These slides were made by
Tim Brody and Stevan Harnad
(Southampton University)

Permission is granted to use them
to promote open access and self-archiving
as long as their source is acknowledged.
The Research-Impact Cycle

Open access to research output maximizes research access maximizing (and accelerating) research impact (hence also research productivity and research progress and their rewards)
Limited Access: Limited Research Impact

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

12-18 Months
Maximized Research Access and Impact Through Self-Archiving

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

Pre-Print is self-archived in University’s Eprint Archive

Post-Print is self-archived in University’s Eprint Archive

New impact cycles:
Self-archived research impact is greater (and faster) because access is maximized (and accelerated)

New impact cycles:
New research builds on existing research

12-18 Months
Research Impact

I. measures the size of a research contribution to further research ("publish or perish")

II. generates further research funding

III. contributes to the research productivity and financial support of the researcher’s institution

IV. advances the researcher’s career

V. promotes research progress
“Online or Invisible?” (Lawrence 2001)

“average of 336% more citations to online articles compared to offline articles published in the same venue”

http://www.neci.nec.com/~lawrence/papers/online-nature01/
“Correlation between RAE ratings and mean departmental citations +0.91 (1996) +0.86 (2001) (Psychology)”

“RAE and citation counting measure broadly the same thing”

“Citation counting is both more cost-effective and more transparent”

(Eysenck & Smith 2002)

http://psyserver.pc.rhbnc.ac.uk/citations.pdf
The objective of open-access
(and the motivation that will induce researchers to provide it)
is:

• **not** to quarrel with, ruin or replace journals (at all)

• **nor** is it to solve the budgetary problems of libraries (and yet…)

• **nor** is it to provide access to teachers - students - the general public (and yet…)

• **nor** is it to provide access to the Developing World (and yet…)}
The objective of open-access is:

to maximize research impact

by maximizing research access
Some old and new scientometric ("publish or perish") indices of research impact

- Peer-review quality-level and citation-counts of the journal in which the article appears
- Citation-counts for the article
- Citation-counts for the researcher
- Co-citations, co-text, "semantic web" (cited with whom/what else?)
- Citation-counts for the preprint
- Usage-measures ("hits," webmetrics)
- Time-course analyses, early predictors, etc. etc.
Time-Course of **Citations (red)** and **Usage (hits, green)**


1. Preprint or Postprint appears.
2. It is downloaded (and sometimes read).
3. Eventually citations may follow (for more important papers)...
4. This generates more downloads...
5. More citations...
Most papers are not cited at all

Usage Impact

is correlated with Citation Impact

(Physics ArXiv: hep, astro, cond, quantum; math, comp)

http://citebase.eprints.org/analysis/correlation.php

(Quartiles Q1 (lo) - Q4 (hi))

\textbf{All} \quad r=.27, \quad n=219328

Q1 (lo) \quad r=.26, \quad n=54832

Q2 \quad r=.18, \quad n=54832

Q3 \quad r=.28, \quad n=54832

Q4 (hi) \quad r=.34, \quad n=54832

\textbf{hep} \quad r=.33, \quad n=74020

Q1 (lo) \quad r=.23, \quad n=18505

Q2 \quad r=.23, \quad n=18505

Q3 \quad r=.30, \quad n=18505

Q4 (hi) \quad r=.50, \quad n=18505

(correlation is highest for high-citation papers/authors)

Average UK downloads per paper: 10

(UK site only: 18 mirror sites in all)
The golden rule for open access: reciprocity

(i) Researchers share a common stake with their own Institutions (not their Disciplines) in maximizing their joint research impact

(ii) Institutions share a reciprocal stake in access to one another’s (give-away) research output.
“Self-archive unto others as ye would have them self-archive unto you.”

http://www.ecs.soton.ac.uk/~harnad/Temp/unto-others.doc

http://www.ecs.soton.ac.uk/~harnad/Temp/self-archiving.ppt
1. **Universities:**

   Adopt a policy mandating open access for all university research output:
   
   Extend existing “Publish or Perish” policies to
   
   “Publish with Maximal Impact”

2. **Departments:**

Adopt a departmental policy mandating **Open Access for All Research Output**

Create (and Fill): **OAI-compliant Eprint Archives**

3. University Libraries:

Provide digital library support for university research self-archiving and archive-maintenance

(and if/when university toll-cancellation savings begin to grow, prepare to redirect 1/3 of annual windfall savings to cover open-access journal peer-review service-costs for university research output)

http://www.eprints.org/self-faq/#libraries-do
4. Universities and Research Institutions:

Mandate *open access* for all research output.

http://www.eprints.org/signup/sign.php

Adopt a standardized online-CV
with harvestable performance indicators
and links to open-access full-texts
(*template and demo below*)

http://paracite.eprints.org/cgi-bin/rae_front.cgi
5. Research Funders:

Mandate *open access* for all research output.

