CWTS
Meaningful metrics

What Open Access ?
Various perspectives at the uptake of OA
scholarly publishing

Thed van Leeuwen *, Nicolas Robinson-Garcia **

& Rodrigo Costas *

* Leiden University, Centre for Science and
     Technology Studies (CWTS),
** Delft University of Technology

OAI 2019 Conference, Geneva, Switzerland
June 19-21th, 2019

Universiteit
Leiden

Outline

- **Criteria for analysis of OA uptake**

- **Sources for analysis of OA uptake**

- **How do we define OA uptake ?**

- **Various perspectives on OA**

- **Results of the perspectives analysis**

- **Conclusions & discussion**

SLI.DO code is OAI11

CWTS criteria for OA

❖ CWTS has defined criteria for the analysis of OA uptake *

- **Sustainable**

  – Publications are OA in the public domain, without immediate and direct risk of disappearing behind a pay-wall.

- **Legal**

  – Identification as OA should not be based on 'illegal acts' and should not be based on copyright infringement.

Sources with evidence of OA, compliant with criteria

DOAJ list   ➔ **Gold OA**

ROAD list   ➔ **Gold OA**

CrossRef   ➔ **Green OA**

PubMedCentral   ➔ **Green OA (?)**

OpenAIRE   ➔ **Gold OA & Green OA**

Methodology used until July 2018

Unpaywall   ➔ **Green OA & Gold OA**

Unpaywall   ➔ **Hybrid OA & Bronze OA**

Simplification!

Data sources that do <u>not</u> comply with the two requirements *(Sustainability* and *Legality)* **for OA detection are:**

– *ResearchGate*

– *SciHub*

… **some people still suggest to use these sources !**

CWTS

Additional considerations concerning criteria for OA

❖ An important aspect in opening up: **Engagement**

➢ **Green OA**: by archiving final, peer-reviewed drafts in a freely accessible institutional repository or disciplinary repository

  ➢ **Self-archiving?** *When any of the authors has archived the publication*

  ➢ **Institutional self-archiving?** *When any of the authors or the librarian of the institution has archived the publication*

  ➢ **PMC OA?** *Is it the same as self-archiving? Does it capture **engagement** in OA? However, it takes care of the sustainability dimension of OA publishing*

