

Measuring what matters: metrics, incentives and openness

David Nichols

Department of Computer Science
University of Waikato
New Zealand

cs.waikato.ac.nz/~daven



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

Michael Twidale

School of Information Sciences
University of Illinois
USA

ischool.illinois.edu/people/faculty/twidale



ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Outline

- Metrics for Openness
- Metrics and Incentives

All shall have metrics

"It would be quite easy to generate a list of over 100 different (nested) measures to which each individual academic in the UK is now (potentially) subject. However, for our purposes here, we will consider just six domains: citations; workload models; transparent costing data; research assessments; teaching quality assessments; and university league tables. ...

The [h-]index has become reified; it has taken on a life of its own; a number that has become a rhetorical device with which the neoliberal academy has come to **enact 'academic value'**."

Burrows (2012) Living with the h-index? Metric assemblages in the contemporary academy. *The Sociological Review*, <https://doi.org/10.1111/j.1467-954X.2012.02077.x>

A surprisingly obvious way to incentivize openness in academia



Stacy Konkiel, 25th October 2016

"I'm starting to see that academia's been approaching evaluation metrics from the wrong angle: most institutions simply measure what can be **easily counted**, rather than using carefully chosen data to measure their progress towards embodying important scholarly values."

"let's use value-aligned evaluation practices to **incentivize** the "enriching" practices we want to **encourage**"

<https://www.altmetric.com/blog/a-surprisingly-obvious-way-to-incentivize-openness-in-academia/>

LSE Impact Blog

Getting our hands dirty: why academics should design metrics and address the lack of transparency.

Chris Elsdon, Sebastian Mellor and Rob Comber

“complement critiques of metrics with getting our hands dirty in reflectively and critically **designing metrics.**”

<http://blogs.lse.ac.uk/impactofsocialsciences/2016/04/06/getting-our-hands-dirty-why-we-should-design-metrics/>

Starting point

- “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>
- Data unavailability → openness in general
- Ideas initially developed over talks at iSchools (UNC, UIUC, IU)
 - 2012

Nichols, Twidale & Cunningham (2012) Metadatapedia: a proposal for aggregating metadata on data archiving. *Proceedings of the 2012 iConference*
<https://doi.org/10.1145/2132176.2132224>

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

Citation-based Metrics

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

altmetrics



- Citations are not enough
- Diverse article-level metrics
- *Impactstory*
- Views, downloads, bookmarking etc. in 'non-academic' venues:
 - Wikipedia
 - Blogs, Twitter, Facebook
 - Mendeley, CiteULike
 - SlideShare



<http://blog.impactstory.org/new-impactstory-logo/>



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.”

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

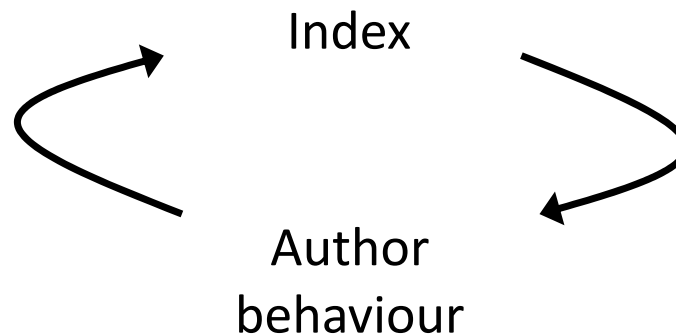
Unfortunately it's 'publish or perish' and not 'share or be shunned' in the scientific ego-system. #SOPA #RWA

- Let's fix that...

Nichols, D. M., & Twidale, M. B. (2017). Metrics for openness. *Journal of the Association for Information Science and Technology*, 68(4), 1048–1060.

