

# Open access monitoring in Finland

AT2OA workshop, Vienna, April 9, 2018

Jyrki Ilva <jyrki.ilva@helsinki.fi>

# Background: Open Science and Research (1)

- Open Science and Research Initiative (2014-2017)
  - A national project funded and coordinated by the Ministry of Education and Culture
  - “The objective is for Finland to become one of the leading countries in openness of science and research by the year 2017 and to ensure that the possibilities of open science will be widely utilized in our society”

The screenshot shows the homepage of the Open Science and Research initiative. At the top, there is a logo with a globe and the text 'OPEN SCIENCE AND RESEARCH'. Below the logo is a navigation menu with links: 'Frontpage', 'About the Initiative', 'Services', 'Training and Documentation', and 'Exploring Open Science'. A search bar is located in the top right corner. The main content area is divided into several sections:

- For Researchers:** Features an image of a person sitting in a field with a laptop. Below the image are three bullet points: 'Find services that support research', 'Read the Open Science Handbook', and 'Explore the Data Management Guide'.
- Support for Research:** Features an image of a laptop on a blue background with data lines. Below the image are two bullet points: 'Explore UNESCO's Open Access Curriculum' and 'Read selected articles on openness'.
- For Organisations:** Features an image of a group of people holding hands. Below the image are two bullet points: 'Open Science policy in other countries' and 'Explore other Open Science Initiatives'.
- Nordic Open Science F...:** Features an image of a building. Below the image is a text snippet: '21.-23.11. Thank you for attending. Materials are coming soon' and a 'Read more' link.
- Science Landscape in Finland:** Features a text block starting with 'Being one of the European and world innovation leaders, Finland invests in knowledge-based competence and aims to increase the overall standard of education. In 2012, R&D expenditure represented 3.6 % of the gross domestic product (GPD), a total sum of approximately seven billion euros and that of public research funding was around 1%. The Government Programme stresses the role of research as the foundation of knowledge and know-how, which promotes sustainable economic growth and immaterial as well as material welfare. To foster the research system in Finland towards better competitiveness and higher quality, Finland's strategy and roadmap for research infrastructures 2014-2020 was developed. The Finnish policy and actions in enhancing open science and research and a panoramic view on the'.
- News and Events:** Features a list of news items: 'Registration to Nordic Open Science and Research Forum is open', 'Open Citizen Science project begins', and 'Transparency and openness to scientific publishing: the Finnish research organisations pay millions of euros annually to the large publishers'.
- More news:** A link to view more news.
- We also tweet!:** A section for tweets, showing a tweet from @AvoinTiede.

At the bottom of the page, there is a footer with the text 'The Ministry of Education and Culture of'.

# Background: Open Science and Research (2)

- A top-down initiative, with many stakeholders involved
  - Successful in raising awareness both among the researchers and decision makers
  - Led to the creation of new open science policies at many Finnish research organizations
  - Provided funding for several projects and services
  - Main focus of the initiative was on research data
- There were ambitious national targets set for open access as well: 65% OA in 2017, 75% in 2018 and 100% in 2020



# National data collection

- The Finnish OA monitoring efforts are integrated with the national data collection conducted for the Ministry of Education and Culture
  - The Ministry has been collecting publication data from the universities since 2011
  - The universities of applied sciences joined in 2012, central hospital districts and some of the state research institutes in 2015
- The universities have a lot of motivation to provide quality information
  - 13% of the state funding to the universities is distributed based on the number and quality of the publications (= 200 million euros a year)

# Open access and state funding?

- Would it be possible to use the open access availability of research output as a factor in the funding model for the universities?
  - Possibly starting from 2019, as the funding is based on the publication data from the previous three years
  - The committee responsible for the university funding model has discussed this, but no decision has been made

# Collecting the publication data

- CSC - IT Center for Science is responsible for the actual data collection
  - Close collaboration with the research organizations, the Ministry, the Federation of Learned Societies and the National Library
- The publication metadata is usually collected from the Current Research Information System of each organization
  - An automated data collection process was launched in 2016
  - Many organizations still upload their data on annual basis - deadline at the end of March