Data sources: advantages/disadvantages

- **Web of Science**
  - **Advantages**: consolidated database, citation linkages, and complete metadata *(author affiliations and classification scheme available)*
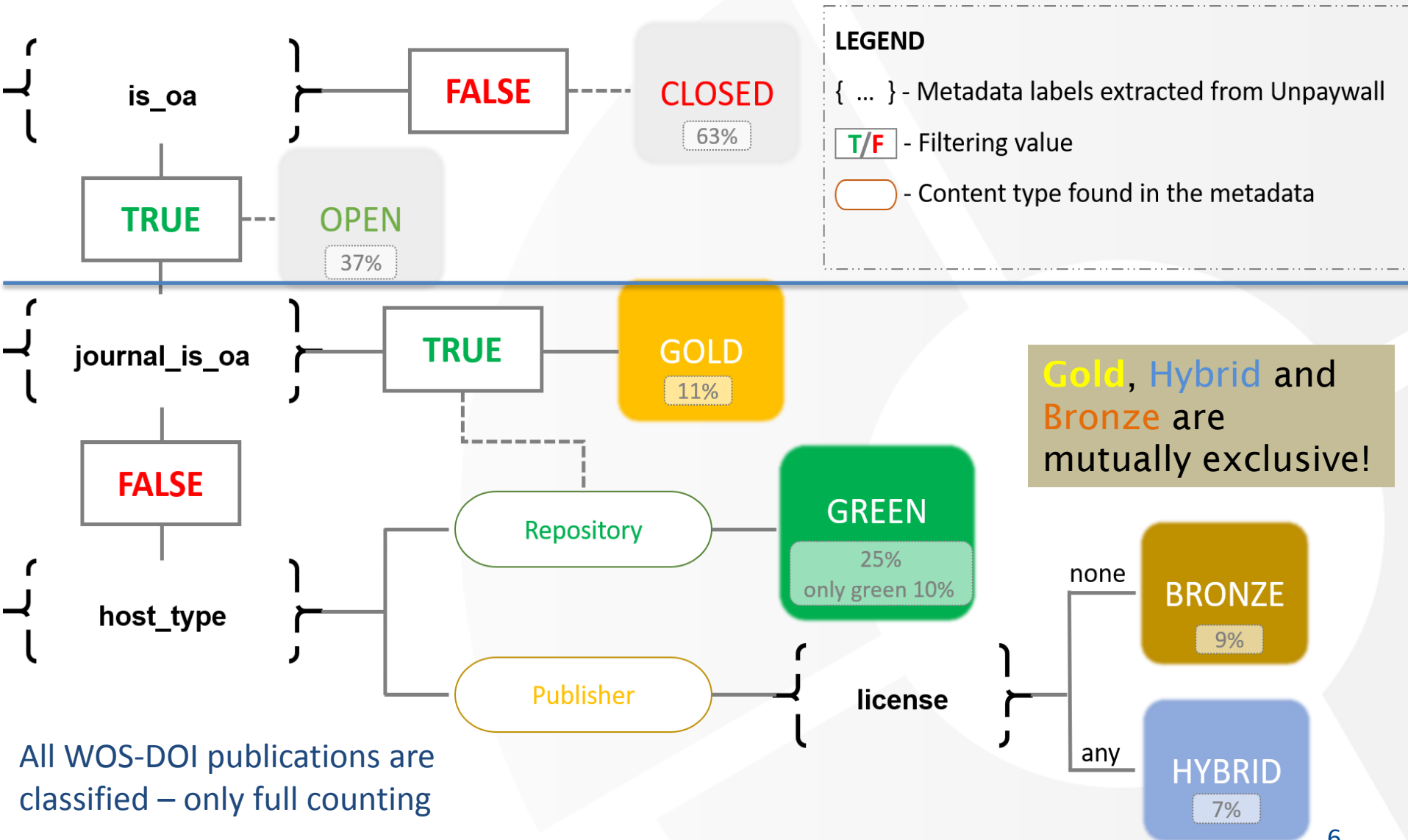 - **Disadvantages**: commercial/proprietary, coverage issues *(SSH, books, conference papers)*

- **Unpaywall**
  - **Advantages**: comprehensive *(multiple sources considered in the identification of OA evidence)*, systematic, large coverage *(Crossref publications)* , 'free' source, becoming 'standard in the business'
  - **Disadvantages**: lack of relevant metadata *(affiliations, classification, doc types)*, dependency of DOIs *(Crossref)*

➢ More validation and research needed *(How good is the data? does Unpaywall track all possible OA evidence?)*

CWTS

How OA?

Approach to OA (from Unpaywall) - all evidence classification diagram *

LEGEND

{ ... } - Metadata labels extracted from Unpaywall

T/F - Filtering value

- Content type found in the metadata

is_oa → FALSE → CLOSED 63%

TRUE → OPEN 37%

journal_is_oa → TRUE → GOLD 11%

FALSE

host_type → Repository → GREEN 25% only green 10%

Publisher → license

none → BRONZE 9%

any → HYBRID 7%

Gold, Hybrid and Bronze are mutually exclusive!

All WOS-DOI publications are classified – only full counting

6

It's all about your perspective …

- **Context counts:** therefore we want to create various points of view on how OA publishing is up taken:
  - **Which perspective?** Readers' perspective? Authors' perspective? Institutions' perspective? Funders' perspective? Publishers' perspective?

