



Mining the messiness of Open Access for insights about research evaluation in the Humanities and Social Sciences

Clifford Tatum and Thed van Leewan
Centre for Science and Technology Studies
Leiden University, The Netherlands



**Universiteit
Leiden**

ESSH 2015 : Research Evaluation for the Social Sciences and the Humanities
Rennes, France —17 February 2015

Outline

Open Access & the Dutch situation

Challenges in measuring Open Access

International Comparison of OA (NL, DK, CH) - overview

Focus on Humanities and Social Sciences in the Netherlands

Situated communication practices

Open Access & the Dutch Situation

Public Statements (BBB of OA)

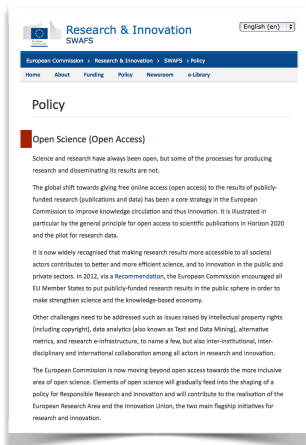
The collage features three key documents in the history of open access:

- Budapest Open Access Initiative (2002):** A document with a green header, defining open access as the free availability of the full text of journal articles and other types of digital content. It emphasizes the importance of removing barriers to the literature and education, and the role of libraries and other organizations in ensuring access.
- Bethesda Statement on Open Access Publishing (2003):** A document with a blue header, stating that the Internet has fundamentally changed the practical and economic realities of distributing scientific knowledge and cultural heritage. It calls for the creation of a global and interactive representation of human knowledge.
- Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (2003):** A document with a blue header, stating that the Internet has fundamentally changed the practical and economic realities of distributing scientific knowledge and cultural heritage. It calls for the creation of a global and interactive representation of human knowledge.

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.

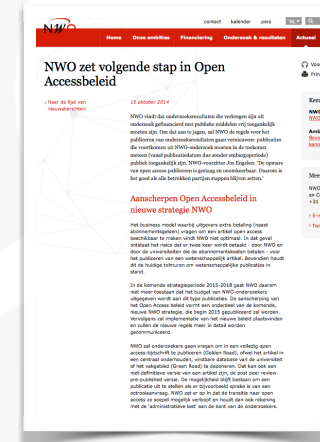
Budapest Open Access Initiative (2002)

OA Policies - funders




The global shift towards giving free online access (open access) to the results of publicly-funded research (publications and data) has been a core strategy [...] **the European Commission is now moving beyond open access towards the more inclusive area of open science.** Elements of open science will gradually feed into the shaping of a policy for Responsible Research and Innovation

Free and simple access to research results: facilitates the research having as big an impact as possible and everyone throughout the world being able to benefit from the research results; **helps science to make more rapid progress because insights can be shared early and data are available for reuse;** ensures that conclusions from scientific research are more easily verifiable as everyone has access to the underlying data.



Open Access Mandates

<p>CBS PUBLICATIONS</p> <p>CBS Open Journals</p> <p>Open Access at CBS</p>	<h2>Open Access at CBS</h2> <p>Since 2009 CBS has had an Open Access policy, the aim of which is to ensure free access to digital versions of CBS scholarly publications.</p> <p>CBS' Open Access-politik (PDF) > The Berlin declaration ></p>
<p>NEWS</p> <p>03.12.2014 EU growth plan - a way to save Europe? Find a CBS expert</p> <p>19.11.2014 CBS must save 60m from 2017 onwards</p> <p>See all news</p>	
<p>EVENTS</p> <p>No current events.</p> <p>See all events</p>	<p>In 2008 CBS signed the "The Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities" which in 2009 led to the formulation of an actual policy on Open Access at CBS.</p> <p>The CBS policy on Open Access stipulates that CBS and its researchers will work towards securing free access for all to read and use in digital form CBS scholarly publications. The aim is to consolidate the effect of CBS research by making it available without financial, technical or legal restrictions.</p>

Greater access to scholarly publications from CBS
The CBS Open Access Policy – 2009

Background
Universities find themselves in a situation in which research becomes more and more international through increased cooperation with other universities inside and outside of Europe. This has been the case for CBS for several years, where internationalization has been one of the key strategic goals.