Type of Publication	
A Peer-reviewed sci...	27,114
A1 Journal articl...	19,421
A2 Review article...	1,102
A3 Book section, ...	2,835
A4 Conference pr...	3,756
B Non-refereed scie...	4,039
C Scientific books (mo...	768
D Publications intend...	8,950
E Publications intend...	4,567
G Theses	1,756

# Technical platforms

- The data is collected in XML format into the [Virta publication information service](#) built by CSC
- The de-duplicated publication data can be browsed at Juuli, [www.juuli.fi](http://www.juuli.fi), a VuFind-based discovery interface run by the National Library
- The data in Juuli is updated automatically from Virta each night

Showing 81 - 100 of 7,280 for search: "", query time: 0.08s

Sort

Narrow Search

Remove Filters

- Open Access: Green OA
- Year of Publication: 2017
- Language: English

Institution

- University
- Health Care District
- Research Institute
- University of Applied Sciences


Author


- Eerola, P.
- Luukka, P.
- Härkönen, J.
- Tuovinen, E.
- Lampén, T.
- Karimäki, V.
- Kinnunen, R.
- Lassila-Perini, K.
- Lehti, S.
- Tuominiemi, J.


more ...


Field of Science


- Natural sciences
- Medical and health sciences
- Engineering and technology

81  **3D Printing : Challenging Existing Business Models**  
by: Ojala, Arto, Hamäläinen, Mervi  
Springer International Publishing, 2017  
A3 Book section, Chapters in research books  
[Get full text](#)  
[Get full text](#)  
[Get full text](#)

82  **3D Super-Resolution Optical Profiling Using Microsphere Enhanced Mirau Interferometry**  
by: Kassamakov, Ivan; Lecler, Sylvain; Nolvi, Anton; Leong-Hoi, Audrey; Montgomery, Paul; Haeggström, Edward  
Scientific Reports, 2017  
A1 Journal article (refereed), original research  
[Get full text](#)  
[Get full text](#)

83  **3D histopathological grading of osteochondral tissue using contrast-enhanced micro-computed tomography**  
by: Nieminen, H. J.; Gahunia, H. K.; Pritzker, K. P. H.; Ylitalo, T.; Rieppo, L.; Karhula, S. S.; Lehenkari, P.; Haeggström, E.; Saarakkalla, S.  
osteoarthritis cartilage, 2017  
A1 Journal article (refereed), original research  
[Get full text](#)  
[Get full text](#)

84  **3D mosquito screens to create window double screen traps for mosquito control**  
by: Khattab, Ayman; Jylha, Kaisa; Hakala, Tomi; Aalto, Mikko; Malima, Robert; Kisinza, William; Honkala, Markku; Nousiainen, Pertti; Meri, Seppo  
parasites vectors, 2017  
A1 Journal article (refereed), original research  
[Get full text](#)  
[Get full text](#)  
[Get full text](#)

85  **3D printed and photonicly cured graphene UHF RFID tags on textile, wood, and cardboard substrates**  
by: Akbari, M.; He, H.; Juuti, J.; Tentzeris, M. M.; Virkki, J.; Ukkonen, L.  
international journal antennas propagation, 2017  
A1 Journal article (refereed), original research  
[Get full text](#)  
[Get full text](#)  
[Get full text](#)

# OA status in the national data collection

- There used to be serious issues with the quality of the OA status data
  - Starting from 2016, improved definitions and categories
- Two fields for indicating the OA status
  - One field indicates whether the publication is OA via a Gold or a Hybrid channel
  - The second field is used for indicating whether it has been deposited into a repository (Green OA)
- URLs should be reported for each version, preferably based on DOIs, URNs, Handles

Year of Publication	▲
2016	✓

Open Access	▲
Gold OA publication channel	12,106
Green OA	5,902
Other OA availability	1,119



# What qualifies as Open Access?

- It would be very nice to have a short and easy-to-understand definition for Open Access
- However, Open Access is actually a relatively complicated issue, with many different flavors

Gold OA?

Green OA?

Hybrid OA?

Bronze OA?

Gratis OA?

Libre OA?

Embargoes?

Versions?