Openness Index

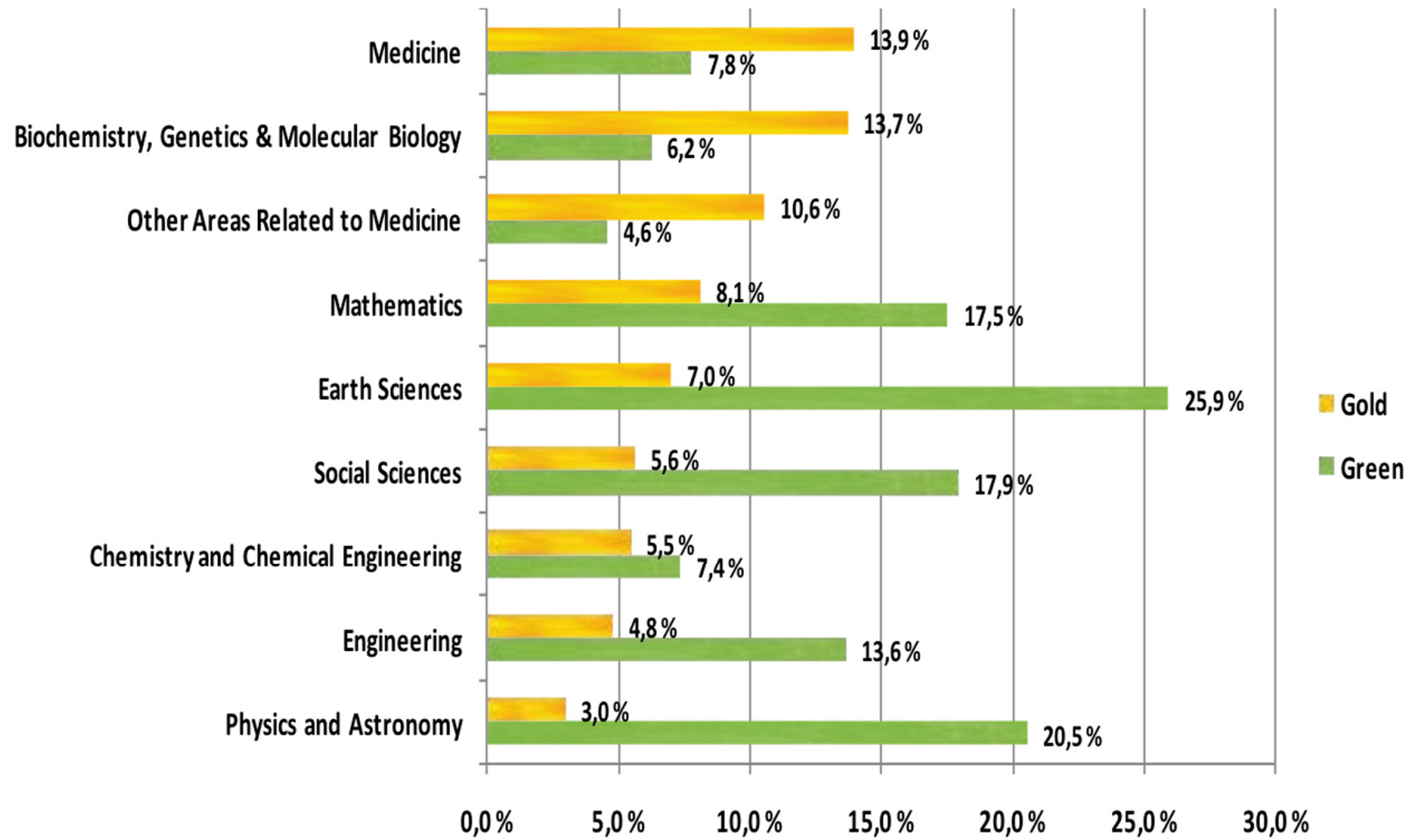
- Of the items *you* have published how many are free for anyone to read?
 - Practically: if someone searches for the title of your paper in Google do they find a copy they can read?



How to measure Openness?

- Grain size:
 - Person, Group, Dept, School, Institution, System
 - Journal/Conference/etc, Publisher
 - Sub-discipline, Discipline
 - Funder
 - State/Province, Country

Cross-Discipline



Björk et al. (PLoS ONE 2009)

Institution

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

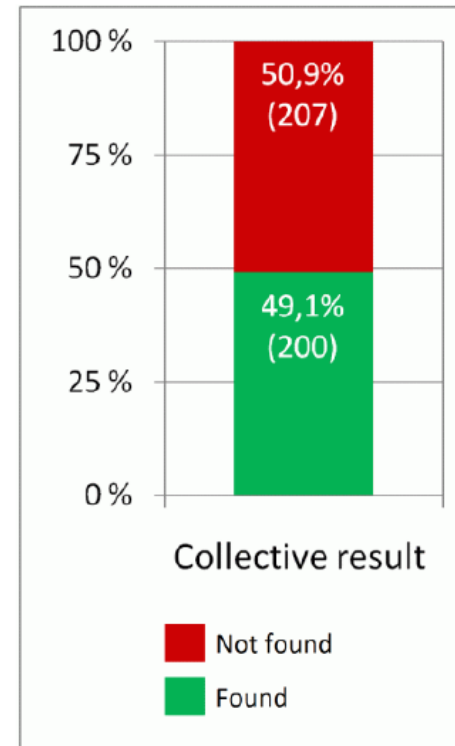


Figure 1: Collective Open Access availability.