Scientific information is increasingly digitized, journals appear more and more in e-format only and references to e-science and e-research gain ground. More and more researchers expect rapid access to research material and information and prefer search tools for information that provides easy access to content via the internet.


Many funders have recognized that the job of research is only half done if the results of that research cannot reach the widest audience. Some are formulating policies to require Open Access to their research, and the European Research Council has recommended open access policy for all EU funded research.

CBS as other universities find themselves in a transitional process which access to the results of their research is an important prerequisite to participating in the international research community and research competition.

And as a publicly funded university CBS has a duty to inform the general public about its research activities and results and to give access to published results of the research to industry and business stimulate knowledge exchange and further innovation.

In line with these considerations CBS last year signed the Berlin Declaration, which calls for unrestricted Open Access to Knowledge in the Sciences and Humanities. In order to implement the Berlin Declaration institutions should:

- Implement a policy to require their researchers to deposit all their published articles in an open access repository and
- encourage their researchers to publish their research articles in journals where a suitable journal exists and provide the support needed that to happen.



CBS Open Journals

CBS Open Journals offers free access to a string of scholarly journals co-published by CBS researchers.

[CBS Open Journals](#) >



CBS Open Journals is the CBS online platform for open access to a wide range of research articles and all Open Access.

Berlin Declaration
The Internet has fundamentally changed the practical and economic realities of distributing scientific knowledge and cultural heritage. For the first time ever, the Internet now offers the chance to constitute a global and interactive representation of human knowledge, including cultural heritage and the guarantee of worldwide access. We, the undersigned, feel obliged to address the challenges of the Internet as an emerging functional medium for distributing knowledge. Obviously, these developments will be able to significantly modify the nature of scientific publishing as well as the existing system of quality assurance.

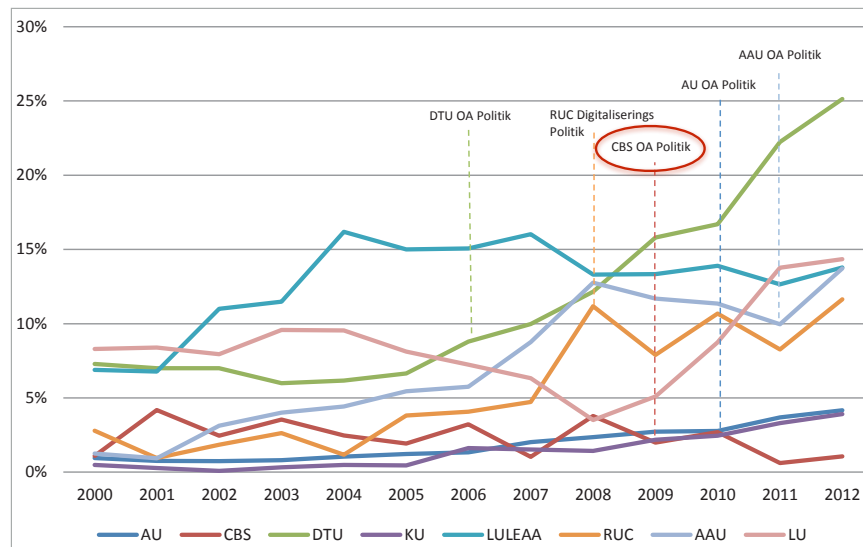
Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities
The Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities of 22 October 2003 was written in English. It is one of the milestones of the Open Access movement. The wording of the English version shall prevail.

Preface
The Internet has fundamentally changed the practical and economic realities of distributing scientific knowledge and cultural heritage. For the first time ever, the Internet now offers the chance to constitute a global and interactive representation of human knowledge, including cultural heritage and the guarantee of worldwide access. We, the undersigned, feel obliged to address the challenges of the Internet as an emerging functional medium for distributing knowledge. Obviously, these developments will be able to significantly modify the nature of scientific publishing as well as the existing system of quality assurance.

