Open Access

Lecture 5

TU Wien, 193.067 Free and Open Technologies (WS 2019/2020)
Christoph Derndorfer and Lukas F. Lang

Organization

- First project meeting: **Friday, November 8**
  - max. 45 min., discussion of your project idea, scope, milestones, etc.
- Location: Projektraum (Argentinierstr. 8/2. Stock/Mitte)
- Time slots:
  - 1PM: Open map based on pictures
  - 2PM: Open Data for Academicians at TU Wien
  - 3PM: VoWikiathon
  - 4PM: Smart Plant Pot

Lecture outline

1. ~~October 8, 2019: FLOSS (Free/Libre and Open Source Software)~~
2. ~~October 15, 2019: Open Hardware~~
3. ~~October 22, 2019: Open Data~~
4. ~~October 29, 2019: Open Content/Open Educational Resources~~
5. **November 5, 2019: Open Access**
6. November 12, 2019: Open Science/Research
7. November 19, 2019: Open Spaces/Open Practices: Metalab Vienna
   - Location: Metalab, **Rathausstraße 6, 1010 Vienna**
8. November 26, 2019: Guest Lecture: Stefanie Wuschitz (Mz* Baltazar's Lab)
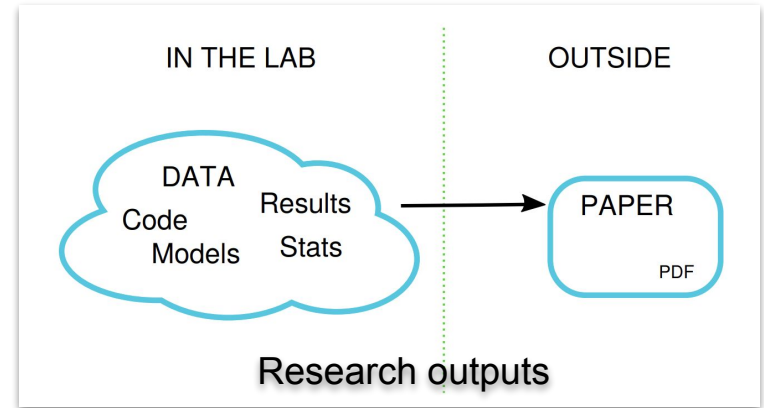
Open content: recap

In a nutshell [1]:

*Creative work that others can **use**, **copy**, **modify**, **distribute** freely, without asking for permission.*

What about research outputs?

- Published peer-reviewed articles ← **Open access**
- Code, data, results, tables, protocols
- Etc.

Example:

- PLOS (Public Library of Science) publishes ~50.000 scientific articles/year under CC license

IN THE LAB   OUTSIDE

DATA
Code   Results
Models   Stats   → PAPER
                              PDF

Research outputs

[1] David Wiley, "Defining the "Open" in Open Content and Open Educational Resources", CC BY, accessed 2019/10/24
[2] Image by Stephen J. Eglen, CC BY

Scientific publishing business

57 major publishing groups received a revenue of €60bn in 2015 [1]:

- Elsevier:
  - £2bn revenue, £724m profit in 2010 [2] → **36% profit margin!**
  - £900m profit in 2018 [3]

- Springer Nature:
  - €1.64bn revenue, €374m profit in 2017 [4] → **23% profit margin!**

- Wiley:
  - US$1.8bn revenue, US$224m profit in 2019 [5]
  - Number of employees: 5,100 (according to WP) → **US$44k profit per employee!**

[1] Wischenbart, R. (2015). "The Global Ranking of the Publishing Industry 2015", accessed 2019/10/30
[2] The Guardian (2019/03/04), "The Guardian view on academic publishing: disastrous capitalism", accessed 2019/10/30
[3] The Guardian (2017/06/27), "Is the staggeringly profitable business of scientific publishing bad for science?", accessed 2019/10/30
[4] Handelsblatt (2018/04/18), "Science publisher hopes IPO will raise €1.2 billion", accessed 2019/10/30
[5] The Bookseller (2019/06/11), "Wiley sees flat revenue, profit dip in 2019 fiscal year", accessed 2019/10/30

How do scientific publishers make money?