See proposal for a UK national policy of *open access* for all refereed research output for research assessment...

http://www.ecs.soton.ac.uk/~harnad/Temp/Ariadne-RAE.doc

...as a model for the rest of the world
Tools for
(a) creating OAI-compliant university eprint archives
(b) parsing and finding cited references on the web,
(c) reference-linking eprint archives,
(d) doing scientometric analyses of research impact,
(e) creating OAI-compliant open-access journals

http://software.eprints.org

http://paracite.eprints.org/


http://citebase.eprints.org/help/

http://psycprints.ecs.soton.ac.uk/
Current Journal Tally: 86% Green!
FULL-GREEN = Postprint, PALE-GREEN = Preprint, GRAY = neither yet
Publishers to date: 100
Journals processed so far: 8689
http://romeo.eprints.org/stats.php
What is needed for open access now:

1. **Universities**: Adopt a university-wide policy of making all university research output open access (via either the gold or green strategy)

2. **Departments**: Create and fill departmental OAI-compliant open-access archives

3. **University Libraries**: Provide digital library support for research self-archiving and open-access archive-maintenance. Redirect 1/3 of any eventual toll-savings to cover open-access journal peer-review service charges

4. **Promotion Committees**: Require a standardized online CV from all candidates, with refereed publications all linked to their full-texts in the open-access journal archives and/or departmental open-access archives

5. **Research Funders**: Mandate open access for all funded research (via either the gold or green strategy). Fund (fixed, fair) open-access journal peer-review service charges. Assess research and researcher impact online (from the online CVs).

6. **Publishers**: Become either gold or green.
OAIster, a cross-archive search engine, now covers over 250 OAI Archives (about half of them Eprints.org Archives) indexing over 3 million items (but not all research papers, and not all full-texts). Below are data for just the full-text research papers with 1990-2003 creation dates.

http://oaister.umdl.umich.edu/o/oaister
The optimal open-access strategy today: **open-access publishing** (5%) plus **open-access self-archiving** (95%):

Open access is possible today for 5% of articles by publishing them in **open access journals**, and for at least 83% (but probably closer to 95%) of the rest by **self-archiving** them.

The optimal dual strategy is hence to

1. **publish your article in an open-access journal** if a suitable one exists and otherwise:
2. **publish your article in a toll-access journal** and also **self-archive** it in your institutional open-access eprint archive.

---

**Proportion of Toll-Access vs. Open-Access Journals Today**

- **1000** Toll Access Journals (Ulrichs)
- **23,500** Open Access Journals

**Romeo "Green/Gold" versus "Gray" Journals**

- Journals already supporting self-archiving (gold)
- Journals already supporting self-archiving (blue/green)
- Journals not yet supporting self-archiving (white)

**The 100% Solution for providing immediate Open Access Today**

- Open Access Journals
- Self-Archiving
Quo usque tandem patientia nostra...?
How long will we go on letting our cumulative daily/monthly/yearly research-impact losses grow, now that the online medium has made it all preventable?

What we stand to gain:

<table>
<thead>
<tr>
<th></th>
<th>Today's Limited Toll-Access Impact</th>
<th>Lost Potential Open-Access Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>23%</td>
<td>77%</td>
</tr>
</tbody>
</table>

.91 correlation with UK research ranking and funding

336% higher impact

Yearly/Monthly/Daily Impact Loss

<table>
<thead>
<tr>
<th>Impact (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>1000</td>
</tr>
<tr>
<td>2000</td>
</tr>
<tr>
<td>3000</td>
</tr>
<tr>
<td>4000</td>
</tr>
<tr>
<td>5000</td>
</tr>
<tr>
<td>6000</td>
</tr>
<tr>
<td>7000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year or Month or Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990/</td>
</tr>
<tr>
<td>1993/</td>
</tr>
<tr>
<td>1995/</td>
</tr>
<tr>
<td>1998/</td>
</tr>
<tr>
<td>2001/</td>
</tr>
<tr>
<td>2004/</td>
</tr>
</tbody>
</table>

open-access impact

<table>
<thead>
<tr>
<th>toll-access impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>open-access impact</td>
</tr>
<tr>
<td>toll-access impact</td>
</tr>
</tbody>
</table>
The two open-access strategies: **Gold** and **Green**