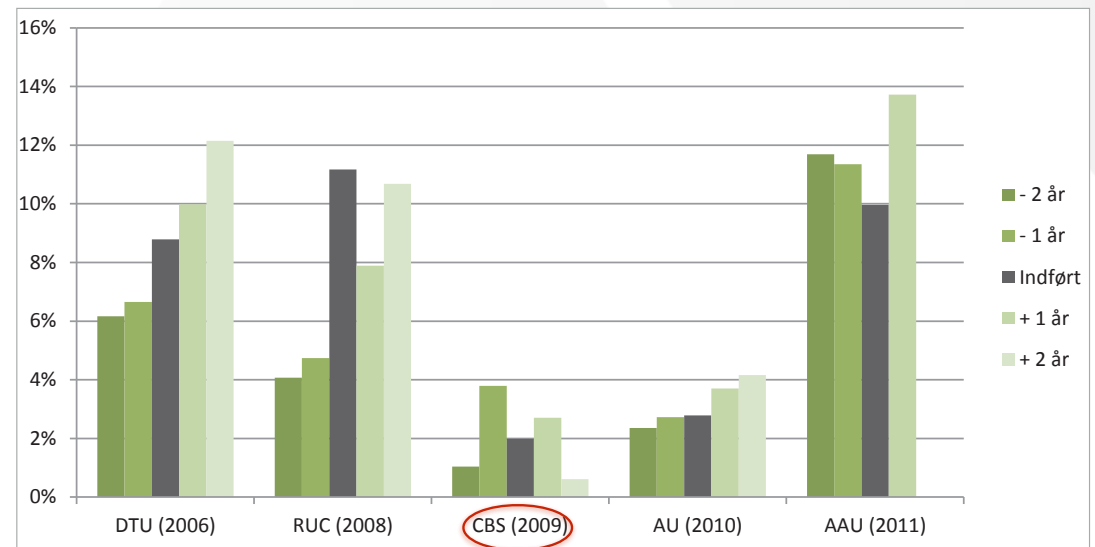
In accordance with the spirit of the Declaration of the Budapest Open Access Initiative, the ECDO Charter and the Bethesda Statement on Open Access Publishing, we have drafted the Berlin Declaration to promote the Internet as a functional instrument for a global scientific knowledge base and human reflection and to specify measures which research policy makers, research institutions, funding agencies, libraries, archives and museums need to consider.



OPEN ACCESS POLITIKKER PÅ DANSKE UNIVERSITETER HAR INGEN EFFEKT



Percentage of articles published as green open access per year



Evolution of the percentage of self-archived articles (green open access) before and after the introduction of open access policy.

Dalsgaard, Søren, and Andreas Larsen. 2014. "Open Access Politikker På Danske Universiteter Har Ingen Effekt." Revy 37 (1): s. 6–8.

Machine Access - interoperability?



The rise of CRIS systems



The screenshot shows a news article on the OpenAIRE website. The article title is "euroCRIS and OpenAIRE work together to enable data exchange with CRIS". It is dated "Updated on 13 September 2013". The article text states: "OpenAIRE and euroCRIS have agreed on the use of CERIF-XML as the standard format by which OpenAIRE will harvest information from CRIS. This was decided during the OpenAIRE Interoperability Workshop at Minho University in Braga, Portugal, held on the 7th and 8th of February." It further explains that a joint working group will work out a dedicated CERIF-XML Application profile. The article also mentions that euroCRIS and OpenAIRE have highlighted this as a substantial step forward in the cooperation of the research information and the repository community, and the use of CERIF-XML as the basis for an exchange format within the research information domain. The article concludes that these results will serve as an essential ingredient for OpenAIRE's data exchange with research information systems (CRIS). The article includes social media sharing buttons for Twitter, LinkedIn, Facebook, and Google+. At the bottom, there is a contact information for Najla Rettberg, OpenAIREplus Scientific Manager, with the email address nrettbe@gwdg.de. The euroCRIS logo is also visible.

OpenAIRE and euroCRIS have agreed on the **use of CERIF-XML as the standard format by which OpenAIRE will harvest information from CRIS.** This was decided during the OpenAIRE Interoperability Workshop at Minho University in Braga, Portugal, held on the 7th and 8th of February,

As a next step, it was decided to facilitate a joint working group, which, in the next few months, **will work out a dedicated CERIF-XML Application profile (including vocabularies) for this.** The profile will be officially presented / launched as a joint cooperation result at the next euroCRIS membership meeting in Bonn.

euroCRIS and OpenAIRE have highlighted this as a substantial step forward in the **cooperation of the research information and the repository community, and the use of CERIF-XML as the basis for an exchange format** within the research information domain.

