

Open Access Indikator for 2016

Del 1

Overblik over datagrundlag, processer og output

Version 1, den 10. april 2018

0 Indhold

0	Indhold	1
1	Forord.....	2
2	Hvilke typer Open Access måles?	3
3	Hvad er datagrundlaget?	3
3.1	Anerkendte og kompatible Open Access repositorier.....	4
3.2	Tidsskrifter med særligt lange embargoperioder	5
4	Hvordan forløber processen i hovedtræk?.....	5
4.1	Proces for kvalitetskontrol.....	6

1 Forord

Den Nationale Styregruppe for Open Access¹ har indstillet til Styrelsen for Forskning og Innovation og Danmarks Elektroniske Fag- og Forskningsbibliotek, at der udvikles en dansk Open Access Indikator. Denne skal støtte implementeringen af den nationale Open Access strategi² - jf. strategiens bemærkninger om monitorering: *"Implementeringen af Open Access skal løbende monitoreres for at sikre, at alle parter gør deres ypperste for at udvikle og udbrede fri tilgængelighed til danske forskningsresultater"*.

Open Access Indikatoren beregnes en gang årligt med genstandsfeltet: *Videnskabelige og fagfællebedømte artikler og konferencebidrag i tidsskrifter og proceedings med ISSN.*

EU kræver i forbindelse med Horizon 2020³, at Open Access realiseres senest 6 måneder efter publicering for hovedområderne naturvidenskab, teknologi og sundhedsvidenskab og senest 12 måneder efter publicering for hovedområderne humaniora og samfundsfag. Denne forsinkelse skyldes, at mange tidsskrifter opretholder såkaldte embargoperioder, hvor de udelukker forskerne fra at etablere Open Access til artiklerne før embargoperiodens udløb.

Da OA Indikatoren beregnes én gang årligt for alle publikationer inden for genstandsfeltet, er den indrettet til at acceptere et års forsinkelse i Open Access adgangen til publikationerne. Således er OA Indikatoren for 2016 beregnet primo marts 2018 for at tillade et helt års embargoperiode også for publikationer fra december 2016. I praksis betyder dette, at publikationer fra januar 2016 vil kunne have embargoperioder på helt op til 24 måneder og stadig blive godskrevet af OA Indikatoren.

Beskrivelsen af Open Access Indikatoren er organiseret i to dele:

- Del 1: Overblik over datagrundlag, processer og output
- Del 2: Teknisk beskrivelse af datagrundlag, processer og output

Henvendelser vedr. indikatoren kan rettes til

Adam Baden/Hanne-Louise Kirkegaard
Styrelsen for Forskning og Uddannelse
Uddannelses- og Forskningsministeriet
Bredgade 40
1260 København K
Email: aba@ufm.dk / hki@ufm.dk

¹ <http://ufm.dk/forskning-og-innovation/samspil-mellem-viden-og-innovation/open-access/artikler/den-nationale-styregruppe>

² <http://ufm.dk/forskning-og-innovation/samspil-mellem-viden-og-innovation/open-access/billeder-og-filer/danmarks-nationale-strategi-for-open-access.pdf>

³ https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

2 Hvilke typer Open Access måles?

Open Access Indikatoren for 2016 måler tre typer Open Access:

1. Artikler og konferencebidrag, der er udgivet i dedikerede Open Access tidsskrifter og dermed OA-tilgængelige, så snart de er publiceret (Gylden OA)
2. Artikler og konferencebidrag, der er OA-tilgængelige fra universiteternes egne forskningsdatabaser eller Open Access repositorier (Grøn OA)⁴
3. Artikler og konferencebidrag, der er OA-tilgængelige fra andre anerkendte Open Access repositorier (Grøn OA).

OA Indikatoren måler ikke såkaldt Hybrid OA, hvor artiklen publiceres i et tidsskrift, der kun giver adgang til abonnenter, og hvor et tillægsgebyr er betalt for at frigive den enkelte artikel som OA til ikke-abonnenter. Dog, i det omfang en sådan hybrid OA-artikel gøres tilgængelig fra universitetets egen forskningsdatabase eller andet anerkendt Open Access repositorie, vil den tælle med ligesom alle andre tilfælde i type 2 eller 3 ovenfor.

