Open Access Indicator 2019

Part 1

Overview of data foundation, processes and output
Publications from publication year 2017.

Version 1, 3 April 2019

0   Contents

1 Preface

Denmark's National Strategy for Open Access 2018-2025[1] has the target: *"From 2025 onwards, there should be unimpeded digital access for all to all peer-reviewed scientific articles from Danish research institutions – with a maximum 12- month embargo."*

Furthermore, the strategy specifies: *"Once a year, the Ministry of Higher Education and Science will monitor the implementation of Open Access via the Open Access Indicator (OAI) to support that all parties are doing their utmost to develop and expand free accessibility to Danish research results."*

The Open Access Indicator is calculated for each publication year with the scope/target field set to: "*Scientific articles and conference contributions in journals and proceedings with an ISSN number.*

Many journals maintain embargo periods of 6, 12, or more months, during which researchers are blocked from establishing open Access to their articles. Similar to the EU Horizon 2020[2] research programme, the national strategy allows up to 12 months delay in establishing Open Access to a given article.

The Open Access Indicator (OAI) is calculated once annually with a delay that allows for a potential 12-month embargo period for articles published at the end of a publication year. Thus OAI 2019 is calculated early March 2019 for publications from the publication year 2017.

The description of the Open Access Indicator is organized in two parts:

- Part 1: Overview of data foundation, processes and output
- Part 2: Technical description of data foundation, processes and output

  Note: In Part 2, the technical description, the notion of the indicator's "target field" is expressed using the term "set of scoped records".

Queries regarding the indicator may be directed to

> Hanne-Louise Kirkegaard
> Danish Agency for Science and Higher Education
> Ministry of Higher Education and Science
> Bredgade 40
> DK-1260 København K
> Email: hki@ufm.dk

---

[1] https://ufm.dk/en/research-and-innovation/cooperation-between-research-and-innovation/open-access/Publications/denmarks-national-strategy-for-open-access/denmarks-national-strategy-for-open-access
[2] https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

2 Which types of Open Access are measured?

The Open Access Indicator for 2019 measures three types of Open Access:

1. Articles and conference contributions, published in dedicated Open Access journals and thus available in Open Access as soon as they are published (Gold OA)

2. Articles and conference contributions, which are available in Open Access from the universities' own research databases or Open Access repositories (Green OA)[3]

3. Articles and conference contributions, which are available in Open Access from other recognized Open Access repositories (Green OA).

The OA Indicator does not measure so-called Hybrid OA, where an article is published in a journal only accessible to subscribers, and where an additional fee is paid to release the individual article in OA to non-subscribers. However, if such hybrid OA articles are made available from the universities' research databases or other recognized Open Access repositories, they will be credited just like all other cases belonging to type 2 or 3 above.

3 What is the data foundation?

- Metadata about the year's publications are collected from the local research databases of each of the 8 Danish universities. This constitutes the basic data of the OA Indicator.

- The result of the year's Danish Bibliometric Research Indicator (BFI)[4] is imported, in order to ensure that the OA Indicator reuses the results from the BFI duplicate- and data conflict handling. Duplicates occur when two or more universities collaborate on a publication, which is consequently collected from more than one local research database. Data conflicts occur when duplicate registrations from more than one university disagree on details that require unambiguity on a national level.

- International metadata for dedicated gold Open Access journals is imported from the Directory of Open Access Journals[5]. This enables the OA Indicator to evaluate Open Access type 1 above (Gold OA)

- A Danish authority list of recognized and compatible Open Access repositories other than the Danish University research databases. This enables the OA Indicator to evaluate Open Access type 3 above (Green OA)

- International metadata about the publishers' and journals' policies wrt. Green OA is imported from the Sherpa/Romeo[6] database. This enables the OA Indicator to establish the Open Access potential of articles in subscription journals. I.e. how many

---

[3] Danish research databases (Current Research Information Systems) are primarily metadata registries of the university's publications, but they may also perform the task of being "repositories" for the publications in full text. In other cases, the universities use other dedicated systems as Open Access repositories.

[4] http://ufm.dk/forskning-og-innovation/statistik-og-analyser/den-bibliometriske-forskningsindikator

[5] https://doaj.org

[6] http://www.sherpa.ac.uk/romeo/

articles could become Open Access via the universities' research databases and other recognized OA repositories without breaking publisher policies?

- A Danish authority list of journals with very long embargo periods (more than the 12 months), which block the establishment of Open Access as prescribed by the Danish strategy and EU's Horizon 2020. The list is used to distinguish between publications with Unused OA-rights and those with Blocked OA.

3.1 Recognized and compatible Open Access repositories

The OA Indicator measures Green OA from two types of repositories:

The universities local research databases or OA repositories
- AU http://vbn.aau.dk/en/
- AU http://pure.au.dk/portal/en/organisations/8000/publications.html
- CBS http://research.cbs.dk
- DTU http://orbit.dtu.dk
- ITU https://pure.itu.dk/portal/
- KU http://research.ku.dk/find-a-researcher/
- RUC https://forskning.ruc.dk
- SDU http://findresearcher.sdu.dk/portal/

Other recognized and compatible OA repositories
The following criterias should be met by the repository to be accepted on the authority list of external Open Access repositories:

- **Technical solidity**
  The repository is in stable operation with good up- and response times
- **Sustainability**
  The repository seems organizational and financial sustainable and can be expected to work for many years to come.
- **Credibility**
  The repository is dedicated to Open Access dissemination of scientific publications.
- **Contact**
  The repository has visible contact information and responds to mail inquiries within a reasonable time.
- **Accessibility**
  The repository's publications are freely available without any access barriers in the form of registration requirements or similar.
- **Link quality**
  The repository displays full texts via direct links: If a human or a machine (robot) follows the link (in a browser etc.), the actual text is downloaded directly.