**Open-Access Publishing** (OApub) (BOAI-2)

1. Create or Convert 23,000 open-access journals (1000 exist currently)
2. Find funding support for open-access publication costs ($500-$1500+)
3. Persuade the authors of the annual 2,500,000 articles to publish in new open-access journals instead of the existing toll-access journals

**Open-Access Self-Archiving** (OAarch) (BOAI-1)

1. Persuade the authors of the annual 2,500,000 articles they publish in the existing toll-access journals to also self-archive them in their institutional open-access archives.
Dual open-access strategy

**Gold:** Publish your articles in an open-access journal whenever a suitable one exists today (currently 1000, <5%) and

**Green:** Publish the rest of your articles in the toll-access journal of your choice (currently 23,000, >95%) and self-archive them in your institutional open-access eprint archives.
To Maximize Research Impact:

Research Funders:
1. Mandate open access provision for all funded research via the gold or green strategies
2. (Help cover open-access journal charges)

Research Institutions:
1. Mandate open access provision for all research output via the gold or green strategies
2. (Libraries redirect 1/3 of any eventual toll-cancellation windfall savings toward funding open-access journal charges)

Outcomes:
1. Authors either find an open-access (gold) journal or a green journal to publish in.
2. Gray publishers will turn green.
3. Eventually green publishers might turn gold, but in the meanwhile:
4. Open-access itself increases to 100%.
5. Eventually toll-cancellation savings might increase to 100%
6. If so, then 1/3 of the growing institutional windfall toll-cancellation savings can pay for all institutional gold journal publication charges (peer review)
Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities

http://www.zim.mpg.de/openaccess-berlin/berlindeclaration.html

The pertinent passages:

“Open access [means]:

“1. free… [online, full-text] access

“2. A complete version of the [open-access] work… is deposited… in at least one online repository… to enable open access, unrestricted distribution, [OAI] interoperability, and long-term archiving.

“[W]e intend to… encourag[e].. our researchers/grant recipients to publish their work according to the principles of... open access.”
The BOAI Self-Archiving FAQ (BOAI-1)

http://www.eprints.org/self-faq/

http://www.soros.org/openaccess/
OA vs. Non-OA Citation Impact Advantage (All fields)

On average 319.9 journals/year (incl. self-citations) - Fri Aug 20 20:37:17 2004
Correlations: OAA*OAP = 0.242, OAP*Year = 0.965, OAA*Year = 0.203

- Total Articles: Open Access plus Non-Open Access (Line)
- OAP: Open Access Articles as a Percentage of Total Articles
- OAA: Open Access Citation Impact Advantage: OA/Non-OA Citation ratio minus 100%
UK House of Commons Science and Technology Committee
Recommendation to Mandate Institutional Self-Archiving

“This Report recommends that all UK higher education institutions establish institutional repositories on which their published output can be stored and from which it can be read, free of charge, online.

“It also recommends that Research Councils and other Government Funders mandate their funded researchers to deposit a copy of all of their articles in this way.”

US House of Representatives Appropriations Committee
Recommendation that the NIH should mandate self-archiving

“The Committee… recommends NIH develop a policy… requiring that a complete electronic copy of any Manuscript reporting work supported by NIH grants or contracts be… [made] freely and continuously available upon acceptance of the manuscript for publication in any scientific journal.”
Institutional Archives Registry: (221 Archives Registered)

http://archives.eprints.org/eprints.php

Country
* United States (57)
* United Kingdom (33)
* Canada (17)
* France (15)
* Sweden (13)
* Germany (12)
* Netherlands (12)
* Italy (11)
* Australia (9)
* India (4)
* Brazil (4)
* Hungary (4)
* China (4)
* Denmark (4)
* Mexico (2)
* Ireland (2)
* Austria (2)
* Japan (2)
* Portugal (2)
* South Africa (2)
* Belgium (2)
* Slovenia (1)
* Finland (1)
* Israel (1)
* Norway (1)
* Switzerland (1)
* Croatia (1)
* Peru (1)
* Spain (1)