Quick BOTE calculation:

390 * €41.94 = **€16.4k**

Quick BOTE calculation:

12k * €41.94 = **€500k**

By selling **peer-reviewed** scientific articles (to individuals, researchers, etc.)

Mainly through **subscriptions** (university libraries, companies, etc.)

→ Using university access (WIFI, VPN) you might not have noticed these paywalls!

[1] Screenshots from https://doi.org/10.1007/s10851-014-0513-4 and https://doi.org/10.1007/s10851-010-0251-1

What's the creators' share?

0
(in words: zero)

Why → transfer of copyrights

But: many (non-financial) incentives
- Career reasons (PhDs, profs,...)
    "Publish or perish"
- Peer review
- Prestige, etc.

Example transfer of copyrights (image removed)

How much do journal subscriptions cost?

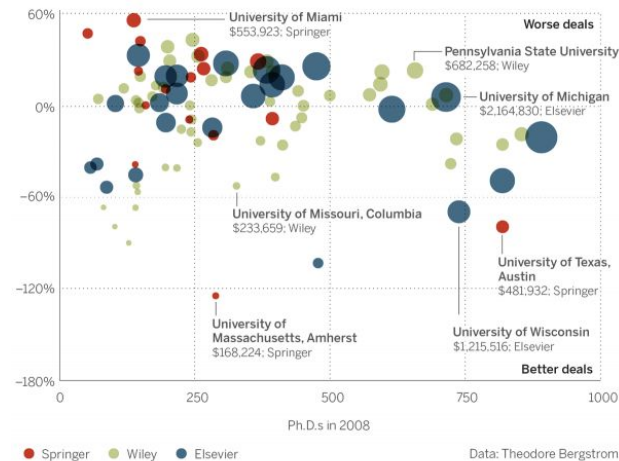Typically, subscription prices of journal **bundles** are negotiated **in secret**

Freedom of Information (FOI) requests revealed:

- In 2014, UK universities spent **£430m** for journal access [1]:

University of Manchester	£15m
UCL	£14m
University of Cambridge	£14m
University of Nottingham	£12m
University of Bristol	£11m

- Massive **discrepancies** in prices paid for the same bundles [2, 3]:

Journal prices: good deal or not?
Payment premium (% > 0) or discount, based on expected price from modeling; circle size reflects university's 2009 payment

University of Miami
$553,923; Springer

Pennsylvania State University
$682,258; Wiley

University of Michigan
$2,164,830; Elsevier

University of Missouri, Columbia
$233,659; Wiley

University of Texas, Austin
$481,932; Springer

University of Massachusetts, Amherst
$168,224; Springer

University of Wisconsin
$1,215,516; Elsevier

Worse deals
Better deals
Ph.D.s in 2008
Springer   Wiley   Elsevier
Data: Theodore Bergstrom

[1] Lawson, S. and Meghreblian, B. (2014), "Journal subscription expenditure of UK higher education institutions"
[2] Bergstrom, T. C. et al. (2014), "Evaluating big deal journal bundles", PNAS.
[3] ZME Science (2014/07/17), "How much is your university paying for journal access?", accessed 2019/11/02

How does academic publishing work?
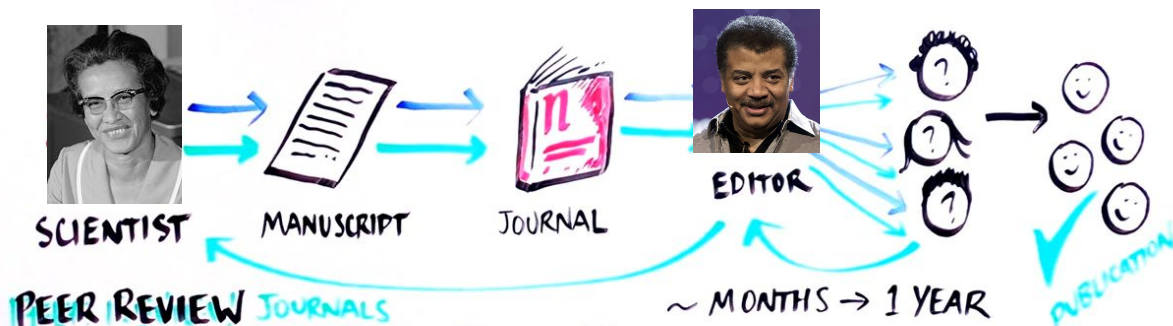
Optical Flow on Evolving Surfaces with Space and Time Regularisation