Sub-Institution

| Faculty (% of the sample) | Google | Google Scholar | HELDA | Open DOAR | Scientific Commons | Collective |
|---|--------|----------------|-------|-----------|--------------------|------------|
| Faculty of Biological and Environmental Sciences (4.9%) | 40.0 % | 65.0 % | 0.0 % | 5.0 % | 10.0 % | 70.0 % |
| Faculty of Veterinary Medicine (3.4%) | 35.7 % | 35.7 % | 7.1 % | 35.7 % | 28.6 % | 50.0 % |
| Independent institutes (10.3%) | 52.4 % | 42.9 % | 2.4 % | 21.4 % | 23.8 % | 59.5 % |
| Faculty of Pharmacy (2.7%) | 9.1 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 9.1 % |
| Faculty of Arts (2.5%) | 40.0 % | 20.0 % | 0.0 % | 0.0 % | 10.0 % | 50.0 % |
| Faculty of Behavioural Sciences (5.2%) | 38.1 % | 28.6 % | 0.0 % | 14.3 % | 9.5 % | 42.9 % |
| Faculty of Medicine (47.2%) | 44.8 % | 41.7 % | 0.5 % | 13.5 % | 13.5 % | 49.5 % |
| Faculty of Agriculture and Forestry (6.9%) | 25.0 % | 32.1 % | 0.0 % | 3.6 % | 7.1 % | 35.7 % |
| Faculty of Science (15.2%) | 50.0 % | 38.7 % | 0.0 % | 21.0 % | 24.2 % | 56.5 % |
| Faculty of Law (1.2%) | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % |
| Faculty of Theology (1.0%) | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % |
| Faculty of Social Sciences (6.1%) | 56.0 % | 40.0 % | 0.0 % | 16.0 % | 16.0 % | 60.0 % |


 Found

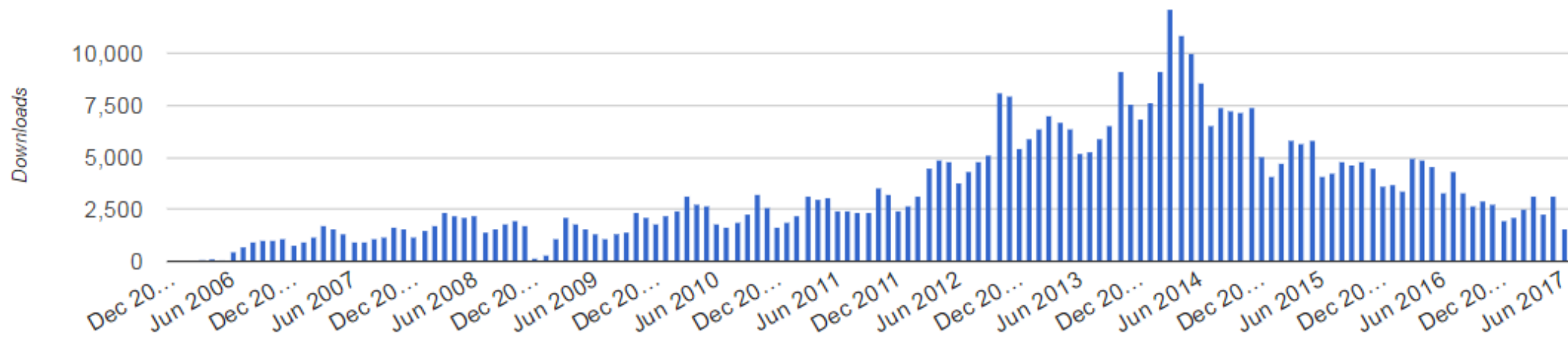
Figure 3: Open Access availability by faculty.

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

Faculties of the University of Helsinki

Statistics Overview

Skitmore, Martin



Key Figures

Collection

394 Works

96.9% Full-text [?]

80.9% Open-access [?]

Total Usage

474,116 Downloads [?]

99.1% External [?]