3 Hvad er datagrundlaget?

- Metadata om årets publikationer indsamles fra hvert af de 8 universiteters lokale forskningsdatabaser. Dette udgør OA Indikatorens basisdata.
- Resultatet af årets Bibliometriske Forskningsindikator(BFI) ⁵ importeres. Dette anvendes til at sikre, at OA Indikatorens dublethåndtering i videst muligt omfang genbruger BFI's dublet- og konfliktløsningsbeslutninger. Doubletter skyldes, at to eller flere universiteter har samarbejdet om en publikation, der følgelig er indsamlet fra mere end én lokal forskningsdatabase. Konflikter opstår, når dublet-registreringer fra mere end én lokal forskningsdatabase ikke er enige om registreringsdetaljer, som på nationalt niveau kræver entydighed.
- Internationale metadata om dedikerede gyldne Open Access tidsskrifter importeres fra Directory of Open Access Journals⁶. Herved kan OA Indikatoren håndtere Open Access type 1 ovenfor (Gylden OA)
- En dansk autoritetsliste over andre anerkendte og kompatible Open Access repositorier ud over de danske universiteters forskningsdatabaser. Herved kan OA Indikatoren håndtere Open Access type 3 ovenfor (Grøn OA)
- Internationale metadata om forlagernes og tidsskrifternes politikker, hvad angår Grøn OA importeres fra Sherpa/Romeo⁷ databasen. Herved kan OA Indikatoren beregne

⁴ Danske forskningsdatabaser er primært registrarer over universitetets publiceringer, men kan også løfte rollen som "repositorier" for selve publikationerne (i fuld tekst). I andre tilfælde anvender universiteterne et andet system til at løfte rollen som Open Access repositorie.

⁵ <http://ufm.dk/forskning-og-innovation/statistik-og-analyser/den-bibliometriske-forskningsindikator>

⁶ <https://doaj.org>

⁷ <http://www.sherpa.ac.uk/romeo/>

Open Access potentialet for artikler i abonnements-tidsskrifter. Dvs. hvor mange artikler kunne gøres OA-tilgængelige fra universiteternes forskningsdatabaser og andre anerkendte OA repositorier uden at bryde forlagernes politikker?

- En dansk autoritetsliste over tidsskrifter med særligt lange embargoperioder, dvs. ud over de 12 måneder, der accepteres af EU.

3.1 Anerkendte og kompatible Open Access repositorier

OA Indikatoren måler grøn OA fra to typer repositorier:

Universiteternes egne forskningsdatabaser eller OA repositorier

- 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/>

Andre anerkendte og kompatible OA repositorier

Følgende kriterier skal være opfyldt for at kunne optages på autoritetslisten over anerkendte eksterne repositorier:

- **Teknisk soliditet**
Repositoriet er i stabil drift med gode oppe- og svartider.
- **Bæredygtighed**
Repositoriet forekommer bæredygtigt organiseret og finansieret og kan forventes at fungere mange år frem i tiden.
- **Troværdighed**
Repositoriet er dedikeret til Open Access formidling af videnskabelige publikationer.
- **Kontaktbarhed**
Repositoriet har synlige kontaktoplysninger og svarer inden for en rimelig tid på mail-henvendelser.
- **Tilgængelighed**
Repositoriets publikationer er frit tilgængelige uden adgangsbarrierer i form af registreringskrav eller lign.
- **Linkkvalitet**
Repositoriet udstiller sine fuldtekster via direkte links: Hvis et menneske eller en maskine (robot) følger linket (i en browser e.l.) downloades selve fuldteksten direkte.

Listen opdateres årligt på basis af input fra de danske universiteter. Universiteterne har mulighed for, i en høringsperiode, at foreslå eksterne repositorier til optagelse. De foreslåede repositorier evalueres i henhold til de seks kriterier og kan, hvis kriterierne opfyldes, optages på listen.

