University of Southampton

Open Access: The Web and Scientific Knowledge

COMP3227 Web Architecture & Hypertext Technologies

Dr Heather Packer – hp3@ecs.soton.ac.uk

The Original Web Challenge?

"In those days, there was different information on different computers, but you had to log on to different computers to get at it. Also, sometimes you had to learn a different program on each computer. Often it was just easier to go and ask people when they were having coffee…"

Sir Tim Burners-Lee
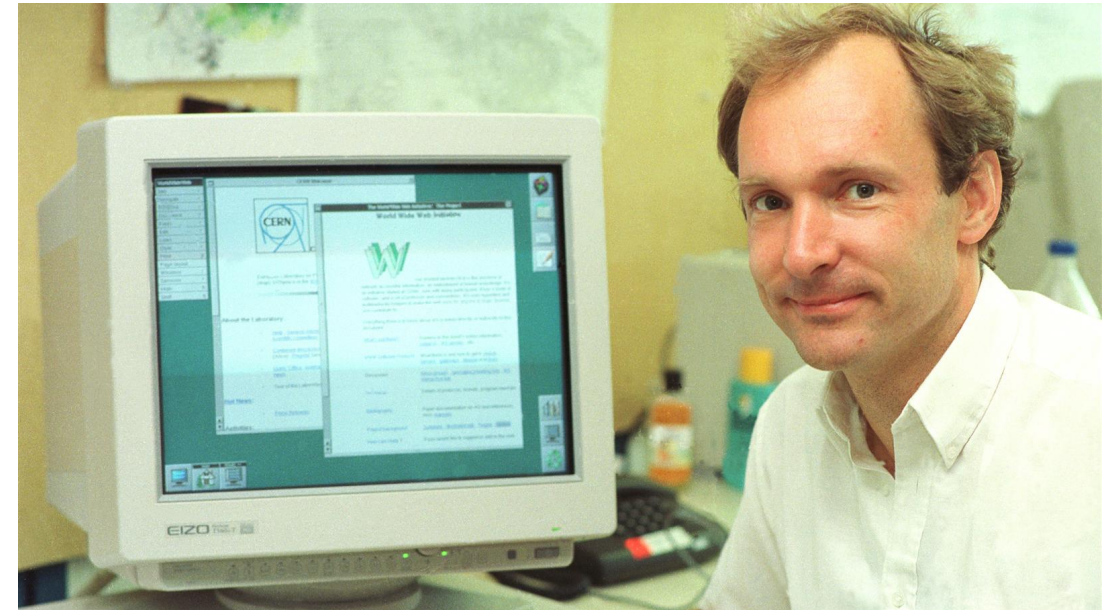
WORLD WIDE WEB

The WorldWideWeb (W3) is a wide-area hypermedia[1] information retrieval initiative aiming to give universal access to a large universe of documents.

Everything there is online about W3 is linked directly or indirectly to this document, including an executive summary[2] of the project, Mailing lists[3] , Policy[4] , November's W3 news[5] , Frequently Asked Questions[6] .

What's out there?[7]Pointers to the world's online information, subjects[8] , W3 servers[9], etc.

Help[10]   on the browser you are using

Software   A list of W3 project components and their current
Products[11]   state. (e.g. Line Mode[12] ,X11 Viola[13] ,
                      NeXTStep[14] , Servers[15] , Tools[16] , Mail
                      robot[17] , Library[18])

Technical[19]   Details of protocols, formats, program internals
                      etc

<ref.number>, Back, <RETURN> for more, or Help:

Open Access: Tim's Problem

- Publishers owned the academic publications
- Authors writes article
  - Author signs over copyright to publishing company
  - Publishing company publishes article
  - Publishing company payment from university libraries
- Author receives academic credit
  - Publication glory
  - Citations
- Universities created a whole new industry and students read and write paid for by external company
- Universities and researchers are knowledge producers and knowledge consumers

researchers   publishers

read

write
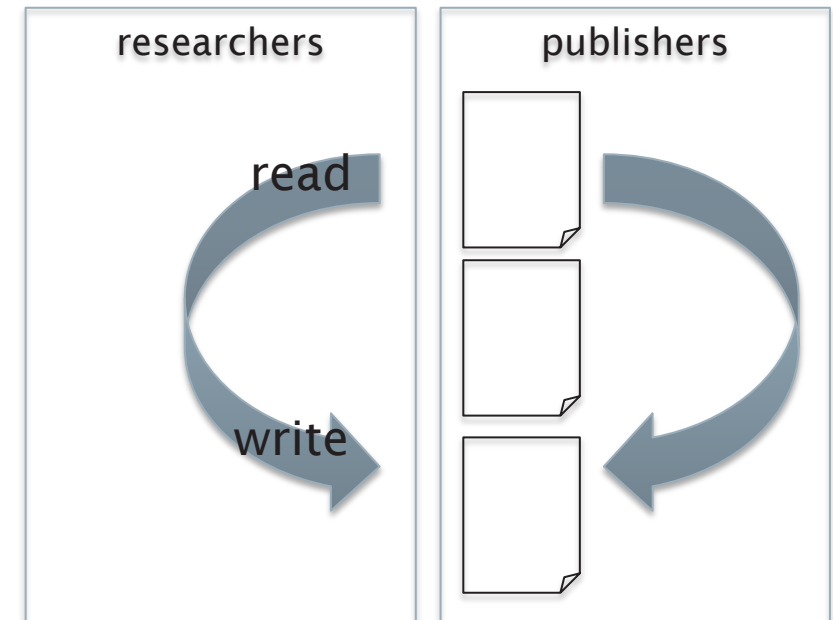
Scholarly Publishing 1665-1960

- Scientific and scholarly societies publish their own journals

- First scientific journal March 1665

   The Royal Society publishes "Philosophical Transactions"