Clemens Kirisits · Lukas F. Lang · Otmar Scherzer

Published in print: **May 2015**

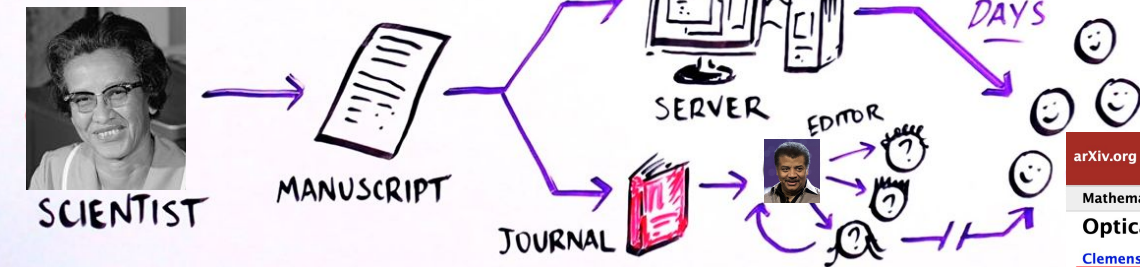arXiv.org > math > arXiv:1310.0322

Mathematics > Optimization and Control

Optical Flow on Evolving Surfaces with Space and Time Regularisation

Clemens Kirisits, Lukas F. Lang, Otmar Scherzer

(Submitted on 1 Oct 2013 (v1), last revised 25 Jun 2014 (this version, v2))

[1] YourekaScience, "What are preprints?", CC BY
[2] Katherine Johnson, public domain [3] Neil deGrasse Tyson, by Norwegian University of Science and Technology, CC BY-SA 2.0

Institutionalization

- Government
- Funding bodies
- Grant agencies

Universities

funding/grants
$

publishing
$

General public

rea____hip

Journals/publisher

rea____hip

salary
$

Scientists
(profs, postdocs, PhDs)