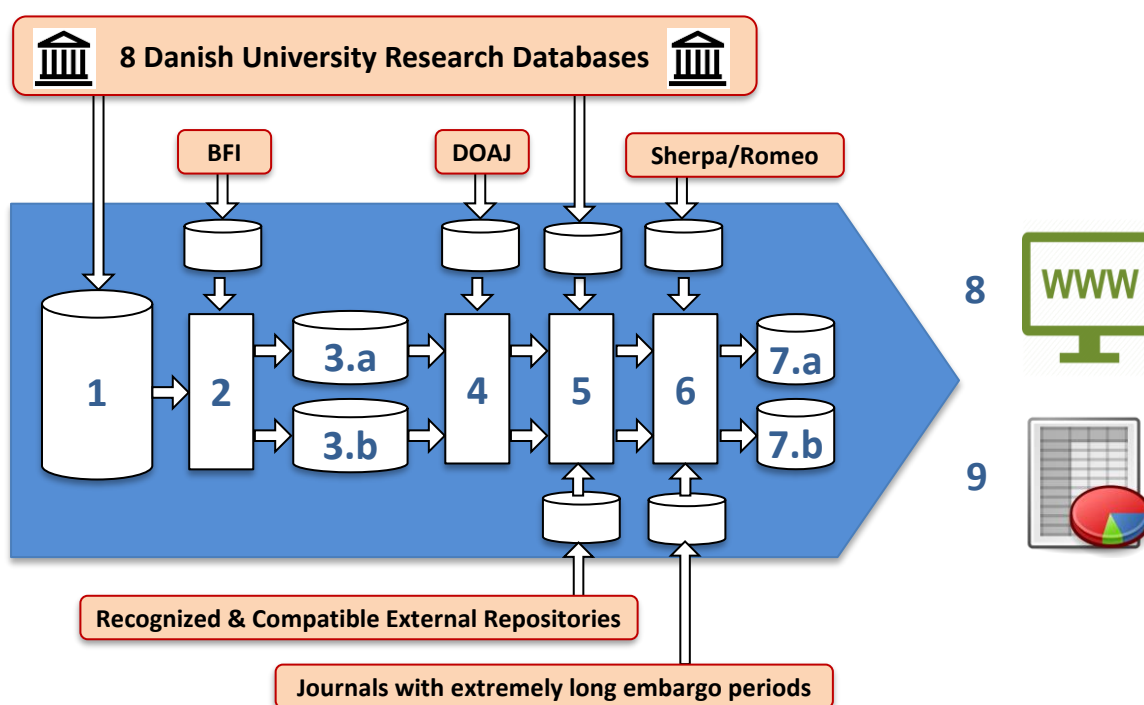
Listen kan downloades fra OA Indikatorens hjemmeside.

3.2 Tidsskrifter med særligt lange embargoperioder

Autoritetslisten over tidsskrifter med embargoperioder udover 12 måneder, opdateres årligt. Der indsamles embargolister fra fem store forlag og universiteterne kan derudover, i en høringsperiode, foreslå yderligere tidsskrifter med særligt lange embargoperioder. Forslagene evalueres hvorefter de tilføjes autoritetslisten.

Listen kan downloades på OA Indikatorens hjemmeside.

4 Hvordan forløber processen i hovedtræk?



1. Universiteternes publikationsdata indsamles automatisk fra universiteterne.
2. Den delmængde, der svarer til OA Indikatorens genstandsfelt, isoleres og to versioner af datasættet produceres mhp. yderligere analyser:
 - 3.a Versionen "**genstandsfelt med dubletter**", der indeholder dubletter når en publikation har forfattere fra flere universiteter og dermed indsamles flere gange. Denne version af genstandsfeltet bruges som grundlag for beregninger, der vedrører de enkelte universiteter.
 - 3.b Versionen "**genstandsfelt uden dubletter**" hvor der er foretaget deduplicering af posterne, bl.a. på baggrund af data fra BFI (Den Bibliometriske Forskningsindikator). Denne version benyttes som grundlag for beregninger, der vedrører hele Danmark og opdelt på hovedforskningsområder.
4. Check om artiklen er publiceret i et gyldent OA tidsskrift. Her anvendes data fra DOAJ (Directory of Open Access Journals) til at sikre, at tidsskriftet er et fuldt dedikeret gyldent OA tidsskrift. Data fra DOAJ giver endvidere information om hvorvidt

tidsskrifter opkræver publiceringsgebyr i form af APC (Article Processing Charge). Denne information er for indikatorens genstandsdelt at finde i et regneark til download OA Indikatorens hjemmeside.