- University Presses not for profit but were for prestige

- For the benefit of their members, for the benefit of science
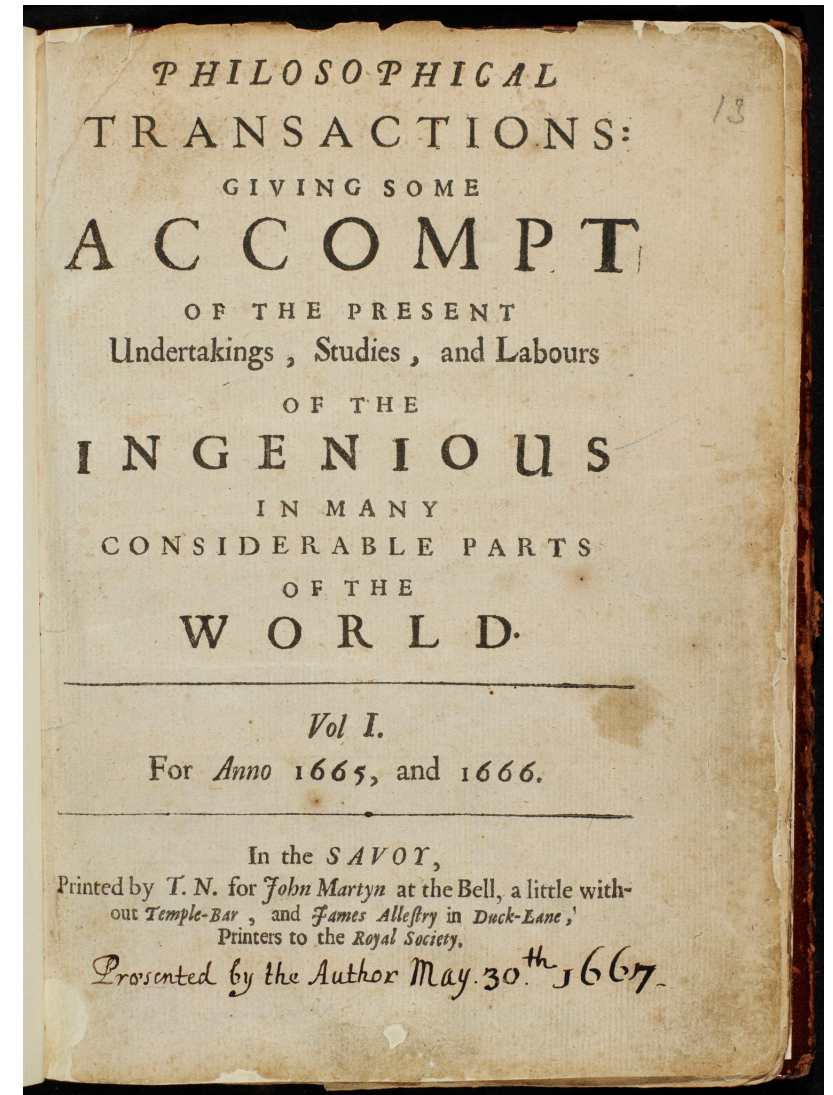
Private Sector

- After the war 1948 **Robert Maxwell** decided to publish scientific journals and set up **Pergamon Press** which was quickly and hugely profitable

  - Bringing "aggressive publishing in science"

- In 1960's Universities were expanding

- UK Robbins Report/expansion of higher education & science budget

- In 1968/9 STM Association founded for publishers

- The **new demand** made for a very profitable system – with an increasing number of commercial publishers moving into STM

STM - The International Association of **S**cientific, **T**echnical and **M**edical Publishers

The Serial Crisis

- Subscription prices continued to rise

- Some research fields were more affected than others

- Association of Research Libraries investigated budgets

- In 2002 the Modern Language Association hoped that electronic publishing would solve the crisis