- copyright
- reviewing
- editing

- fame
- prestige
- citations
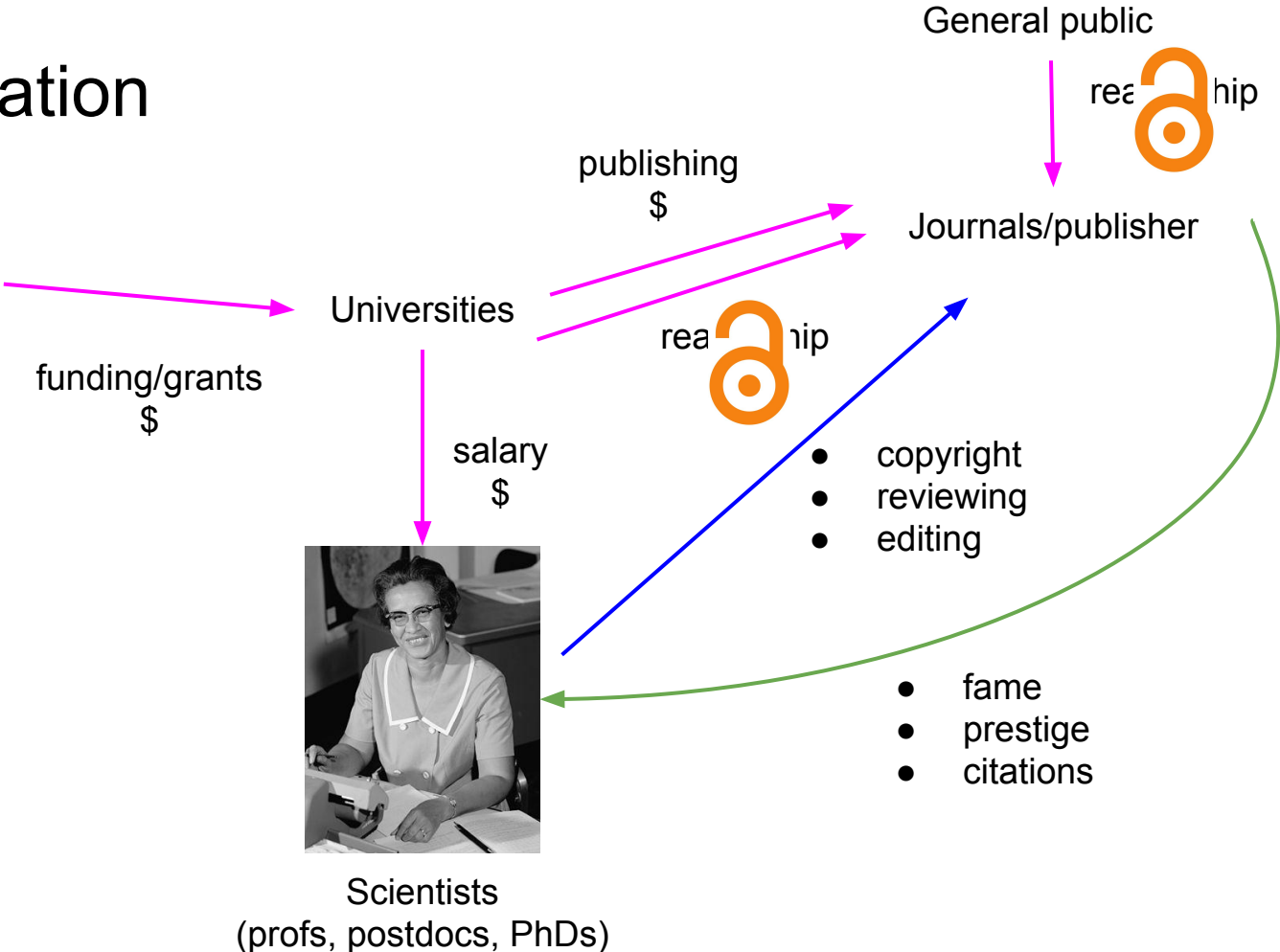
Motivation

- Massive increase in journal subscription costs [1]
  - Referred to as "serials crisis"

- Digitalization & internet
  - Scientific publishing affected by wide adoption of the internet (distribution, speed, costs, ...)
  - Added value by publishers increasingly negligible/irrelevant, outdated processes

- Increased pressure by public and scientific community
  - Most research is funded by taxpayer
  - Publications should be available to everyone free to read

- Funding bodies, institutions, and libraries
  - Pushed to change cost models
  - Introduced open access mandates
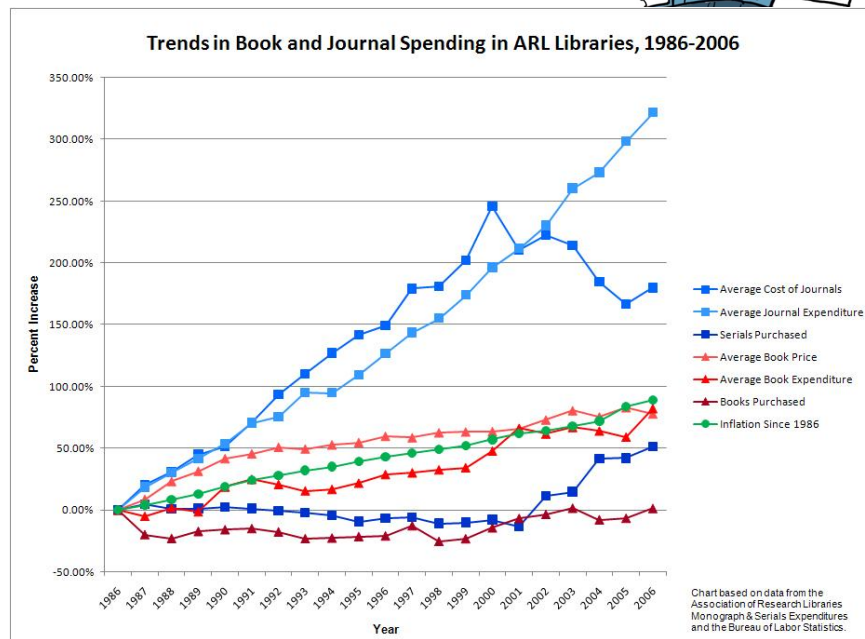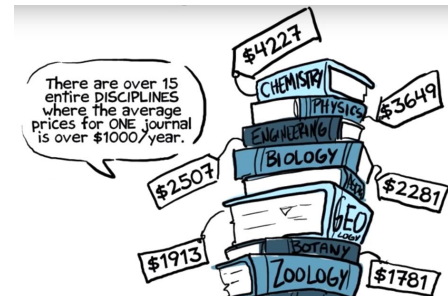
[1] Western Illinois University, "Open Access and Scholarly Publishing", accessed 2019/11/04
[2] Piled Higher and Deeper (PHD Comics), CC BY