Last 28 Days

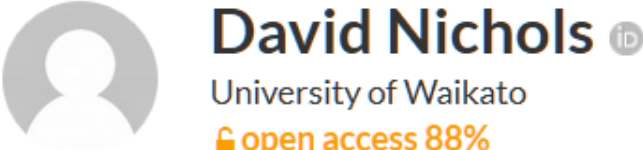
2,701 Downloads





<https://eprints.qut.edu.au/statistics/creators/Skitmore%2C%20Martin/>

Individuals


- Missing?
- Each author gets an *Openness Index* defined across their publications
- Conferences + journals →
 - Twidale = 0.82
 - Nichols = 0.91



David Nichols 
University of Waikato
 open access 88%

[OVERVIEW](#) [ACHIEVEMENTS](#) [TIMELINE](#) [PUB](#)

ACHIEVEMENTS [view all](#)

 **Open Access** Top 10%
87% of your research is free to read online. This level of availability puts you in the top 8% of researchers.

<https://impactstory.org/u/0000-0003-0321-7267>

What to measure?

- 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?
 - *Practical Openness Index* (POI)

Effective Openness Index (EOI)

- *After* taking account of existing copyright agreements
- Of the items an author *could* have made open, how many *are* actually open?
 - Difficult to automate
 - Individual author addenda
special issue one-off copyright arrangements

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?

Preservation-Friendly Openness Index (PFOI)

= open items in legacy-friendly locations / all items

Yes IR, open access journal, publisher DL

No personal web/FTP space

? research group web space

PFOI values are Twidale = 0.45 and Nichols = 0.84

Acce\$\$ Index

- How much does it cost to access 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.

Twidale = US\$1,484 (over 20 items), Nichols = US\$183 (over four items)
in 2016

John Q. Public Institutional Cost Index

- If one person had the time ...
- How much would it cost to access all of your university's research output in 2016?

UI Mission

“To create a brilliant future for the University of Illinois in which the students, faculty and staff thrive and the *citizens* of Illinois, the nation and the world benefit”

<https://www.uillinois.edu/about/mission/>

citizens ≠ people with access via institutional subscriptions
citizens ≠ people with money who can go through paywalls

iBiosphere Rich Jorgensen

Yes, ethical science now requires each of us to publish in a way that makes our publicly funded research truly accessible to all [#RWA](#)

26 minutes ago

<https://whoneedsaccess.org/>

Actual Individual Purchase Index

- How much was actually paid to publishers to access *your* outputs?
- How much was *actually paid* to publishers to access your institution's research outputs?
- Royalty statements for books do contain this data
 - for the rest ...?

The 'Open Paper'

- Defining measures across sets of papers
 - Interest comes from the set membership
- Set = references in a paper
- *Open Reference Index* (ORI) is the proportion of all the cited works of a paper that are themselves OA.
- The ORI for the *Metrics for Openness* paper at the time of writing was 0.92

Open Papers?

- a *Fully Open Paper*:
 - Which itself is Open Access, and
 - Where all of the references it relies on are Fully Open Papers
- Does an *Open Paper* exist in the literature?
- Does a *Fully Open Paper* exist?

Acce\$\$ Support Index

- What does it cost to access the closed items in a reference list?
- an *Acce\$\$ Index* calculated over the set of closed references it is an *Acce\$\$ Support Index*.
- For *Metrics for Openness* the *Acce\$\$ Support Index* for the eight non-open items cited was US\$3,662 (2016)

Open Science/Research

“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

Data Archiving

- “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 ...
 - Implies existence of paper and author-specific data archiving metadata
 - and a ranking of researchers ordered by their data archiving rates



Empty rhetoric over data sharing slows science

Governments, funders and scientific communities must move beyond lip-service and commit to data-sharing practices and platforms.

12 June 2017

<http://www.nature.com/news/empty-rhetoric-over-data-sharing-slows-science-1.22133>

Journals' Retreat From Data-Sharing Mandate Puts Onus on Universities and Government

By Paul Basken | JUNE 12, 2017 ✓ PREMIUM

<http://www.chronicle.com/article/Journals-Retreat-From/240323>

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

- “anything less than the release of source programs is intolerable for results that depend on computation”

Ince, Hatton & Graham-Cumming (Nature 2012)

Biostatistics (2011), **12**, 4, pp. 637–652

doi:10.1093/biostatistics/kxr002

Advance Access publication on February 5, 2011



Estimating the acute health effects of coarse particulate matter accounting for exposure measurement error

Illegality Index

- How much of your work is available in contravention of existing copyright agreements?
 - = illegally available papers / all papers
- At the Repository level this is partially a measure of:
 - Workflow copyright checking effectiveness and academics' behaviour
- ... But other people can illegally distribute your work!