# OA definition

- Starting from 2016, the OA rules were made as clear as possible
  - The publication should be at least free to read (“gratis OA”)
  - Gold/Hybrid and Green Open Access are all accepted
  - Embargoes allowed for Green OA but not for Gold or Hybrid
  - Depositing into an institutional or subject-based repository is OK, but papers available at personal/project websites or at walled gardens like Research Gate or Academia.edu don't count
  - The OA version must be peer-reviewed to qualify

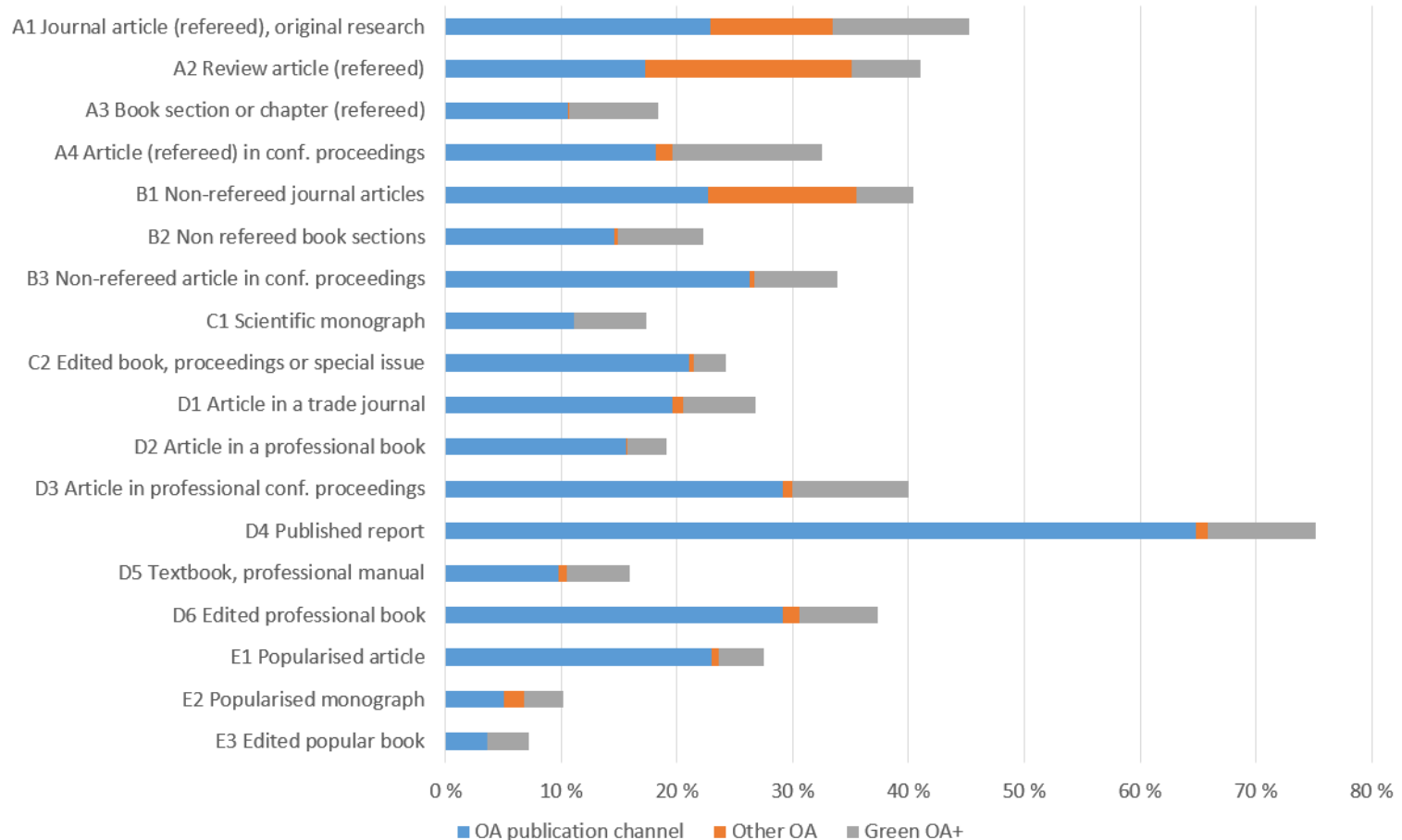


# Observations after two years...

- We now have two years of data (2016, 2017) that has been collected using the new guidelines
- Although the quality of the data is better than before, it is still far from perfect
  - Some organizations are putting more effort into it than others
  - Some of the organizations have been struggling with the identification of OA publication channels vs. Hybrid/Other OA
  - There are still some technical issues with de-duplication, especially with Gold and Hybrid OA
- No data on embargoes, licenses or APC costs collected at the moment
  - Embargoed publications can be reported as Green OA, if the final URL is already available