What is open access (OA)?

OA to scientific literature means [1, 2]:

- Free availability on the public internet

- Permits users to:
    - **read**, **download**, **copy**, **distribute**, **print**, **search**, or **link full texts** of articles
    - **crawl** for indexing, **pass on** as data to software, or **use** for any other lawful purpose

- Without financial or technical barriers other than access to the internet

- Only constraint on reproduction, distribution, and copyright:
    - should be to give authors control over the integrity of their work
    - the right to be properly acknowledged and cited

Note: different definition of "openness" → gratis vs. libre OA

[1] Peter Suber (2004/06/21), "Open Access Overview", accessed 2019/10/30
[2] Budapest Open Access Initiative (2002, February 14), "Read the Budapest Open Access Initiative", accessed 2019/10/11

Benefits

More exposure for your work

Practitioners can apply your findings

Higher citation rates

Your research can Influence policy

The public can access your findings

Compliant with grant rules

Taxpayers get value for money

Researchers in developing countries can see your work

[1] Jisc, "An introduction to open access", accessed 2019/11/03

History

1991: [arXiv](#) (preprint repository) by Paul Ginsparg (Los Alamos NL)

2000s:
- 2000: Public Library of Science (PLOS) founded
- 2001: Creative Commons (CC) founded
- 2002: Budapest Open Access Initiative ([BOAI](#))
  - First and most widely used definition of OA
  - Advised strategy: self-archiving & open-access journals
- 2003: Bethesda Statement of Open Access Publishing
- 2003: Berlin Declaration on Open Access to Knowledge
- 2004: Launch of Google Scholar
- 2005 onwards: Wellcome Trust, National Institutes of Health (NIH), etc. adopt OA policies

2010s: public initiatives and boycotts
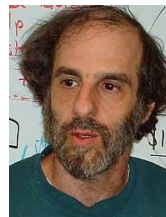
14k new submissions each month

Against the status quo: Aaron Swartz

- Downloading articles from JSTOR:
  - Aaron Swartz (involved in Creative Commons, Demand Progress, Markdown, Reddit, RSS, etc.) started downloading almost 5 million articles from JSTOR (a largely subscription based digital library for academic journals) between September 2010 and January 2011
  - MIT's network was used for these bulk downloads and MIT's investigation determined that Swartz was responsible for it in January 2011
  - JSTOR and Swartz reached a settlement that the copies would be deleted and had not been distributed
  - Federal government continued with its hacking prosecution with potential fines of up to $1m and 35 years of imprisonment
  - Swartz committed suicide in January 2013 before a trial could take place
  - His motivation for downloading these materials is not known (but there is lots of speculation)

Against the status quo: Timothy Gowers

- The Cost of Knowledge boycott against Elsevier (2012):
  - Mathematician and Fields Medal winner Timothy Gowers (University of Cambridge) called for a boycott of Elsevier [1]
  - Cites four main reasons: high prices, bundling, bullying practices, support for restrictive laws: Research Works Act / SOPA / PIPA
  - Gathered significant support from many mathematicians within a month (>18,000 to date) [2]
  - Elsevier then announced a stop of its support for the Research Works Act (which was announced not to be pursued further on the same day) and other measures (price reduction, opening up some publications' archives, etc.)
  - Subsequent research indicates that the actual impact of the boycott is unclear and hard to assess [3]

