about CRU weather station data

http://en.wikipedia.org/wiki/Phil_Jones

Data Archiving Index

- “Lonnie Thompson is **one of the worst archiving offenders in paleoclimate**, and that’s a real beauty contest.”

<http://climateaudit.org/2006/07/06/new-thompson-article-at-pnas/#comment-55284>

- This statement cannot be evaluated with current metadata – but it implies the existence of paper and author-specific data archiving metadata
 - and a ranking of researchers ordered by their data archiving rates

Data Archiving Index

- If a paper creates/uses a dataset does it
 - Archive or uniquely identify the data in an open location
- = papers with open data / papers with data
- Lots of data release issues to get a 'fair' index
 - Privacy, commercial, ...

Code Archiving Index

- Repeat Data Index but for source code
- “anything less than the release of source programs is intolerable for results that depend on computation”
 - Ince, Hatton & Graham-Cumming (Nature 2012)

Faculty & Copyright

- Unaware, confused, ignoring
 - Antelman (2006), Kim (2010)
- “Many faculty members are simply not aware of publisher policies. Many have a meager understanding of copyright. Some faculty have little respect or concern for publisher policy or copyright.” – Covey (2006)

- “many authors are not aware of the terms of their copyright agreements and many believe they hold copyright to their own works. At the same time, authors who know their rights or publisher requirements do not necessarily abide by them”

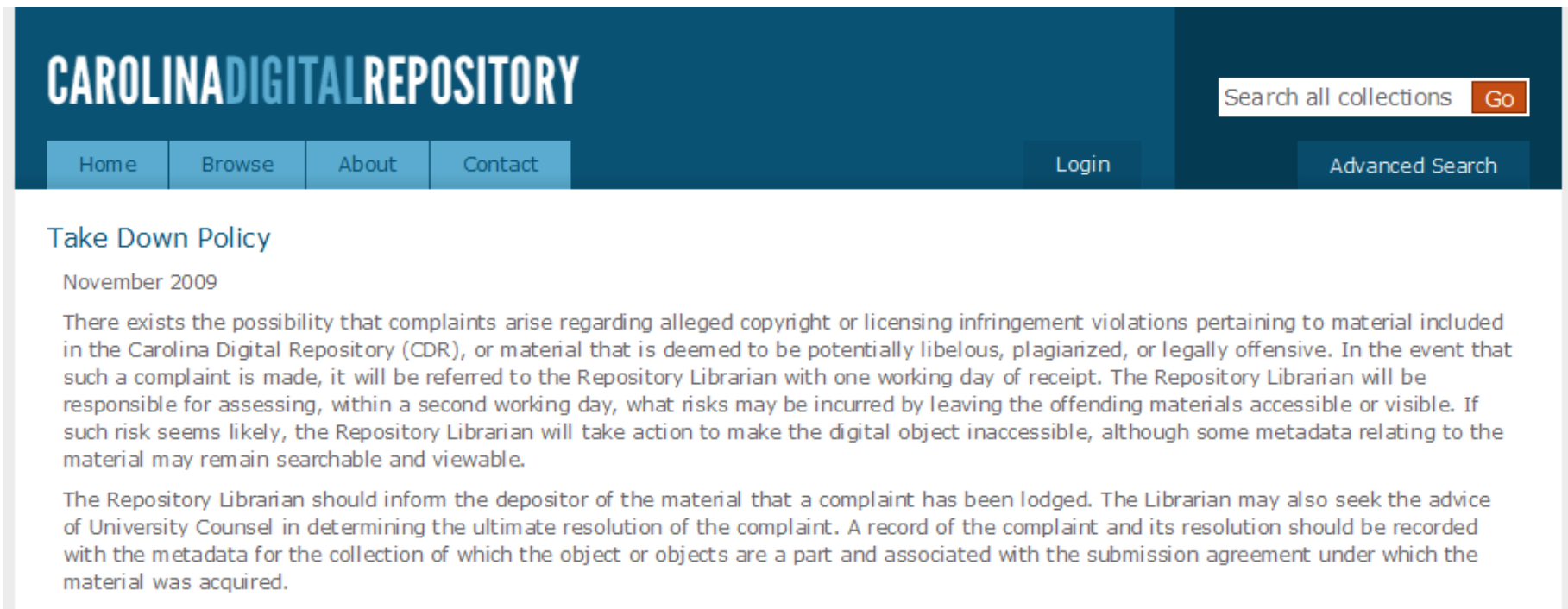
— Antelman (2006)

Illegality Index

- How much of *your* work is available in *contravention* of existing copyright agreements?
- = num. illegally available / num. of items
- At the Repository level this is a measure of:
 - Workflow copyright checking effectiveness
 - Faculty mistakes

Reversion Index

- Open access can disappear
 - Change in OI, but worth separating?