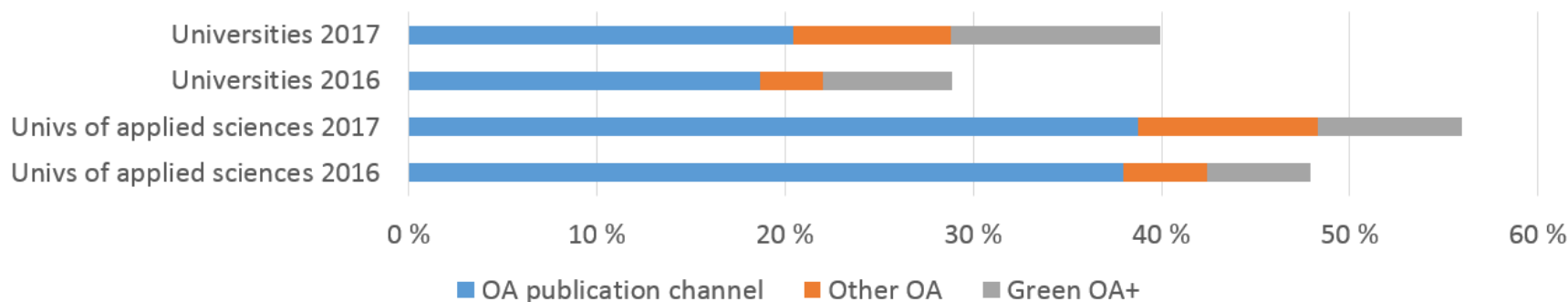
# OA in 2017 – some (very) preliminary data

The percentage of OA publications at Finnish universities in 2017,  
by publication type, data collected on 26/3/2018



# The good news: OA is increasing

The share of OA among the peer-reviewed articles and book chapters (A1-A4) at Finnish universities and universities of applied sciences in 2016 and 2017, data collected on 26/3/2018



- The share of peer-reviewed OA articles at the Finnish universities rose from 28.9 % in 2016 to 39.9 % in 2017
- In the universities of applied sciences the amount OA went from 49.9 % to 56.0 %
- Most of this increase is probably based on actual change, but some of it may be explained by improvements in the quality of the data

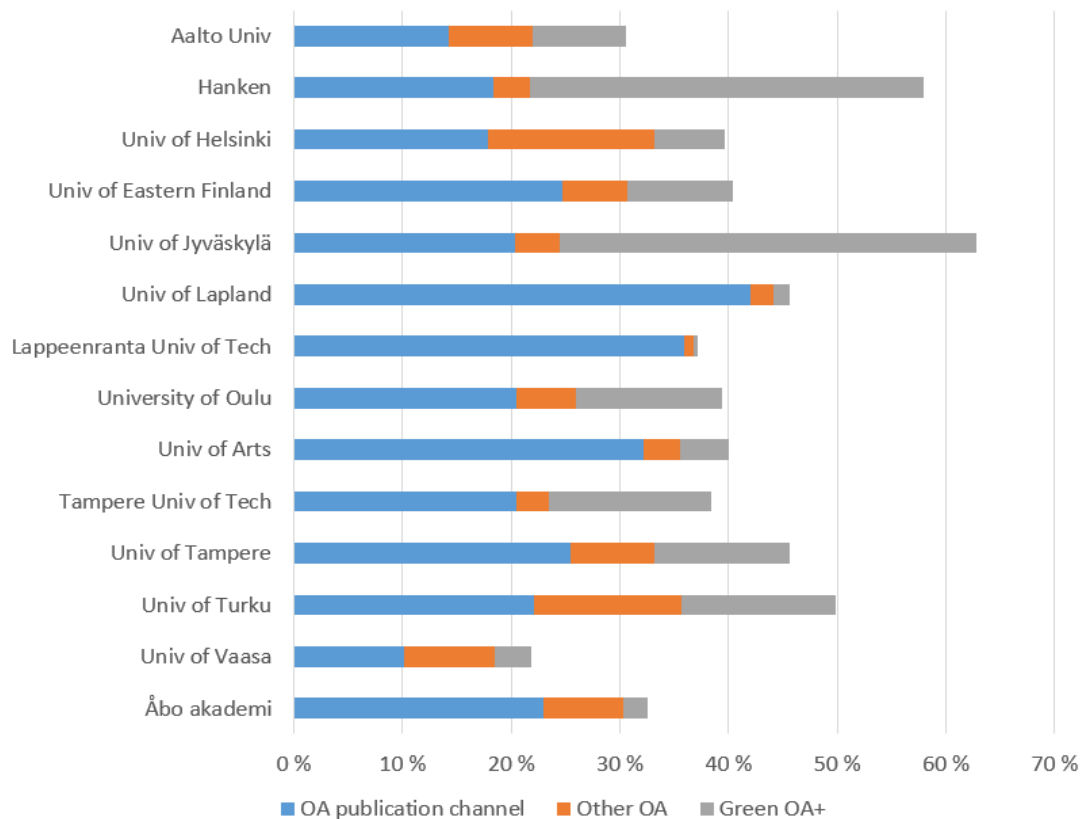
# The 2017 data: Peer-reviewed articles (A1-A4) and OA at Finnish universities