Archive Type
* Research Institutional or Departmental (117)
* Research Cross-Institution (32)
* e-Theses (27)
* Demonstration (22)
* e-Journal/Publication (11)
* Other (10)
* Database (2)

Software
* GNU EPrints v2 (122)
* GNU EPrints v1 (18)
* DSpace (28)
* ARNO (2)
* DiVA (1)
* CDSWare (1)
* other (49)
Institutional OAI Archive Growth

How OAI Archives for institutional research output have been growing – and how to accelerate their growth

(Data collected and analysed by Tim Brody, doctoral candidate, Electronics and Computer Science, Southampton University)
RoMEO Directory of Publishers who have given their **Green Light** to Self-Archiving
http://www.sherpa.ac.uk/romeo.php
http://romeo.eprints.org

Proportion of journals already formally giving their **green light** to author/institution self-archiving (**already 86%**) continues to grow:

<table>
<thead>
<tr>
<th>Green light to self-archive:</th>
<th>Journals</th>
<th>%</th>
<th>Publishers</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8689</td>
<td>(100%)</td>
<td>100</td>
<td>(100%)</td>
</tr>
<tr>
<td>Neither yet</td>
<td>1283</td>
<td>14%</td>
<td>35</td>
<td>35%</td>
</tr>
<tr>
<td>Preprint</td>
<td>1976</td>
<td>+22% (=86%)</td>
<td>67</td>
<td>+6% (=65%)</td>
</tr>
<tr>
<td>Postprint</td>
<td>5733</td>
<td>64%</td>
<td>59</td>
<td>59%</td>
</tr>
</tbody>
</table>
Percentage of green PUBLISHERS grew from 42% - 58% from 2003-2004.

Percentage of green JOURNALS grew from 55% - 83% from 2003-2004.
Growth of University Eprints.org Archives and Contents

http://archives.eprints.org/eprints.php

Growth of GNU EPrints Archives and Contents
Generated by http://archives.eprints.org/
Archives flagged as 'Research Institutional'. The datestamps of records as exported by the archive's OAI-PMH interface is used to plot a cumulative graph of records over time. The date of the earliest OAI-PMH record is used to show the number of cumulative archives over time.

http://archives.eprints.org/eprints.php
Declaration of Institutional Commitment to implementing the Berlin Declaration on open-access provision

Our institution hereby commits itself to adopting and implementing an official institutional policy of providing open access to our own peer-reviewed research output -- i.e., toll-free, full-text online access, for all would-be users webwide -- in accordance with the Budapest Open Access Initiative and the Berlin Declaration.

**UNIFIED OPEN-ACCESS PROVISION POLICY:**

(OAJ) Researchers publish their research in an open-access journal if a suitable one exists otherwise

(OAA) Researchers publish their research in a suitable toll-access journal and also self-archive it in their own research institution's open-access research archive.

To sign: [http://www.eprints.org/signup/sign.php](http://www.eprints.org/signup/sign.php)

A JISC survey (Swan & Brown 2004) "asked authors to say how they would feel if their employer or funding body required them to deposit copies of their published articles in one or more… repositories. The vast majority... said they would do so willingly."

Gold (OA), Green (S-A) & Gray Journals

- **GRAY**: No green light yet
- **PALE-GREEN**: Green Light for Author Preprint Self-Archiving (S-A)
- **GREEN (spotted)**: Green Light for Author Postprint Self-Archiving (S-A)
- **GREEN (solid)**: Green Light for Author Postprint and Preprint Self-Archiving (S-A)
- **GOLD**: Open Access Journals (OAJ) http://www.doaj.org/

Journals (percent) Journals (total number: 11148) (adding in all OAJs: total number: 11623)
Where the work needs to be done to accelerate growth per Archive:

*These curves must become convex upward:*

*Institutional self-archiving policies are needed*
Even the fastest-growing archive, the Physics ArXiv, is still only growing linearly (since 1991):

At that rate, it would still take a decade before we reach the first year that all physics papers for *that year* are openly accessible (Ebs Hilf estimates 2050!)
Four reasons for research impact
(shared by researcher and institution but not by researcher and discipline)

1. Contributions to Knowledge
2. Employment, Salary, Promotion, Tenure, Prizes
3. Research Funding, Resourcing
4. Institutional Overheads, Prestige (attracting teachers, students, researchers, industrial collaboration)
Don’t conflate the different forms of institutional archiving:

Only the 5th is relevant here

1. Institutional digital collection management
2. Institutional digital preservation
3. Institutional digital courseware
4. Institutional digital publishing
5. Institutional self-archiving of refereed research output
Would-be peer review reformers, please remember:

• The pressing problem is to free peer-reviewed research access and impact from **tolls**: not from **peer review**!