- 2009-2010 saw 70% budget cuts

1986

Books
44%

Subscrip…
56%

44

56

1998

Books
28%

Subscrip…
72%

28

72
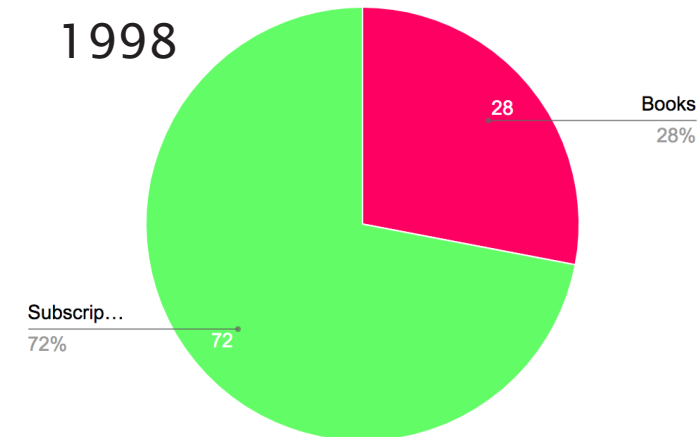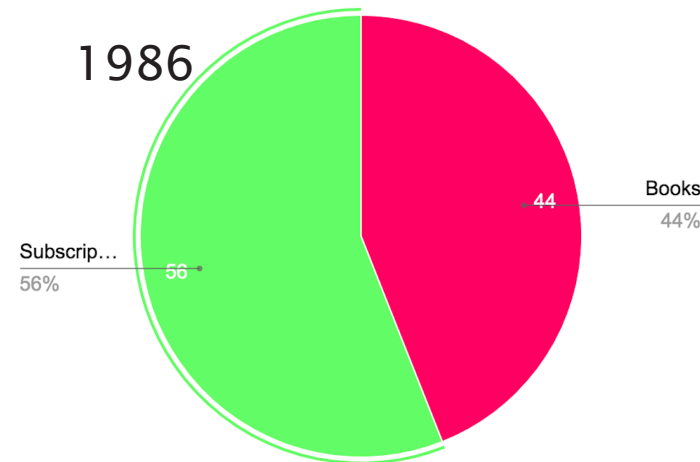
Publications

Year	Not OA	OA
2009	807,672	361,932
2010	824,427	389,100
2011	876,706	445,341
2012	917,504	503,875
2013	971,428	569,939
2014	1,010,275	625,503
2015	1,017,187	677,841
2016	1,044,540	731,812
2017	1,086,564	736,656
2018	1,208,464	684,636

Data, gathered through the analysis of Scopus data and Unpaywall data

The Problem with Paywalls

- Restricts access to research papers

- Prevents research being used

- Prevents research from being built upon

The Twin Peaks Problem

- ≈24,000 journals with 2,500,000 article/yr

- Pay wall for access

- Limited access based on universities budgets
  - Well funded universities can only afford a fraction of subscriptions
  - Less well funded universities can afford even fewer subscriptions

- ≈£10,000 per year for access to some journals, and prices increasing

- Universities are measured by their impact

- Fewer readers means less impact (fewer citations)

The Literature: As We Imagine

- Integrated
  - Citations
- Readily available
- Standing on the shoulders of giants

  "we see more and farther than our predecessors, not because we have keener vision or greater height, but because we are lifted up and borne aloft on their gigantic stature"
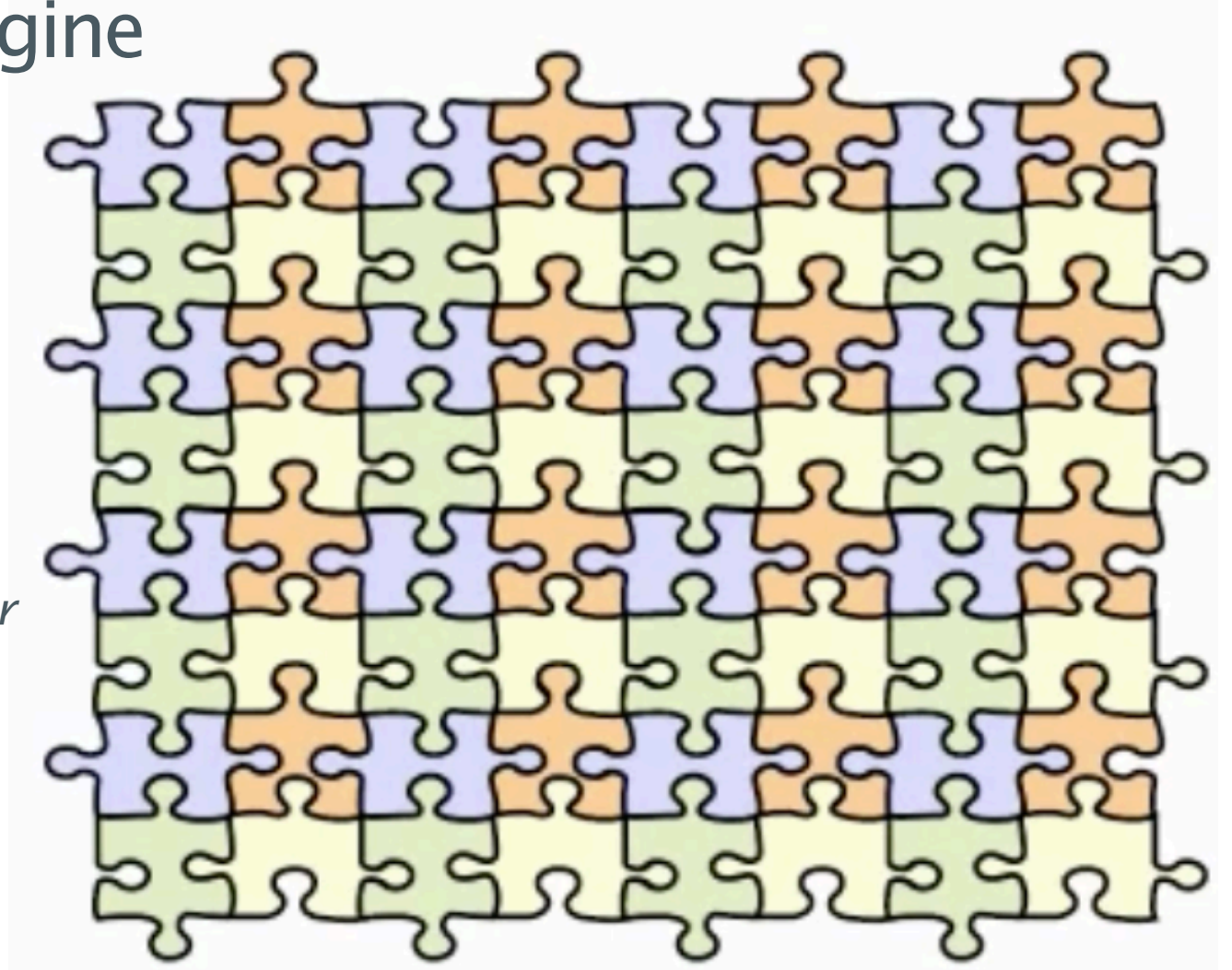
The Literature: As It Is