University	All publications	OA publication channel	Other OA	Green OA	OA total	% of OA	Green OA+
Aalto Univ	3225	459	250	895	986	30,6 %	277
Hanken	202	37	7	98	117	57,9 %	73
Univ of Helsinki	7084	1264	1087	2423	2814	39,7 %	463
Univ of Eastern Finland	2317	574	137	708	936	40,4 %	225
Univ of Jyväskylä	1965	400	82	1209	1235	62,8 %	753
Univ of Lapland	335	141	7	16	153	45,7 %	5
Lappeenranta Univ of Tech	838	301	7	8	312	37,2 %	4
University of Oulu	2302	471	126	743	906	39,4 %	309
Univ of Arts	90	29	3	8	36	40,0 %	4
Tampere Univ of Tech	1428	293	43	455	548	38,4 %	212
Univ of Tampere	1955	497	152	716	893	45,7 %	244
Univ of Turku	3320	734	449	1300	1655	49,8 %	472
Univ of Vaasa	265	27	22	18	58	21,9 %	9
Åbo Akademi	935	215	69	70	305	32,6 %	21
<b>All universities</b>	<b>23006</b>	<b>4707</b>	<b>1923</b>	<b>7126</b>	<b>9186</b>	<b>39,9 %</b>	<b>2556</b>
<b>All universities of applied sciences</b>	<b>627</b>	<b>243</b>	<b>60</b>	<b>140</b>	<b>351</b>	<b>56,0 %</b>	<b>48</b>