[1] Gowers, T. (2012/01/21), "Elsevier — my part in its downfall", last accessed 2019/10/30
[2] The Cost of Knowledge, last accessed 2019/10/30
[3] Heyman T, Moors P and Storms G (2016) On the Cost of Knowledge: Evaluating the Boycott against Elsevier. Front. Res. Metr. Anal. 1:7. doi: 10.3389/frma.2016.00007
[4] By Martin Greuel - https://opc.mfo.de/detail?photo_id=14720, CC BY-SA 2.0 de, https://commons.wikimedia.org/w/index.php?curid=31227309

Types of open access

- Two main types of open access articles [1]:
  - Gold Open Access: publication in an open access journal
  - Green Open Access: "self-archiving" of pre- or post-prints on own website or platforms
    - arXiv.org by Cornell University offers access to more than 1.6 million publications
- Also hybrid models, e.g. non-open access journals offering authors to pay a fee to have their articles published open access
  - Hybrid models can contain additional restrictions
    - 6 or 12 months embargos
    - self-archiving, etc.
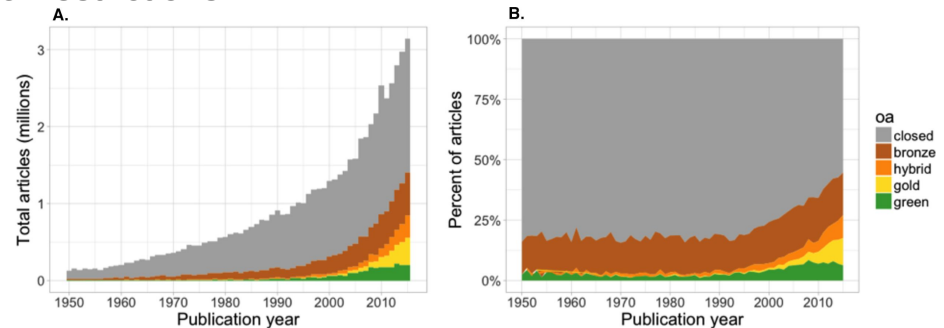
[1] Cornell University Library, "Open Access Publishing : What is Open Access?", last accessed 2019/11/03
[2] Piwowar, H. et al. The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles, PeerJ 6 (2018), e4375.

Open access and licenses

- CC BY used by almost 50% of Open Access journals in Directory of Open Access Journals
- CC BY-NC and CC BY-NC-ND each used by roughly 20% of OA journals [1]

[1] Zerkee, J. (2018/10/22), "CC-What? What's behind the Creative Commons licenses used by OA journals?", last accessed 2019/11/03

SHERPA RoMEO Project

RoMEO colour	Archiving policy	Publishers	%
green	Can archive pre-print and post-print	1064	42
blue	Can archive post-print (ie final draft post-refereeing)	844	33
yellow	Can archive pre-print (ie pre-refereeing)	182	7
white	Archiving not formally supported	472	18

Summary: **82%** of publishers on this list formally **allow** some form of self-archiving.

● Provides an overview of journals and publishers' OA policies

Journal:	**Journal of Mathematical Imaging and Vision** (ISSN: 0924-9907, ESSN: 1573-7683)
RoMEO:	This is a RoMEO **green** journal
Paid OA:	A paid open access option is **available** for this journal.
Author's Pre-print:	✔ author **can** archive pre-print (ie pre-refereeing)
Author's Post-print:	✔ author **can** archive post-print (ie final draft post-refereeing)
Publisher's Version/PDF:	✘ author **cannot** archive publisher's version/PDF
General Conditions:	- Author's pre-print on author's personal website, non-commercial pre-print server - Author's post-print on author's personal website immediately - Author's post-print on institutional repository or funder designated repository after 12 months embargo from first online publication - Publisher's version/PDF cannot be used - Published source must be acknowledged - Must link to publisher version with DOI - Post-prints are subject to Springer Nature re-use terms
Mandated OA:	(*Awaiting information*)
Paid Open Access:	Springer Open Choice
Copyright:	Self-archiving Policy
Updated:	18-Oct-2018 - **Suggest an update for this record**
Link to this page:	http://sherpa.ac.uk/romeo/issn/0924-9907/