  - **Approach**:
    - consider all different forms of OA that a publication can have.
    - In principle, not imposing any preference on any perspective

CWTS

Importance of different 'perspectives'

|	Type of 'preferred' OA by order of importance!			
	1	**2**	**3**	**4**
Readers perspective [access to the closest to published version of the publication]	Gold	Hybrid	Bronze	Green
Authors perspective (Ease) [minimum engagement by the authors in making OA]	Gold	Hybrid	Green*	
Authors perspective (Cost) [minimum cost for the authors in making OA]	Green *	Gold	Hybrid	
Funders perspectives [archiving and access to the closest to published version at a minimum cost]	Gold	Green		
Research policy perspective (Now) [legal access and sustainability at a minimum cost]	Gold	Green	Hybrid	
Research policy perspective (Future) [legal access and sustainability at a minimum cost after Plan S]	Gold	Green		
Institutional/Library perspective [archiving and optimizing access]	Green	Gold		
Publishers perspective [promotion and visibility of publishers publications & activities]	Bronze	Hybrid	Gold	

* Author and Institutional **Green OA** is based on self-archiving only, no PMC Green (for now)
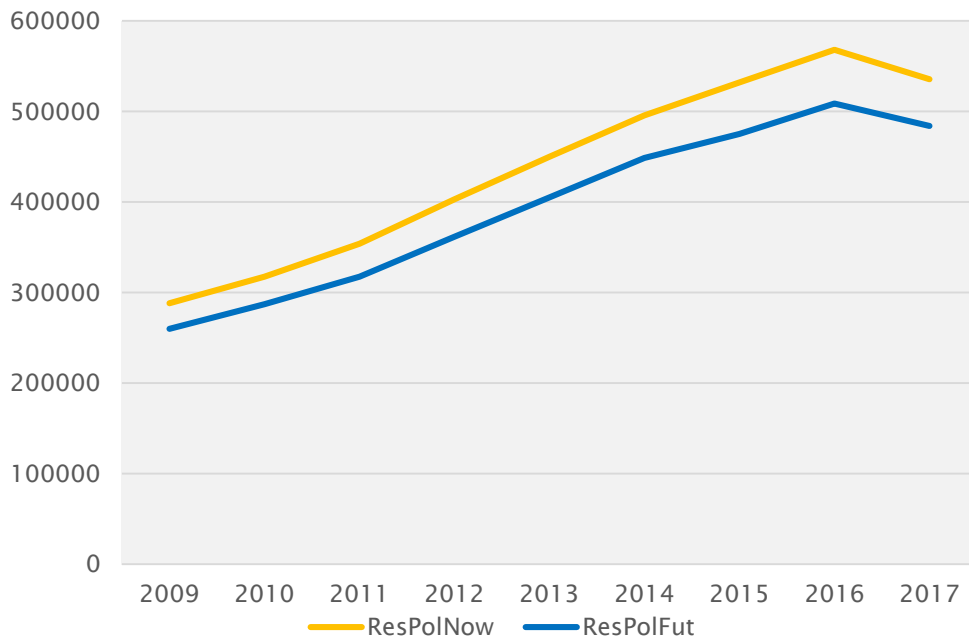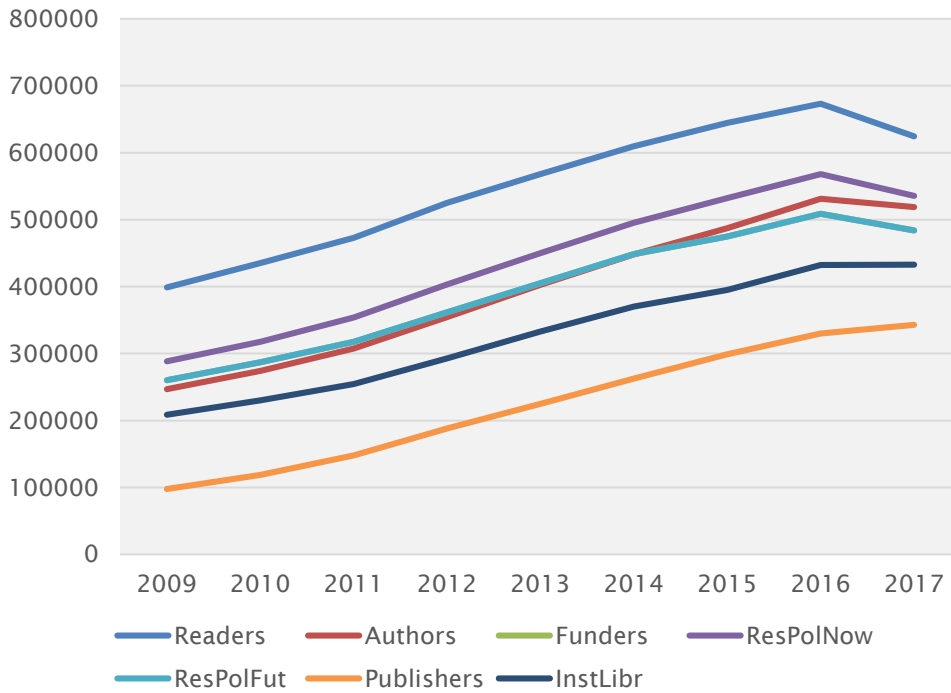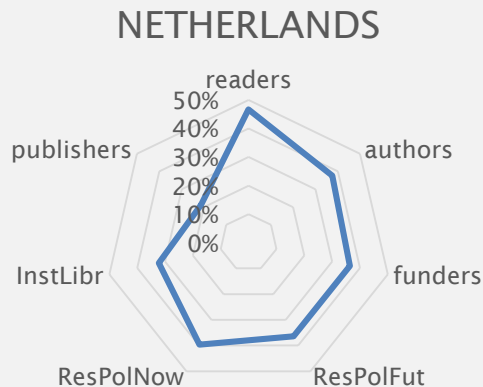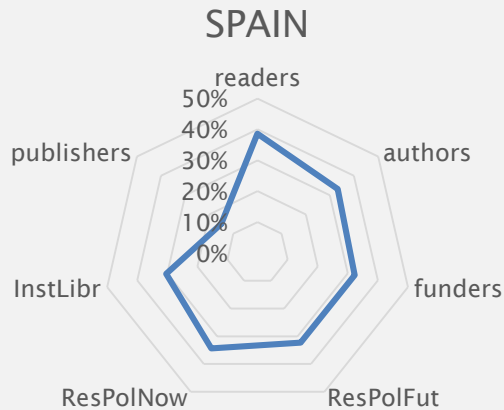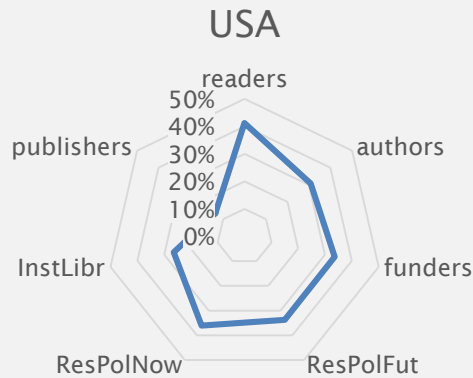
At a global scale