• If you have a peer-review reform hypothesis, please take it elsewhere, and test it, and then let us all know how it comes out…

• Meanwhile, please let us free peer-reviewed research such as it is!
Universal Access Through Affordable Licensing?

Open access through author/institution self-archiving is a parallel self-help measure for researchers, to prevent further impact-loss now. Open access is a supplement to toll-access, but not necessarily a substitute for it.

One possible outcome is that the toll access and open access versions will peacefully co-exist in perpetuity, with all researchers using the toll-access versions of the research their own institutions can afford and the open-access versions of the rest. The more affordable the toll-access licenses, the less researchers will need to use the open-access versions.

Even if the growth of the open-access versions is destined eventually to reduce the demand for the toll-access versions, that is a long way off, because self-archiving proceeds gradually and anarchically, and journals cannot be cancelled while only random parts of their contents are openly accessible.

If and when open accessibility does reduce the demand for the toll-access versions, this will at the same time be creating windfall savings for institutions on their periodical budgets -- savings which will then be available to institutions to pay for peer-review service provision up-front to those journals that are ready to convert to becoming open-access journals.
Proportion of Toll-Access and Open-Access Journals Today

- Toll-Access Journals (Ulrichs)
- Open-Access Journals (DOAJ)

1000
23,500
Proportion of Toll-Access and Open-Access Articles Today

- Total Toll Access Articles: 2,000,000
- Open Access Journal Articles (estimate): 100,000
- Oaster (estimate): 250,000
- Citeseer (estimate): 150,000
Romeo "Green/Gold" versus "Gray" Journals

- Journals already supporting self-archiving (gold)
- Journals already supporting self-archiving (green)
- Journals not yet supporting self-archiving (gray)

- 78%
- 17%
- 5%
The 100% Solution for providing immediate Open Access Today

- Open Access Journals
- Self-Archiving
39% of authors self-archive
69% would self-archive willingly if required

Actual and potential proportions of Open Access Articles

Authors already self-archiving at least one TA article 39%
Authors already publishing at least one OA Journal 4%
Authors unwilling to provide OA even if required 3%
Authors who would self-archive if required 26%
Authors who would self-archive willingly if required 28%
LIMITED ACCESS

LIMITED IMPACT
What-is/why/how FAQs:
  - What is self-archiving?
  - What is the Open Archives Initiative (OAI)?
  - What is OAI-compliance?
  - What is an Eprint Archive?
  - How can I or my institution create an Eprint Archive?
  - How can an institution facilitate the filling of its Eprint Archives?
  - What is the purpose of self-archiving?
  - What is the difference between distributed and central self-archiving?
  - What is the difference between institutional and central Eprint Archives?
  - Who should self-archive?
  - What is an Eprint?
  - Why should one self-archive?
  - What should be self-archived?
  - Is self-archiving publication?
  - What about copyright?
  - What if my copyright transfer agreement explicitly forbids self-archiving?
  - Peer-review reform: Why bother with peer review?
  - Is self-archiving legal?
  - What if the publisher forbids preprint self-archiving?

What-to-do FAQs:
  - What can researcher/authors do to facilitate self-archiving?
  - What can researchers' institutions do to facilitate self-archiving?
  - What can libraries do to facilitate self-archiving?
  - What can research funders do to facilitate self-archiving?
  - What can publishers do to facilitate self-archiving?

"I-worry-about..." FAQs

1. Preservation
2. Authentication
3. Corruption
4. Navigation (info-glut)
5. Certification
6. Evaluation
7. Peer review
8. Paying the piper
9. Downsizing
10. Copyright
11. Plagiarism
12. Priority
13. Censorship
14. Capitalism
15. Readability
16. Graphics
17. Publishers' future
18. Libraries'/Librarians' future
19. Learned Societies' future
20. University conspiracy
21. Serendipity
22. Tenure/Promotion
23. Version control
24. Napster
25. Mark-up
26. Classification


Harnad, S. (2001) For Whom the Gate Tolls? How and Why to Free the Refereed Research Literature Online Through Author/Institution Self-Archiving, Now. [http://cogprints.soton.ac.uk/documents/disk0/00/00/16/39/]