[1] SHERPA/RomEO, "Publisher copyright policies & self-archiving", accessed 2019/11/04
[2] SHERPA/RomEO, "RoMEO Statistics", accessed 2019/11/04

OA mandates by funding bodies

An increasing number of funding bodies mandate open access/data:

- FWF (AT, €230 million for basic research in 2018), EPSRC (UK)
- EU Framework Programme "Horizon 2020" (€75 billion) defaults to an open access obligation [1]
  - Publication:
    - Researchers can choose between green OA and gold OA
    - Only applies to peer reviewed publications
    - APC costs are eligible for funding!
  - Data:
    - Requires all research data for publication to be open access (via the Open Research Data Pilot [2])
    - Requires enabling of further use of the data
    - **But:** projects may still opt out!

More info: Registry of Open Access Repository Mandates and Policies (ROARMAP)

[1] FFG, "OLD EN: Open Access in Horizon 2020", accessed 2019/11/04
[2] FFG, "OLD EN: Open Data in Horizon 2020", accessed 2019/11/04

Business models

- Traditional publishing
  - "A 2005 Deutsche Bank report referred to it as a 'bizarre' 'triple-pay' system, in which 'the state funds most research, pays the salaries of most of those checking the quality of research, and then buys most of the published product'." [1]

- Open Access publishing
  - Article or book processing charges (paid for by the authors, their institutions, or funding agencies) → an additional fourth payment to the three ones mentioned above!
    - Also found in hybrid models which tend to be more expensive than OA journals [2]
    - Discussions about the cost of OA publishing vs. traditional publishing
  - Charging for physical copies (e.g. printed versions of articles or books)
    - E.g. OpenEdition
  - Donations & grants

[1] Buranyi, S. (2017/06/27), "Is the staggeringly profitable business of scientific publishing bad for science?", last accessed 2019/10/30
[2] Pinfield, S., Salter, J. and Bath, P.A. (2016) The 'total cost of publication' in a hybrid open-access environment. Journal of the Association for Information Science and Technology, 67 (7). pp. 1751-1766. ISSN 2330-1635. https://dx.doi.org/10.1002/asi.23446

Business models

- Related services
  - E.g. OpenBookPublishers offers additional services such as multimedia book production, marketing and distribution services, etc.
  - E.g. Mendeley offers a reference management platform and related tools and operates with a freemium model (bought by Elsevier in 2013)
  - Altmetrics
    - Alternative/new indicators of impact (news mentions, blogs, Twitter, GitHub, etc.)
    - Complementary to classical citation metrics (journal impact factor, citations, h-index)

New approaches

Overlay journals

- Use preprint servers (e.g. arXiv) to deposit papers
- Organize peer review on top
  - 1990s: Paul Ginsberg invented concept [1]
  - 2016: Journal "Discrete Analysis" launched by Tim Gowers and others ("academic spring")

Open Evaluation [2]

- OpenReview.net - some prestigious CS conferences
  - Idea: improve quality & uncover bias by making them public
- Personal blogs
  - Nikolaus Kriegeskorte (Cognitive comp. neuroscientist) publishes his reviews on his blog.

[1] Brown, J. (2010), "An introduction to overlay journals". Repositories Support Project: UK
[2] Kriegeskorte, N., Walther, A., Deca, D. (2012), "An emerging consensus for open evaluation: 18 visions for the future of scientific publishing", Front. Comput. Neurosci.
[3] Screenshot from https://openreview.net/forum?id=HJenmmF8Ir

New approaches

Alexandra Elbakyan
(Kazakhstani comp. scientist)

Unpaywall

- Chrome browser plug-in finds free versions of (paywalled) scientific articles
- E.g. pre/postprint servers, institutional repositories
- Provides free access to 20 million articles [1]
- Roughly 50% of searched articles available for free [2]

Pirate sites

- Sci-Hub, Library Genesis (LibGen)
  - Started by Alexandra Elbakyan
  - Disregard copyright in favor of OA
- **Warning:** do not use

(working links can be found on Wikipedia)