- Inaccessible
  - Hidden behind pay walls
  - Not subscribed too
- Disjoint
  - Books
  - Online journals
  - Scientific magazines
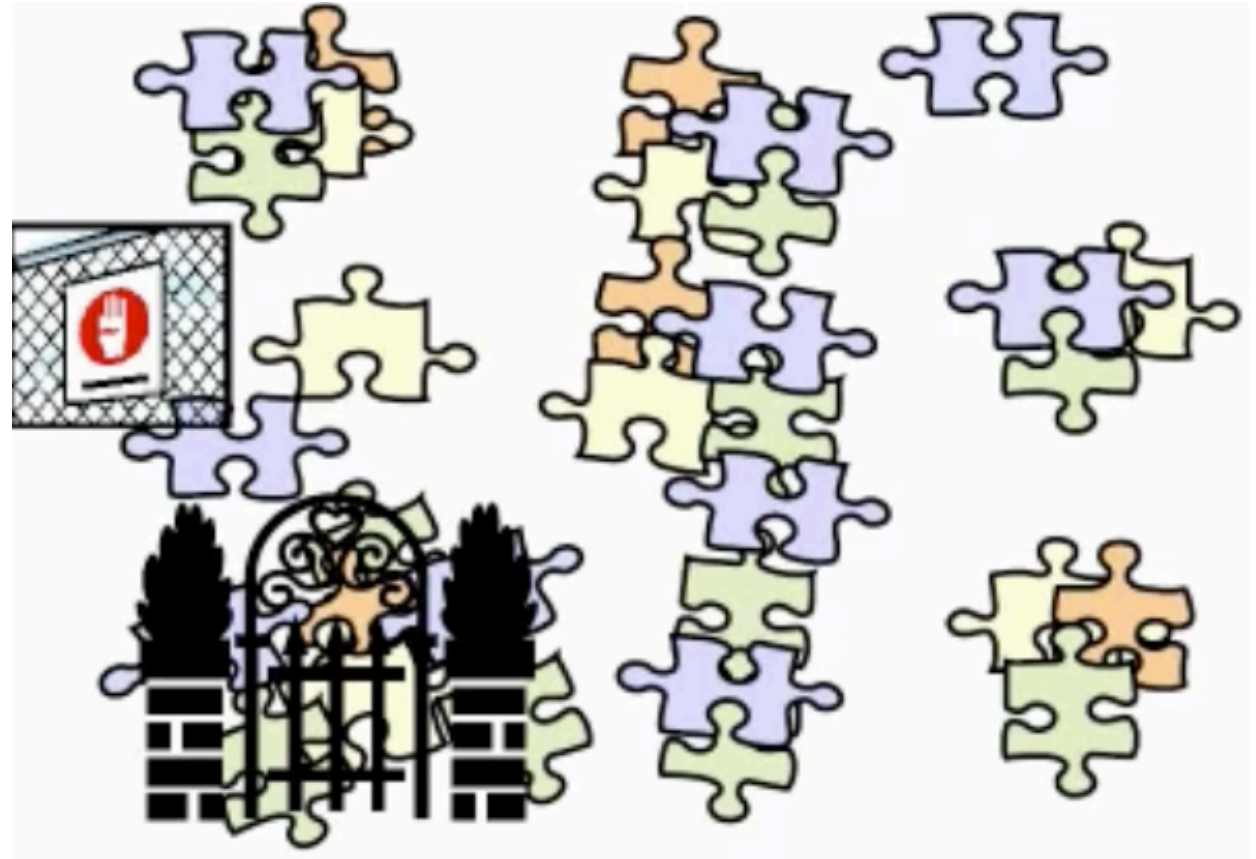
Open Access

Open Access (OA)

- Open Access to Scientific and Scholarly Research Literature

- OA is movement for organising and disseminating the worlds research knowledge through Web technology

    – Free of cost

    – Other access barriers

- Focuses on Peer Reviewed Literature

Open access logo, designed by Public Library of Science

The Budapest Open Access Initiative (BOAI)

- The BOAI publishes a set of principles for open access and research literature

- Released February 14th, 2002

- It arose from a conference hosted in Budapest by the Open Society Institute in 2001

  - This small gathering of individuals is recognised as one of the major defining events of the open access movement.