# Different types of open access

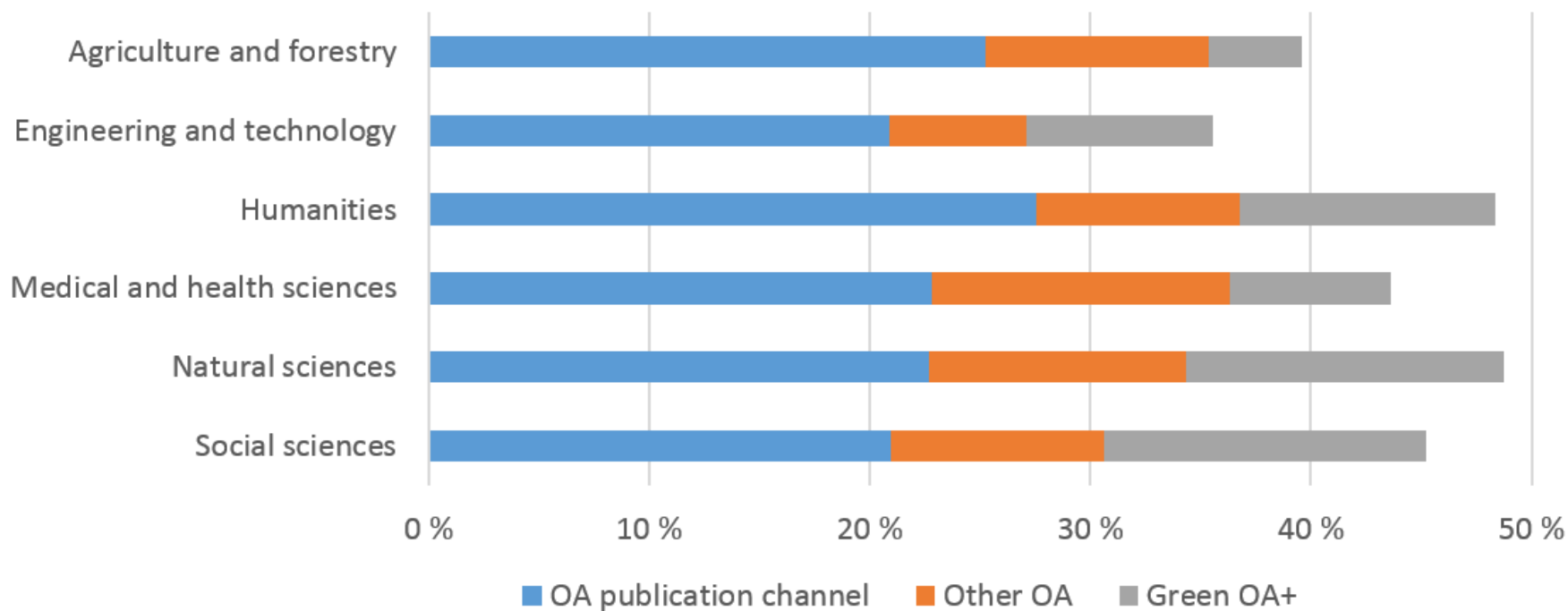
- Lots of overlap between Gold/Hybrid and Green OA
- Some universities systematically collect all of their Gold and Hybrid OA publications into their own repositories
- Green OA+: Additional OA provided by Green OA was about 11 % of the total peer-reviewed article output in 2017

The share of OA among the peer reviewed articles and book chapters (A1-A4) at each Finnish university in 2017, data collected on 26/3/2018



# OA in different research fields in 2017

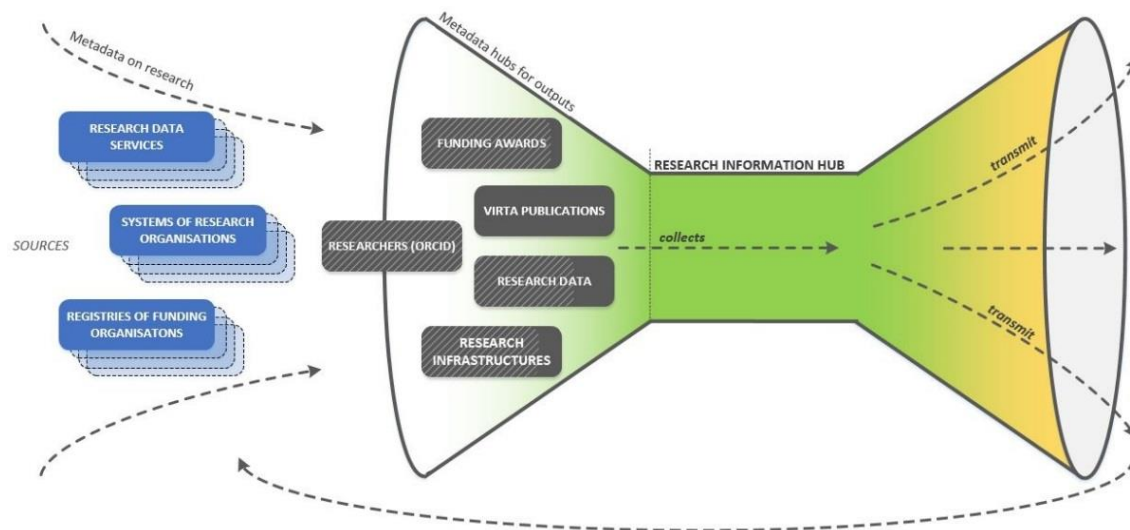
Peer-reviewed journal articles (A1-A2) at the Finnish universities in 2017,  
percentage of OA by field of science, data collected on 26/3/2017