13 February 2014 - <https://www.openaire.eu/en/component/content/article/9-news-events/447-eurocris-and-openaire-work-together-to-enable-data-exchange-with-cris>

Dekker Letter(s)

The screenshot shows a document from the Tweede Kamer der Staten-Generaal (Dutch House of Representatives). The document is titled "Open Access to publications" and is a parliamentary document from 21-01-2014. The text discusses the principle of open access to research publications, journals, and books free of charge. It states that the principle of open access involves offering readers worldwide access to research publications, journals, and books free of charge. The document also mentions that the principle of open access should be rolled out in the near future and that open access promotes knowledge sharing and knowledge circulation, which in their turn contribute to the Netherlands' innovative capacity. The document is signed by the Secretary of State for Education, Culture and Science, and is dated 18 April 2013.

Open Access to publications
Parliamentary document | 21-01-2014

During the general parliamentary consultations on Science Policy of 18 April 2013, I promised to explain my ideas about the practice of providing open access to research publications and how that practice will continue to evolve. The principle of open access involves offering readers worldwide access to research publications, journals and books free of charge. My contention is that access to the results of publicly and publicly-privately funded research should always be unrestricted. Because such research is paid for from the public purse and technical impediments are essentially non-existent, I believe open access should be rolled out in the near future. Open access promotes knowledge sharing and knowledge circulation, which in their turn contribute to the Netherlands' innovative capacity.

The first steps towards open access were taken twenty years ago, when researchers began sharing their publications with one another on the Internet. In the past ten years, various parties in the Netherlands have been working towards creating an open access system. A wide variety of rules, agreements and options for open access publishing have emerged in the research community. The situation is confusing for authors, readers and publishers alike, and the stakeholders would like this confusion to be resolved as quickly as possible.

It is for that and for the following reasons that I wish to regulate open access:

1. The stakeholders – researchers, universities and publishers – have, for one reason or another, been unable to arrive at a single system for making access to publications arising from publicly and publicly-privately funded research free for everyone at the point of use. The relevant organisations in the Netherlands, including the publishers, have indicated that a fast transition from subscription-based publishing to open access would be beneficial to them.
2. Open access is a cross-border matter and I will therefore be discussing it with my counterparts in various other countries. I will invite them to join me in considering how best to put an international system of open access into

<http://www.government.nl/documents-and-publications/parliamentary-documents/2014/01/21/open-access-to-publications.html> Page 1 of 8

The principle of open access involves offering readers worldwide access to research publications, journals and books free of charge. My contention is that access to the results of publicly and publicly-privately funded research should always be unrestricted.

My preference is “golden” open access; in other words, publication in journals that make research articles available online free of charge. My aim is for the Netherlands to have switched entirely to the golden road to **open access within ten years, in other words by 2024.** In order to achieve this, at least 60 per cent of all articles will have to be available in open access journals in five years' time.

Secretary Dekker - 2013, 2014

(other national mandates: UK, DK, SE...)

Challenges in Measuring Open Access

Open Access in practice

Green OA	full text (draft or published) manuscripts self-archived on the Web (institutional or subject repository)
Gold OA - Journal	OA journals with immediate free access, some of which (e.g. PLoS) operate on an author pays model
Gold OA - Article	author pays the article processing costs (APC) to make articles OA in a subscription based journal (hybrid)
Delayed OA - Green	Publisher embargo period (e.g. 6, 12, 18, or 24 months), after which a published article may self-archived
Delayed OA - Journal	subscription-based journals that convert to OA after a specified period (e.g. 6, 12, 18, or 24 months)
Transient OA	OA during a finite period (e.g. journal promotion); updated or deleted manuscripts in repositories
Restricted OA	access requires registration and/or membership in a group; limited use, such as read-only
Rogue OA	published manuscripts posted on websites or self-archived, but in conflict with licensing/copyrights