    Old tradition of scholarly publishing + New technology of the Internet =

    Public good: free and unrestrictive access to peer-reviewed journal literature

Open Access Definition

*"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. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited."*

- The Budapest Open Access Initiative

Full Open Access Ideal

- The entire full-text refereed corpus online

- On every researcher's desktop, everywhere

- 24 hours a day

- All papers citation-interlinked

- Fully searchable, navigable, retrievable

Researchers in developing countries can see your work

More exposure for your work

Practitioners can apply your findings

Higher citation rates

Taxpayers get value for money

Compliant with grant rules

The public can access your findings

Your research can Influence policy

Traditional Subscription Publishing

- Limited Dissemination, economic efficiency & social impact
- Paywall
- Libraries purchase subscriptions
- Very limited permissions
- Slow process

Publicly funded researchers conduct research and write up results.

Manuscripts submitted to subscription journals & reviewed by peers.

Manuscripts accepted for publication.

Authors transfer copyright to publishers. No rights retained by authors.

Slow scientific progress, poor return on public investment.

Even after paying for access, readers are granted little or no reuse rights beyond permissions to read.

Libraries purchase subscription or public pays per article to view on publisher's website.

Published articles are locked behind paywalls.

Limited Access: Limited Research Impact

- Research is a process
- Can take more than a year and a half for a journal
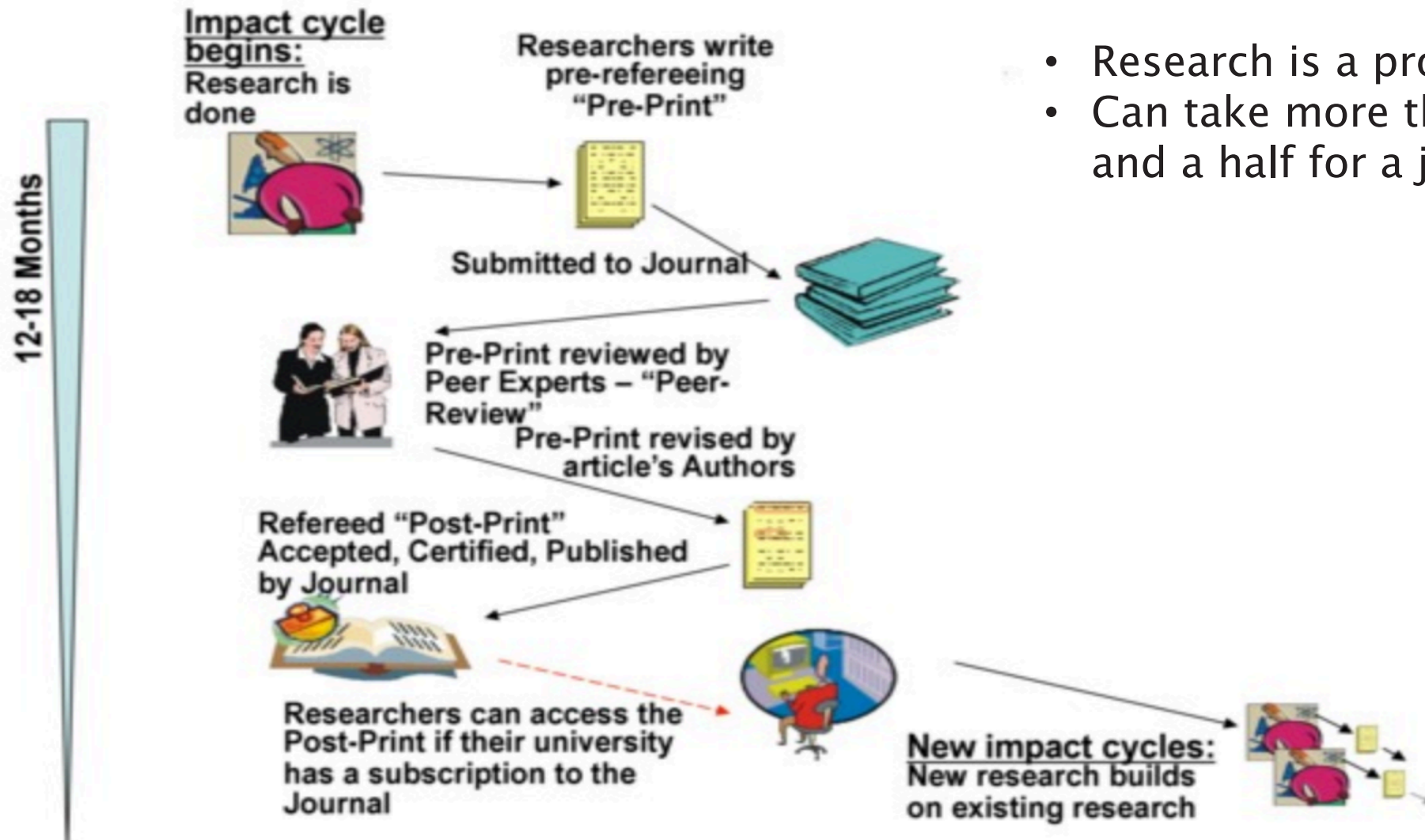
Gold Open Access