# Plans for the future

- The Ministry of Education and Culture has launched a new project (2017-2020) which is developing [a National Research Information Hub](#)
  - The development of the Virta service will continue, but the new system will also cover e.g. research data and infrastructures
- The collection of OA data is one of the areas which may require further development



# New developments

- The current Finnish OA monitoring guidelines were created in 2015 - there have been quite a few developments on an international level since then
  - New tools and data sources like [Unpaywall](#) now available
  - A fast and fairly reliable way to check the current OA status of a large number of publications (as long as they have a DOI)
- We haven't yet looked at what would be the best way to utilize these tools, either on a national level or at a local level at each organization
  - Some of the organizations may be already using them

# Changing categories

- A recent article by [Piwowar et al.](#) adopted a strict definition of Gold OA and introduced a new category, Bronze OA
  - Bronze OA includes both delayed OA and the OA publication channels that don't use CC licenses
- Not compatible with the current Finnish OA definition
  - We have a more inclusive definition of Gold OA and don't accept delayed OA at all

✓ PEER-REVIEWED

## The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles

PeerJ

View  
394  
tweets

Related  
research  
▼

Research article | Legal Issues | Science Policy | Data Science

Heather Piwowar<sup>1\*</sup>, Jason Priem<sup>1\*</sup>, Vincent Larivière<sup>2,3</sup>, Juan Pablo Alperin<sup>4,5</sup>, Lisa Matthias<sup>6</sup>, Bree Norlander<sup>7,8</sup>, Ashley Farley<sup>7,8</sup>, Jevin West<sup>7</sup>, Stefanie Haustein<sup>3,9</sup>

Published February 13, 2018

📌 Note that a [Preprint of this article](#) also exists, first published August 2, 2017.

PubMed [29456894](#)

> Author and article information

▼ Abstract

Despite growing interest in Open Access (OA) to scholarly literature, there is an unmet need for large-scale, up-to-date, and reproducible studies assessing the prevalence and characteristics of OA. We address this need using oaDOI, an open online service that determines OA status for 67 million articles. We use three samples, each of 100,000 articles, to investigate OA in three populations: (1) all journal articles assigned a Crossref DOI, (2) recent journal articles indexed in Web of Science, and (3) articles viewed by users of Unpaywall, an open-source browser extension that lets users find OA articles using oaDOI. We estimate that at least 28% of the scholarly literature is OA (19M in total) and that this proportion is growing, driven particularly by growth in Gold and Hybrid. The most recent year analyzed (2015) also has the highest percentage of OA (45%). Because of this growth, and the fact that readers disproportionately access newer articles, we find that Unpaywall users encounter OA quite frequently: 47% of articles they view are OA. Notably, the most common mechanism for OA is not Gold, Green, or Hybrid OA, but rather an under-discussed category we dub Bronze: articles made free-to-read on the publisher website, without an explicit Open license. We also examine the citation impact of OA articles, corroborating the so-called open-access citation advantage: accounting for



# Categories vs. attributes?

- It seems that the commonly used OA categories (Gold, Hybrid, Green, Bronze, Black, etc.) and their definitions are constantly evolving
  - Instead of trying to adapt to all of these changes and conflicting definitions it might be a good idea to use attributes instead of categories
  - The categories (if they are needed) could be generated from the attributes
  - But there may be a downside as well: you might end up with quite a few OA fields per publication...
- publication
    - publisher's version
      - current OA status (yes/no)
      - link: publication channel (OA/not)
      - license
      - delay end date (if known)
      - APC paid by the organization (euros)
      - identifier / url
    - pre-print version (open)
      - license (if other)
      - embargo end date
      - identifier / url
    - self-archived version
      - version (final draft, etc.)
      - license (if other)
      - embargo end date
      - identifier / url

# Some final thoughts

- Although it is easy to add new fields into the data collection requirements, each extra piece of information has a cost
  - The collection of additional data requires resources and often also changes to the information systems used by the reporting organizations
  - If you ask the organizations to provide information there should be a strong motivation for doing it
- What is it that you are really trying to measure, and why?
  - New data sources and automatic tools provide opportunities for creating better and more cost-effective processes
  - However, you shouldn't necessarily do things in a certain way just because it happens to be easy

**Thank you!**



THE NATIONAL LIBRARY OF FINLAND