5. Check om artiklen i praksis kan downloades fra et grønt Open Access repository
 - enten et universitets forskningsdatabase
 - eller et andet anerkendt OA repository (optaget på autoritetslisten over sådanne)
6. Check om artiklen er publiceret i et tidsskrift med grønt OA potentiale. Her anvendes data fra Sherpa/Romeo-databasen. I tillæg slås op i autoritetslisten over tidsskrifter med særligt lange embargoperioder for at undgå at signalere grønt OA potentiale i tilfælde af embargoperioder ud over 12 måneder.

Dette resulterer i et todelt statistisk datasæt:

- 7.a Statistik på universitetsniveau beregnet vha. ”genstandsfelt med dubletter”
- 7.b Statistik på nationalt niveau beregnet vha. ”genstandsfelt uden dubletter”
8. Resultatet formidles via websider i den Danske Open Access Indikator og...
9. ... via regneark, der kan downloades fra den Danske Open Access Indikator. I tillæg hertil kan de underliggende publikationsdata downloades som regneark. Som noget nyt i OA Indikatoren for 2016 tilbydes regneark til download med en underinddeling af den realiserede OA kategori:
 - Grøn OA: Publikation med OA via et universitets lokale forskningsdatabase
 - Grøn OA: Publikation med OA via et anerkendt eksternt repository
 - Gylden OA: Publikation i et tidsskrift registreret i DOAJ med APC-opkrævning
 - Gylden OA: Publikation i et tidsskrift registreret i DOAJ uden APC-opkrævning

4.1 Proces for kvalitetskontrol

OA Indikatorens resultater for 2016 kvalitetssikres gennem følgende procedure;

- **Datagrundlaget.** De indsamlede data testes mhp. at sikre, at datagrundlaget er høstet korrekt og klassificeret i overensstemmelse med den fastlagte proces. Testen er stikprøvebaseret og har været repræsentativ på tværs af universiteternes data.
- **Downloadede fuldtekstfiler.** Testen fokuserer på downloadede filer der synes at afvige fra de registrerede metadata. Dette drejer sig f.eks. om afvigelser i sidetal, meget små filstørrelser mv.
- **Stikprøve.** En tilfældigt udvalgt stikprøve på 5%, af det samlede datasæt på realiseret OA potentiale fra hvert universitet, er gennemgået med henblik på et overordnet kvalitetstjek af datakvaliteten.

Bemærk

Sherpa/Romeo data anvendes til at afgøre, om en publikation, der ikke er OA realiseret, bør klassificeres som havende uudnyttet (gul) eller uklart (rød) potentiale. I forbindelse med kvalitetskontrollen af OA Indikator 2016 er der identificeret en potentiel mangel i de Sherpa/Romeo data, der er blevet downloadet og anvendt.

Det viser sig, at Sherpa/Romeo's online-tjeneste på hjemmesiden viderebringer oplysninger om flere tidsskrifter, end der findes i selve Sherpa/Romeo-databasen, som OAI og andre kan downloade. Dette skyldes, at online-tjenesten integrerer supplerende information fra to andre online-databaser i søgesvarene (DOAJ og ZETOC).

Denne diskrepans synes at medføre, at OA Indikatorens røde kategori fremstår lidt for stor og den gule kategori lidt for lille. Den grønne kategori (realiseret OA) er ikke påvirket af diskrepansen. Om dette er et betydningsfuldt problem, og om det kan afhjælpes, afventer en nærmere dialog med Sherpa/Romeo.