The authority list is updated yearly based on input from the Danish universities. In a hearing period, the universities may suggest new repositories to be added to the list. The suggested repositories are evaluated in accordance with the six criteria and if these are met, the repositories will be included on the authority list of external Open Access repositories.

The list is available for download on the OA Indicator website.

3.2 Journals with very long embargo periods

The authority list of journals with embargo periods exceeding 12 months is updated annually. Embargo lists are gathered from five major publishers and, in a subsequent hearing period, the universities may propose other journals to be added to the authority list. Finally, information is obtained on eventual national publisher agreements, which shorten any of these embargo periods. This is collected from the secretariat of the national licensing consortium undertaking the negotiations with the publishers.

The list is available for download on the OA Indicator website.

4 How does the process flow – in brief?
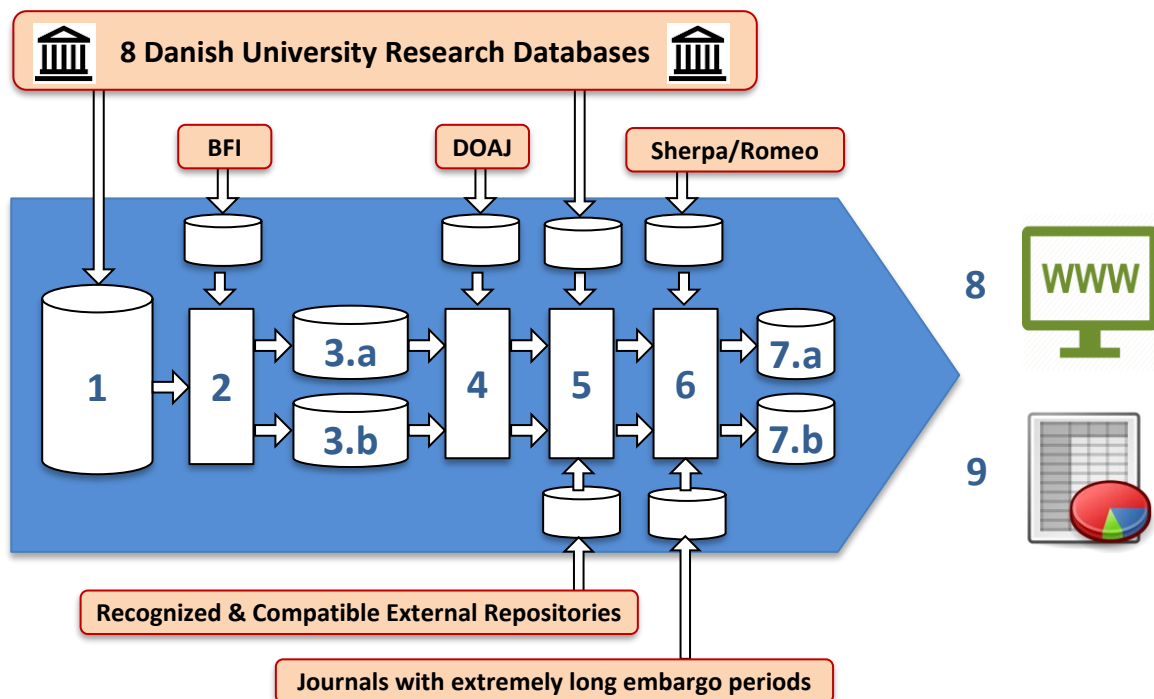
1.   Publication metadata is collected automatically from the universities.

2.   A subset corresponding to the definition of the OA Indicator's target field is isolated and two versions of the data set for further analysis are produced.

3.a: The version "**target field with duplicates**", which will contain duplicates when a given publication has authors from more than one university, as it will have been collected more than once. This version of the target field is used as basis for calculations dealing with individual universities.

3.b: The version "**target field without duplicates**", where deduplication of records has been carried out, among others using data from BFI (the Danish Bibliometric Research Indicator). This version is used as basis for calculations on the national level and by main research area.

4. Check whether the article is published in a Gold OA journal. In which case it is labelled as Realised OA. Data from DOAJ (Directory of Open Access Journals) is used to ensure that the journal is a fully dedicated Gold OA journal. DOAJ also provides information on journals charging a publication fee (APC - Article Processing Charge). This information is available as a spreadsheet for download at the OA Indicator website.

5. Check whether the article in practice may be downloaded from a Green Open Access repository. In which case it is labelled as Realised OA.
    - Either from a Danish university's research database
    - Or from a recognized external OA repository (on the authority list of such)

6. Check whether the article is published in a journal with Green OA rights. In which case it is labelled as Unused OA rights. If not it is labelled as Blocked OA. Here data from the Sherpa/Romeo-database is used. In case of Unused OA rights, the journal is checked against the list of journals with very long embargo periods (more than 12 months). If the journal appear on this list, the article is relabeled to Blocked OA.

    This results in a statistical dataset in two parts:

7.a: University level statistics calculated using "**target field with duplicates**"

7.b: National level statistics calculated using "**target field without duplicates**".

8. The result is communicated via web-pages of the Danish Open Access Indicator and ...

9. ... via spreadsheets, which may be downloaded from the Danish Open Access Indicator. Also the underlying publication data may be downloaded as spreadsheets.

4.1 Quality assurance process

The results of the OA Indicator 2019 have been subjected to the following quality assurance measures:

- **Data Foundation.** The collected data has been tested in order to make sure the data foundation has been harvested correctly and in accordance with the established process. The test is based on a random sample of 2% of each university's data set.

- **Downloaded fulltext files.** The test examins downloaded fulltext files that appear to deviate from registered metadata; deviations in page number, very small files etc.