adapted from Archambault et al 2014 and Laakso and Björk 2013

Issues with disclosure of OA in bibliometric databases

Web of Science database

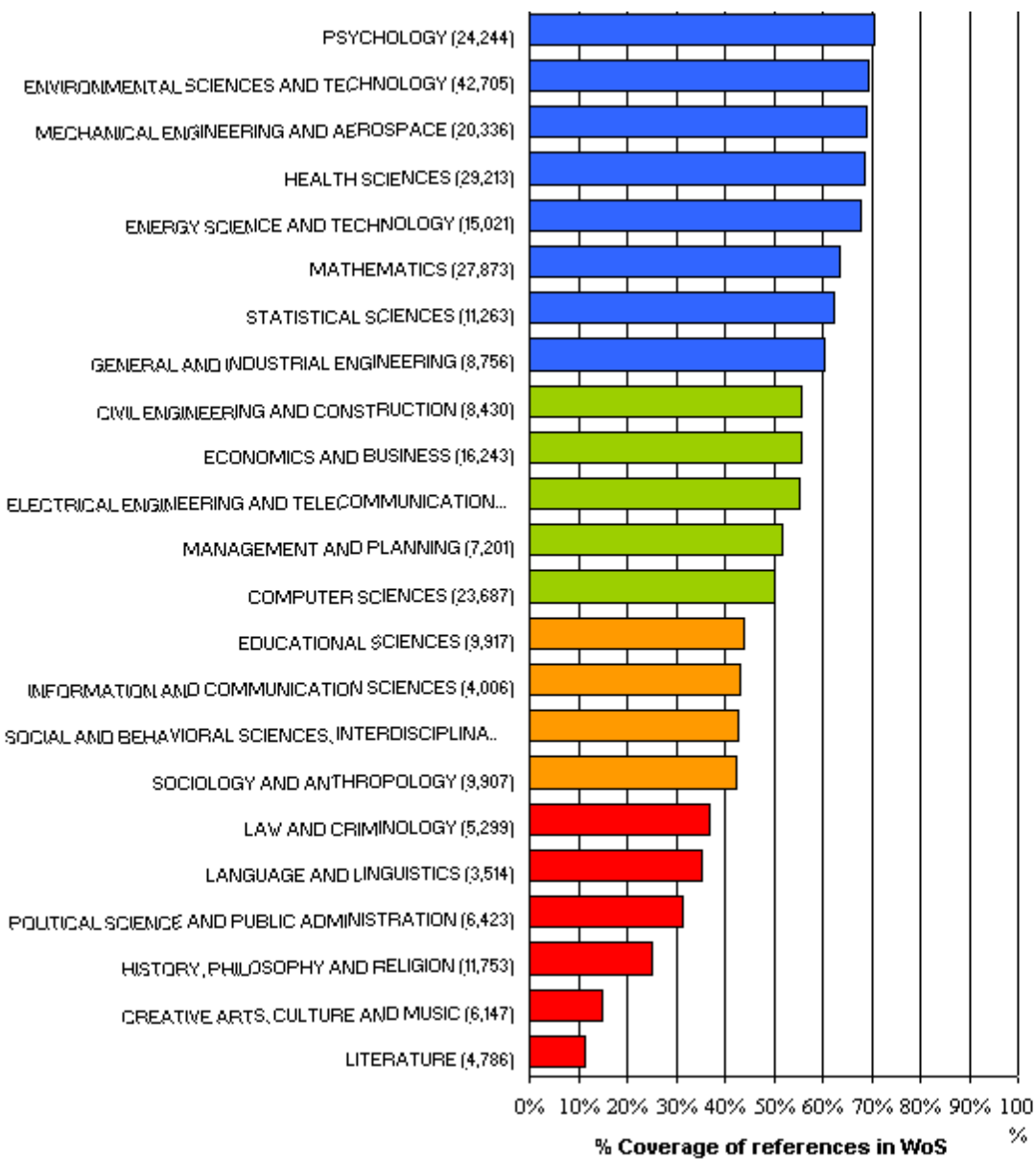
- Desktop version contains OA functionality
- CWTS in-house version of WoS does not have that functionality (yet)

Scopus database

- Desktop version does not contain OA functionality
- Outcomes by Scopus are constructed otherwise ?