The screenshot shows the top navigation bar of the Carolina Digital Repository website. The header is dark blue with the text "CAROLINADIGITALREPOSITORY" in white. To the right is a search bar with the text "Search all collections" and a red "Go" button. Below the header is a navigation menu with links: Home, Browse, About, Contact, Login, and Advanced Search. The main content area is white and displays the "Take Down Policy" page. The page title is "Take Down Policy" in blue, followed by the date "November 2009". The text of the policy is in a smaller, grey font. It states that there is a possibility of complaints regarding copyright or licensing infringement, and that the Repository Librarian will be responsible for assessing the risks and taking action to make the digital object inaccessible if the risk seems likely. It also mentions that the Librarian should inform the depositor of the material that a complaint has been lodged and seek the advice of University Counsel in determining the ultimate resolution of the complaint.

CAROLINADIGITALREPOSITORY

Search all collections **Go**

Home Browse About Contact Login Advanced Search

Take Down Policy

November 2009

There exists the possibility that complaints arise regarding alleged copyright or licensing infringement violations pertaining to material included in the Carolina Digital Repository (CDR), or material that is deemed to be potentially libelous, plagiarized, or legally offensive. In the event that such a complaint is made, it will be referred to the Repository Librarian with one working day of receipt. The Repository Librarian will be responsible for assessing, within a second working day, what risks may be incurred by leaving the offending materials accessible or visible. If such risk seems likely, the Repository Librarian will take action to make the digital object inaccessible, although some metadata relating to the material may remain searchable and viewable.

The Repository Librarian should inform the depositor of the material that a complaint has been lodged. The Librarian may also seek the advice of University Counsel in determining the ultimate resolution of the complaint. A record of the complaint and its resolution should be recorded with the metadata for the collection of which the object or objects are a part and associated with the submission agreement under which the material was acquired.

<https://cdr.lib.unc.edu/external?page=about.takeDownPolicy>

Clayton Weatherston's work erased

NEIL REID

Last updated 05:00 11/12/2011

 Text Size

 Print

 Share



KIRK HARGREAVES/The Press

VICTIMS OF CRIME: Gil and Lesley Elliott, parents of Sophie Elliott, who was murdered by Clayton Weatherston.

.....



Red-faced Otago University officials have deleted killer Clayton Weatherston's work from an online archive after offending the family of murder victim Sophie Elliott.

The university spiked seven papers written by Weatherston between 2001 and 2006 after being approached by the Sunday Star-Times last week.

Weatherston completed a PhD at the university before tutoring in economics. But in 2008 the 36-year-old killed Elliott, 22, stabbing her more than 200 times in a bedroom of her family's Dunedin home.

Weatherston, who had taught Elliott before they embarked on a five-month relationship, was later jailed for a minimum non-parole period of 18 years.

After the case, university officials tried to erase references to Weatherston from its records. But the seven items by the convicted murderer, covering professionalism in rugby, crowd sizes in Super Rugby, and evaluating rugby's former NPC, slipped through the cracks.

A university spokeswoman did not know Weatherston's work was still online when contacted by the Sunday Star-Times, but rang back later, saying references to Weatherston had been removed.

Author deleted

Now showing items 1485-1504 of 1599

[Previous Page](#) [Next Page](#)

Author Name

[Ward, Roy](#)

[Wasicek, Armin](#)

[Wasson, Marilyn](#)

[Watkins, Leah](#)

[Watson, John R](#)

[Watt, Clayton](#)

[Watts, Michael](#)

[Watts, Michael John](#)

[Watts, Mike](#)



[Weinstein, Phil](#)

[Weir, Megan](#)

[Informal institutions and cross-country income differences](#)

..., CR Weatherston - 2006 - otago.ourarchive.ac.nz

There is a growing literature which analyses, using cross-country data, whether institutions or geography is the most important deep determinant of income per capita. The empirical proxies used for institutions in this literature, however, focus only on formal institutions, ...

[Cited by 12](#) - [Related articles](#) - [Library Search](#) - [All 19 versions](#)

[\[PDF\] from ourarchive.ac.nz](#)

[Measuring competitive balance in professional team sports using the Herfindahl-Hirschman Index](#)

..., M Ryan, CR Weatherston - *Review of Industrial Organization*, 2007 - Springer

Abstract Competitive balance in sports leagues is essentially concerned with inequality in match and championship outcomes. Measures of inequality or concentration from the income distribution and industrial organization literatures have, therefore, often been used ...

[Cited by 15](#) - [Related articles](#) - [Library Search](#) - [BL Direct](#) - [All 9 versions](#)

[\[PDF\] from ourarchive.ac.nz](#)



Otago University Research Archive

OUR Archive

Login

[OUR Archive Home](#) → [Sign in](#)

The file is restricted

The file you are attempting to access is a restricted file and requires credentials to view. Please login below to access the file.

Search

Go

(Missing) Openness Metrics

- Openness Index
 - Effective Index: given past copyright
- Openness Cost Index
- Illegality Index
- Data Archiving Index
- Code Archiving Index
- Reversion Index

Value of (yet) more metrics?

- “What gets measured gets noticed”
- “What gets measured gets managed” - Drucker
- “What gets measured gets to frame the issue”
- “The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor.”
 - http://en.wikipedia.org/wiki/Campbell%27s_law
 - Citation cartels, coercive citation etc
 - El Naschie, *Chaos, Solitons & Fractals*
 - <http://elnaschiewatch.blogspot.com/>
- Can you subvert an altmetrics measure?
- Can you subvert an Openness Index?