[1] Else, H. (2018/08/15), "How Unpaywall is transforming open science", Nature, 560 (7718): 290–291
[2] Piwowar, H. et al. "The state of OA", PeerJ 6 (2018), e4375.
[3] Alexandra Elbakyan, by Apneet Jolly, CC BY 2.0

Predatory journals & fake conferences

Pay to publish:
- No or extremely weak peer review
- Leads to "bought" studies to back up claims about products
- Hard to spot (academic SPAM)

Investigation by NDR, WDR, and Süddeutsche Zeitung:

"...more than 5000 German researchers and 400000 researchers worldwide published in pseudo-scientific predatory journals within the last years" [1]

→ **often paid for by taxpayer!**

Revealing example:
- In 2005, MIT students created SCIgen [2] to generate nonsensical computer-science papers [3]
- Got papers accepted & presented at conferences

[1] Spiegel (2018/07/19), "Tausende deutsche Wissenschaftler veröffentlichen in Pseudo-Fachzeitschriften", accessed 2019/11/04
[2] SCIGen - An Automatic CS Paper Generator, https://pdos.csail.mit.edu/archive/scigen/
[3] MIT News (2015/04/14), "How three MIT students fooled the world of scientific journals", accessed 2019/11/04
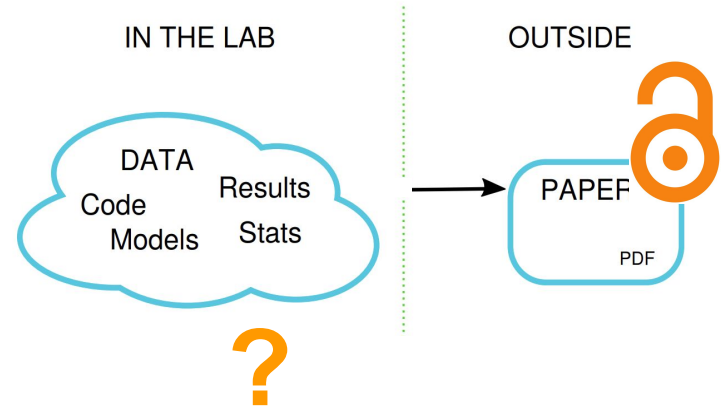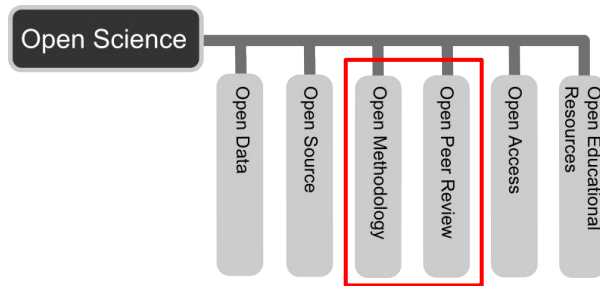
Outlook

Open access is mainly about published peer-reviewed articles

What about other research outputs?

- Data, code, results, protocols, etc.

More general topic: **open science/research**

Upcoming tasks

- Next lecture: Open Science/Research:
  - **Tuesday, November 12**, 17:00–19:00, Argentinierstraße 8, Seminarraum/Bibliothek 194-05

- First project meeting (45 min., discussion of your project idea):
  - **Friday, November 8, 13:00–17:00**, Argentinierstraße 8, **project room**
    - 1PM: Open map based on pictures
    - 2PM: Open Data for Academicians at TU Wien
    - 3PM: VoWikiathon
    - 4PM: Smart Plant Pot

- Paper group forming and topic selection:
  - **Friday, November 29**, via email to **both** lecturers

Literature & resources

Suber, P. (2012), [Open Access](#), MIT Press

Laakso, M. et al. (2011), [The Development of Open Access Journal Publishing from 1993 to 2009](#), PLoS ONE 6(6): e20961

Noorden, R. V. (2013), [Open access: The true cost of science publishing](#), Nature 495, 426–429.

Piwowar, H. et al. [The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles](#), PeerJ 6 (2018), e4375.

Other resources:

- [Open access - Tools and Resources](#)
- Piled Higher and Deeper (PHD Comics) - [Open Access Explained!](#)
- Forschergeist (2015/08/10), Podcast episode on [Open Science](#) (in German)