Google Scholar

- OA functionality is somewhat problematic here ...
- Other critical issues relate to:
 - Unclear situation around calculation of citation statistics
 - No field normalization available



Coverage of SSH in WoS

International comparison - NL, DK, CH

Measuring the portion of Open Access

harvest method (present state of the art)

- random sample from publication index (e.g. WoS, Scopus, and repositories)
- operational definition: full text of peer reviewed publications, freely available on the internet somewhere
- harvest (find and download) all available full-texts
- analyze available metadata to determine OA type

Measuring the portion of Open Access

harvest method (present state of the art)

- random sample from publication index (e.g. WoS, Scopus, and repositories)
- operational definition: full text of peer reviewed publications, freely available on the internet somewhere
- harvest (find and download) all available full-texts
- analyze available metadata to determine OA type



CRIS method (potential complete dataset)

- local enrichment of commercial data
- registration of output in CRIS systems
- CRIS used with/for evaluation events
- increased granularity of RI
- bottom up aggregation
 - individual
 - department
 - institution

Defining open access in the Web of Science

Method I: By selecting OA output from the WoS desktop interface:

- Link that to the CWTS in-house version of WoS
- Problem here: older papers that are backwardly labeled as OA !

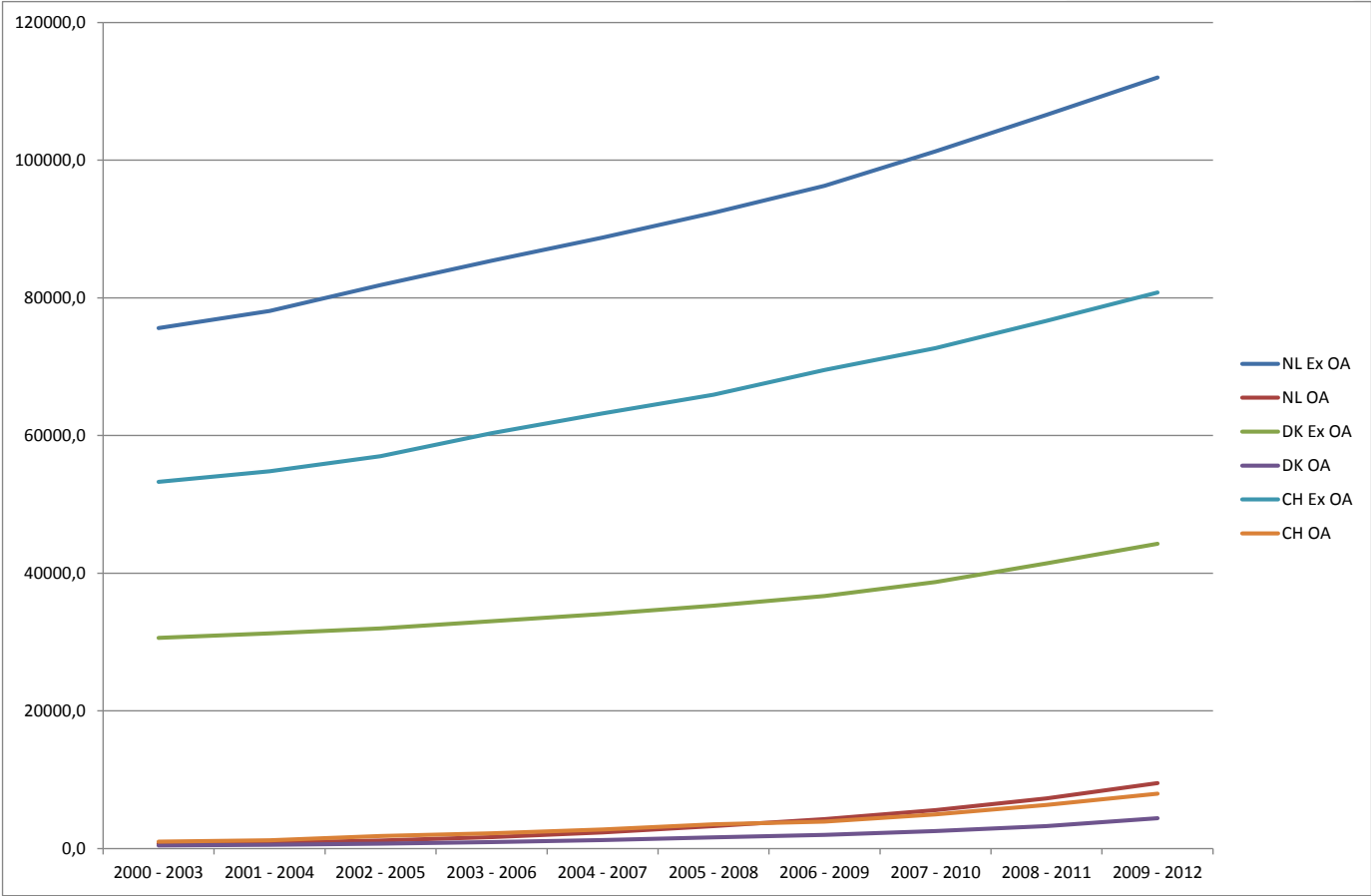
Method II: By linking the DOAJ list with the CWTS WoS database:

- Select the papers from journals on the DOAJ list;
- Use the doi's to link both DOI AJ and WoS to each other.
- Problem here: many older papers in WoS do not carry DOI's yet !

Output, OA and Non OA: NL, DK, CH

	NL Ex OA	NL OA	Share OA	DK Ex OA	DK OA	Share OA	CH Ex OA	CH OA	Share OA
2000 - 2003	75607	712	1%	30616	452	1%	53283	995	2%
2001 - 2004	78087	858	1%	31262	557	2%	54793	1220	2%
2002 - 2005	81849	1180	1%	31972	728	2%	56982	1836	3%
2003 - 2006	85386	1663	2%	33024	949	3%	60319	2217	4%
2004 - 2007	88745	2349	3%	34082	1244	4%	63205	2790	4%
2005 - 2008	92349	3265	4%	35273	1631	5%	65920	3517	5%
2006 - 2009	96278	4269	4%	36672	1997	5%	69518	3912	6%
2007 - 2010	101270	5587	6%	38726	2554	7%	72687	4981	7%
2008 - 2011	106560	7299	7%	41417	3264	8%	76658	6354	8%
2009 - 2012	111990	9504	8%	44264	4420	10%	80786	7990	10%

Output, OA and Non OA: NL, DK, CH

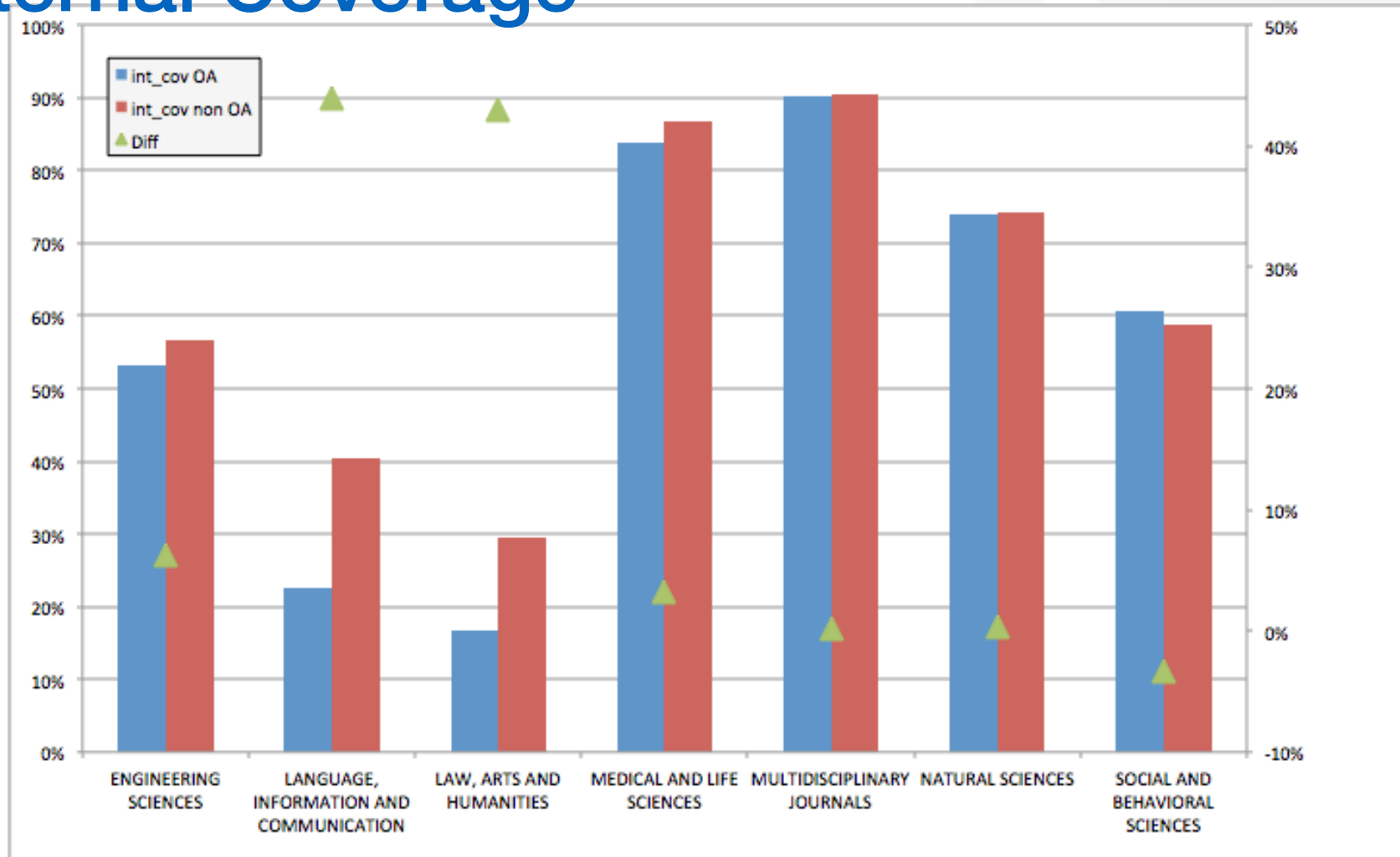


Drilling down: SSH in the Netherlands

Internal Coverage

indicates the degree to which a unit refers themselves to the literature covered in the WoS. The reference behavior of a unit thus indicates whether the journal literature is important for the scholarly communication in a field, and as such can be interpreted as indicator of the applicability of bibliometrics in an assessment context.

Internal Coverage



situated communication practices

organization of work practices

“science [broadly construed to include Humanities] is a professional work organization in so far as it controls how work is carried out, how it is evaluated and its criteria and procedures govern access to material rewards”

Whitley 1986 The intellectual and social organization of science

organization of work practices

Whitley identifies *mutual dependence* and *task uncertainty* as significant for understanding variations in how knowledge is produced.