- Maximised Dissemination, economic efficiency & social impact
- Articles freely, permanently and immediately accessible by everyone
- Copyright retained by authors
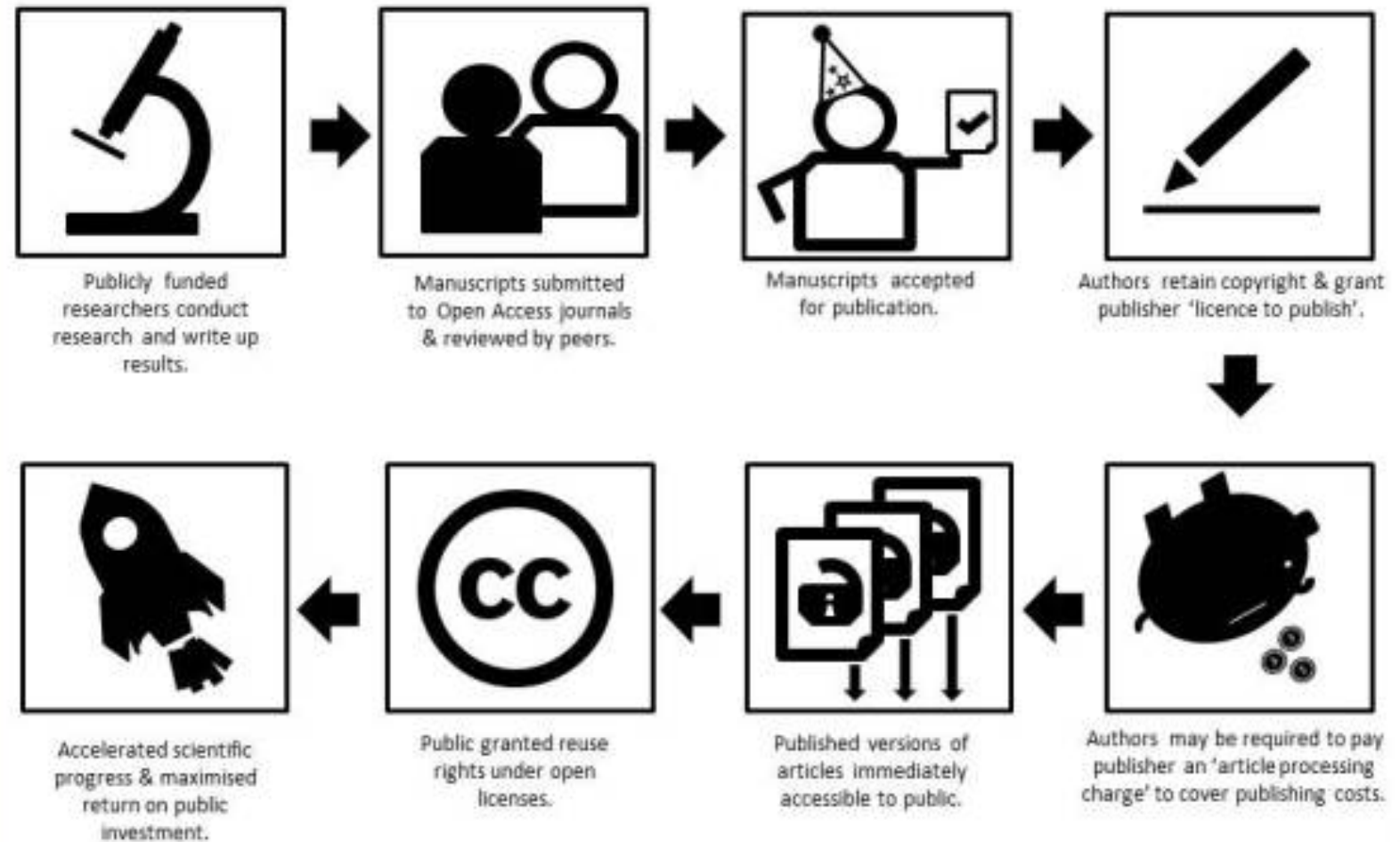- Very few restrictions on use

Publicly funded researchers conduct research and write up results.

Manuscripts submitted to Open Access journals & reviewed by peers.

Manuscripts accepted for publication.

Authors retain copyright & grant publisher 'licence to publish'.

Authors may be required to pay publisher an 'article processing charge' to cover publishing costs.

Published versions of articles immediately accessible to public.

Public granted reuse rights under open licenses.

Accelerated scientific progress & maximised return on public investment.