Metrics for Openness

- Openness Index
 - Effective Openness Index
 - Preservation-Friendly Openness Index
- Acce\$\$ Index
 - *John Q. Public* Institutional Cost Index
- Actual Individual Purchase Index
- The ‘Open Paper’
- Data Archiving Index
- Code Archiving Index
- Illegality Index

Metrics as Designed Artefacts

- Designed not discovered
- A design space
- Multiple, often conflicting goals
- Multiple constraints
- Invidious tradeoffs
- Dealing with edge cases
- Use in context
- Incentives
- Appropriation & misuse

"We should understand metrics as **designed artefacts**."

<http://blogs.lse.ac.uk/impactofsocialsciences/2016/04/06/getting-our-hands-dirty-why-we-should-design-metrics/>

From Social Science to Social Engineering

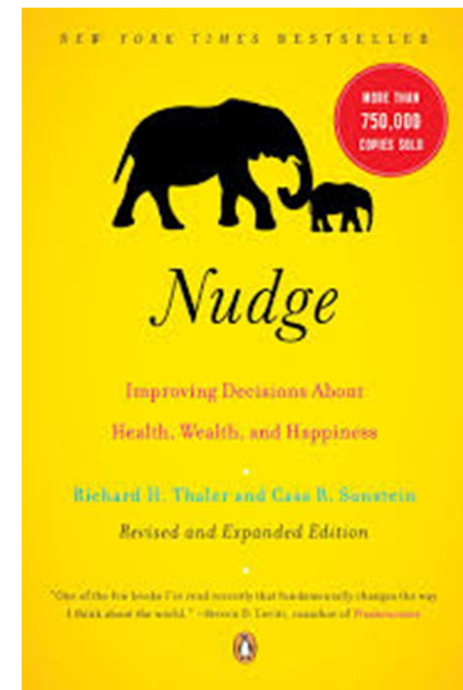
- In the natural sciences a metric typically measures a thing that has no awareness of being measured: the temperature of a star, the capacity of a battery, etc.
- When we measure people, there is the possibility that they know that they are being measured.
- When you are measuring scientists, that possibility becomes a probability,
- When you are measuring scientific publishing, it becomes a certainty –
- Those being measured may change their behaviour.

Campbell's Law

- "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
- But is that always a bad thing?
- Can it sometimes be a good thing?
- What if the change is in a desirable direction?

Metrics as Persuasive Technologies

- Metric design should not be considered purely as an objective scientific process but as a behavioural, managerial and political act, with the potential to nudge behaviour closer or further from a particular desired state.
 - (Thaler & Sunstein, 2008)
- The metric can be constructed as a form of advocacy
 - intended to change in a certain direction



Unintended Consequences

- Unintended OK, but unanticipated?
- Testing
 - White hat testing
- Considering use in context
- Account for human behavior
- Account for human innovation



Exploiting Campbell's Law to Nudge

- Devise metrics that encourage behavioural change towards greater openness.
- Gaming: The negative effect of Campbell's law
 - where people do an action that improves the metric while failing to improve access, or improving access less than if they had made their papers fully open.
- A challenge for metric designers is to minimize the incentives for subversion.
 - Ideally design a metric such that subverting it is more effort than engaging in the desirable change in behaviour.
- Minimize disincentives
 - Don't want a metric that accidentally discourages certain kinds of openness
- Can we measure the effect of the metric in changing that which it measures?

Sociotechnical Systems Design and Power

- Metrics considered as (very small) sociotechnical systems
- Metric design is not just Information Science but also Sociotechnical Engineering.
- The power of the metric designer
- Tendency to design metrics that make you look good
- Consider special cases that matter to you
- Overlook special cases that don't matter to you
- Power erodes inconvenience for the powerful

It's proxies all the way

- Measuring 'impact' or 'research effectiveness' or 'quality'
- Citations, reads, influence, altmetrics
- Errors, type I & II, error bias
- Levels of use and appropriateness

**The Agony and the Ecstasy—
The History and Meaning of the Journal Impact Factor**

Presented by

Eugene Garfield
Chairman Emeritus, Thomson ISI

I had considered as an alternative title for my talk "Citation Sanity and Insanity -- the Obsession and Paranoia of Citations and Impact Factors." Others might have preferred "Uses and Abuses of Impact Factors." – Eugene Garfield