Practices clustered in relation to mutual dependence (e.g., standardized methods) and in relation to task uncertainty (e.g., the particularity of research objects) have implications for how scientific fields are organized and for how they make use of formal communication venues.

Role of scholarly communication

fields in which *mutual dependence* is dominant, published research is often used to establish the priority of common goals and problems, and thus to facilitate coordination of a research agenda within a field or among related fields. Journal publications often address issues of strategic importance, which rely in part on a stable meaning for the results of research.

example: human genetics

- standardized equipment
- standardized materials
- large collaboration
- often multiple authors
- knowledge claims understood on the basis of standards

Role of scholarly communication

In fields where *task uncertainty* is dominant, the object of research is often situated in a social context. Formulation of problems and goals would be unlikely to find common ground on the basis of published research.

Researchers in fields with uncertain and usually unstable research agendas, correspondingly, have more individual autonomy in facilitating a localized research strategy.

example: ethnography

- descriptive
- situated context
- argument
- output interpreted
- unsuited venue for research agenda

Communication practices

an explanation for low internal coverage

mutual dependence - cited references likely align with standard methods, procedures, and materials, and thus a tighter cluster of intellectual resources

task uncertainty - cited references likely aligned with the particularity of research object and context, and thus a more diverse spread of intellectual resources

Thank you!
questions?

email: c.c.tatum@cwts.leidenuniv.nl