Green Open Access

- Increased Dissemination, economic efficiency & social impact
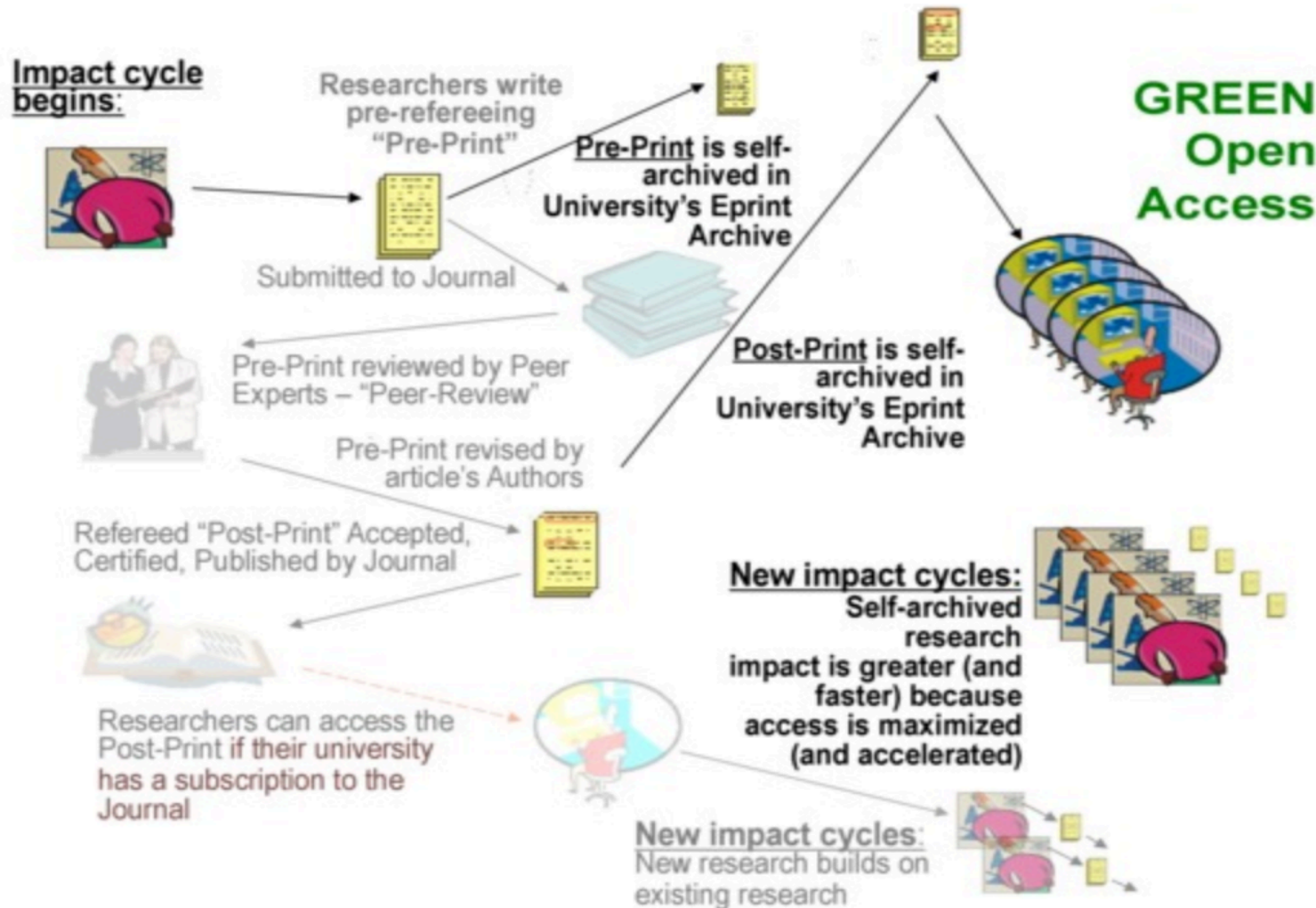- Self Archiving
- A version is freely available
- Publish pre or post publication

Publicly funded researchers conduct research and write up results.

Manuscripts submitted to subscription journals & reviewed by peers.

Manuscripts accepted for publication.

Authors assign copyright to publishers, but retain the right to disseminate an OA copy (of the accepted manuscript) via open access repositories.

Published versions of articles accessible to journal subscribers immediately.

Accepted manuscript versions may be embargoed temporarily.

After embargo period, public can download the open access copies from repositories.