Journal Impact Factor as a case of metric appropriation

- Developed as a measure of quality of the journal to inform library purchasing decisions (a proxy)
- Now also used as a measure of the quality of the **papers in** the journal (a proxy-proxy)
- Even used as a measure of the quality of the **researcher** who publishes papers in that journal (a proxy-proxy-proxy)
- Perhaps some day as a measure of the quality of the department or the **university** that employs the researcher who publishes in that journal (a proxy-proxy-proxy-proxy)

What's in a name? In social engineering: – A Lot

- Rather shocking to the purist scientist, particularly in the natural sciences.
- The value of a metric for drug resistance is unlikely to change more rapidly just because of what you call it.
- But an Acce\$\$ Index score may
- Journal Impact Factor renamed as the **US Librarians' Purchasing Indicator**
 - “In submitting your CV please include the Impact Factor for each journal publication”
 - “In submitting your CV please include the US Librarians' Purchasing Indicator for each journal publication”

Metrics used in other metrics



- Shanghai Jiao Tong Academic Ranking of World Universities
 - uses data about publication venues (Science and Nature) and citations.
- Possible uses of openness score in derived measures
- Must consider the pathological case - even if you don't want to
- Another risk if a poor score correlates with prestige.

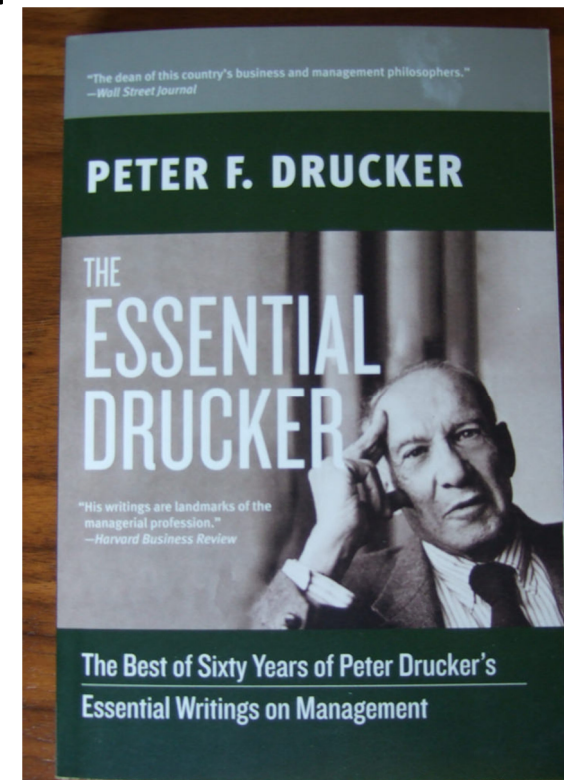
The power of hypothetical metrics

- As critiques
 - Challenge the authority of imposed metrics
 - Fight metrics with metrics
 - Gresham’s Law: “Bad money drives out good”
 - Debase the currency of metrics by minting new ones
 - Critical metrics
 - Administrator Metrics
- As alternatives
- As nudges
- Even if easy to compute openness metrics are never deployed
 - openness metrics can still have a rhetorical effect.
- Just talking about hypothetical openness metrics seems to frame a discussion and seems to change our understanding of a wider set of issues.



Value of (yet) more metrics?

- “What gets measured gets noticed”
 - “What gets measured gets managed” - Drucker
 - “What gets measured gets to frame the issue”
-
- Can you subvert an altmetrics measure?
 - Can you subvert an Openness Index?



Harnessing Paranoia for Good

- Metrics can have effects on behaviour even without explicit enforcement.
- Just a thought experiment, but one that may yet turn out to be a testable hypothesis.
- The case of the tenure process in the USA.
 - Commendable aims of ensuring academic freedom
 - But it seems to engender a substantial amount of fear, even paranoia, amongst junior faculty.
- Departments and universities have various explicit and implicit factors used in tenure cases
 - Junior faculty are advised accordingly.
- Ex: publications in certain journals may be strongly encouraged or even required.
- If a university's promotion rules required a faculty member to submit their OI as part of their tenure case materials, we would confidently expect a rapid and substantial increase in openness scores for junior faculty.
- For our Machiavellian purposes, this may not even be necessary.
- Just the suspicion that the OI *might* be used in the tenure process may be enough to channel paranoia into becoming more open.
- The fear that an external reviewer *might* use the OI in her assessment (assuming it was as easy to discover as say the h-index is on Google Scholar) may be sufficient,
 - even if openness is not explicitly considered in the home department's deliberations.