Accelerated scientific progress & increased return on public investment.
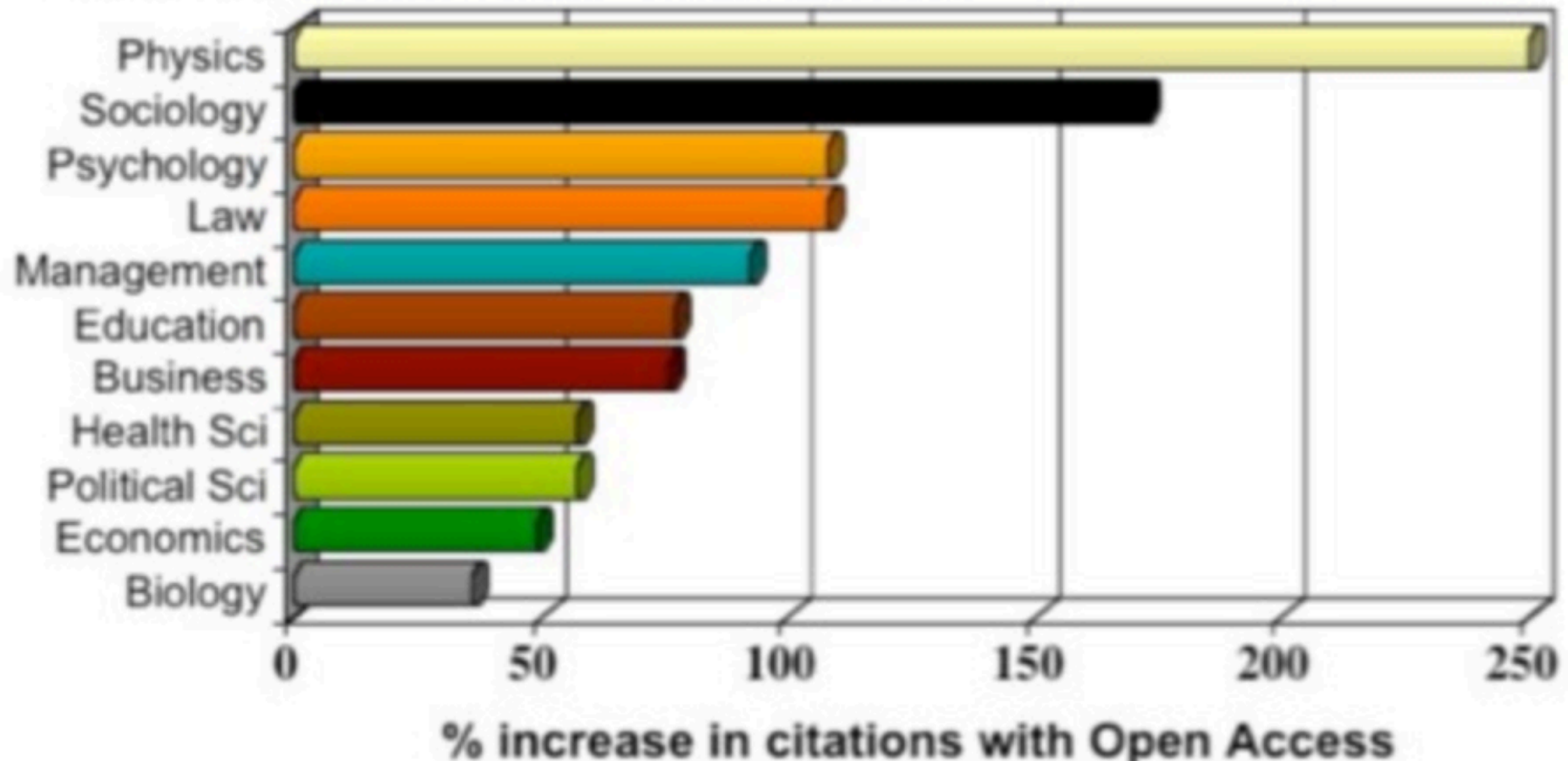
Green Open Access Pre-Prints

- Preprints allow people to view work immediately
- However, impact is greater after peer reviewed publications

Green OA Advantages

- **Early Advantage**: Self archiving **preprints** before publications hastens and **increases usage and citations**

- **Quality Advantage**: Self-archiving **postprints** immediately upon publication hastens and **increases usage and citation**

  – Higher quality articles benefit more

- **Usage Advantage:** Self-archiving both **pre/post prints** increases downloads

  – Higher quality articles benefits more!

Full bibilograhy, see http://opcit.eprints.org/oacitation-biblio.html

Open Access

% increase in citations with Open Access

Publications

University of Southampton

Year	Not OA (%)	OA (%)	Gold OA(%)	Green OA(%)
2009	69.1	30.9	4.69	20.44
2010	67.9	32.1	5.98	21.8
2011	66.3	33.7	7.7	22.92
2012	64.6	35.4	9.86	24.24
2013	63.0	37.0	11.37	25.64
2014	61.8	38.2	13.02	27.15
2015	60.0	40.0	14.32	28.22
2016	58.8	41.2	15.16	28.99
2017	59.6	40.4	16.98	26.91
2018	63.8	36.2	18.47	15.29

Data, gathered through the analysis of Scopus data and Unpaywall data

Open Access Strategies

Green: Self Archiving	Gold: Publishing
Journal processes continue as normal	Journal changes business model
Authors deposit a copy of their papers into an **'open access repository'**	Readers no longer pay to read
Public copy is a supplement to the publishers official article for those who can't afford a subscription	Instead, authors pay to publish or their funders

Problems with OA

- 20% of new papers are Open Access 2019

- 50% of new papers, all versions, are behind a paywall

- Green OA relies on publishers polices

- Romeo service tracks publisher policies

  - http://www.sherpa.ac.uk/romeo

- Gold OA relies on publishers changing their model

- Scientific publications is lucrative (18% profits)

- Academia highly values the peer review process

Open Access Repositories

- Open-access repository or open archive

- Free, immediate and permanent access for anyone to use, download and distribute

- Institutional or disciplinary

- They provide free access to research for users outside the institutional community

Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)

- A protocol developed for harvesting metadata descriptions of records in an archive so that services can be built using metadata from many archives

- Repositories must be interoperable according to OAI-PMH

- Search engines harvest the content of open access repositories, constructing a database of worldwide, free of charge available research

Role of the Repository

- Who takes responsibility for curating the research knowledge of the world?
    - Privately owned, profit making, publishing industry
    - University, research institution, knowledge creator

Next: Telling Tales part 2