- In absolute terms, we find some 300 k publications difference between Author and Publisher perspectives.

- Some perspectives are overlapping (*Funders*, *ResPolFut* & *Inst/Libr*).

- Plan S decreases OA uptake from the policy perspective.

- Plan S aligns funders and policy making
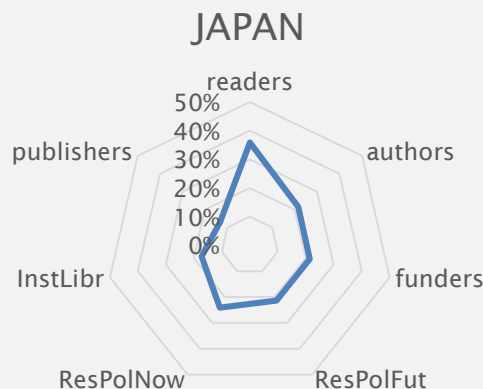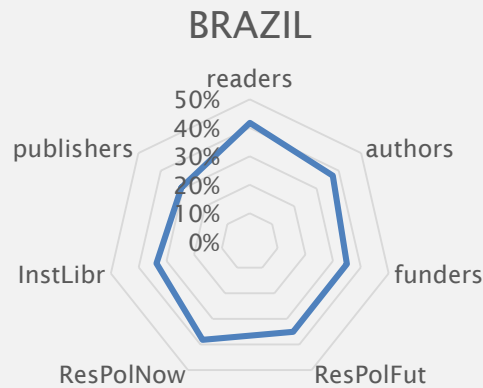
At a global scale, in relative terms
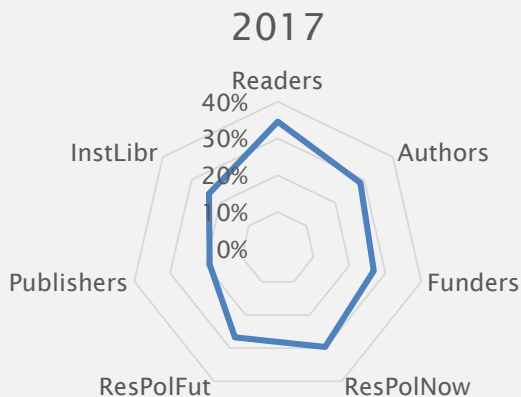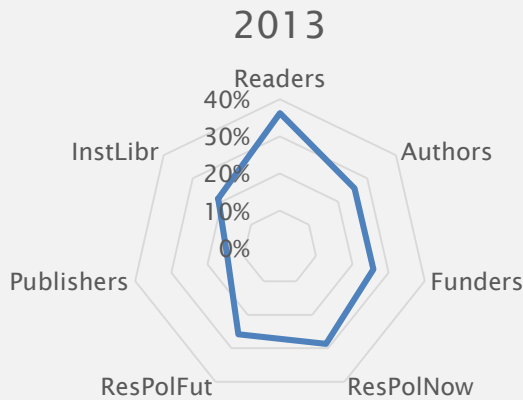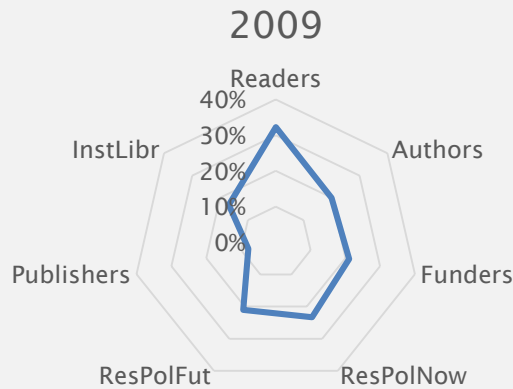
- The surface in the diagram indicates the 'volume' of OA uptake across perspectives.

- Authors perspective is most generous, while Publishers perspective is least generous.

- Through time, we observe that the various perspectives cover more of the surface.

USA — readers, authors, funders, ResPolFut, ResPolNow, InstLibr, publishers (0%–50%)

SPAIN — readers, authors, funders, ResPolFut, ResPolNow, InstLibr, publishers (0%–50%)

NETHERLANDS — readers, authors, funders, ResPolFut, ResPolNow, InstLibr, publishers (0%–50%)

1 - At national level, in relative terms

- The USA displays relatively low visibility through the institutional perspective, as Green OA is less well developed.

- For Spain goes, that the publishers perspective shows less OA uptake, due to a relative small Bronze share.

- For the Netherlands, we see that the Readers perspective on OA uptake is well developed, due to a large Bronze share by publishers.

11

Radar charts: BRAZIL, PEOPLES R CHINA, JAPAN (axes: readers, authors, funders, ResPolFut, ResPolNow, InstLibr, publishers; scale 0%–50%)

2 - At national level, in relative terms

- For Brazil, we notice a balanced representation, even in the Publishers perspective.

- For the PRC, we observe a relatively weak OA uptake throughout all perspectives

- For Japan, Readers and ResPol Now perspectives stand out, due to Bronze in case of the former, and a weakly developed Green OA uptake in Japan in the latter.

At institutional level: LR 2018 data
(>800 universities on a global scale)
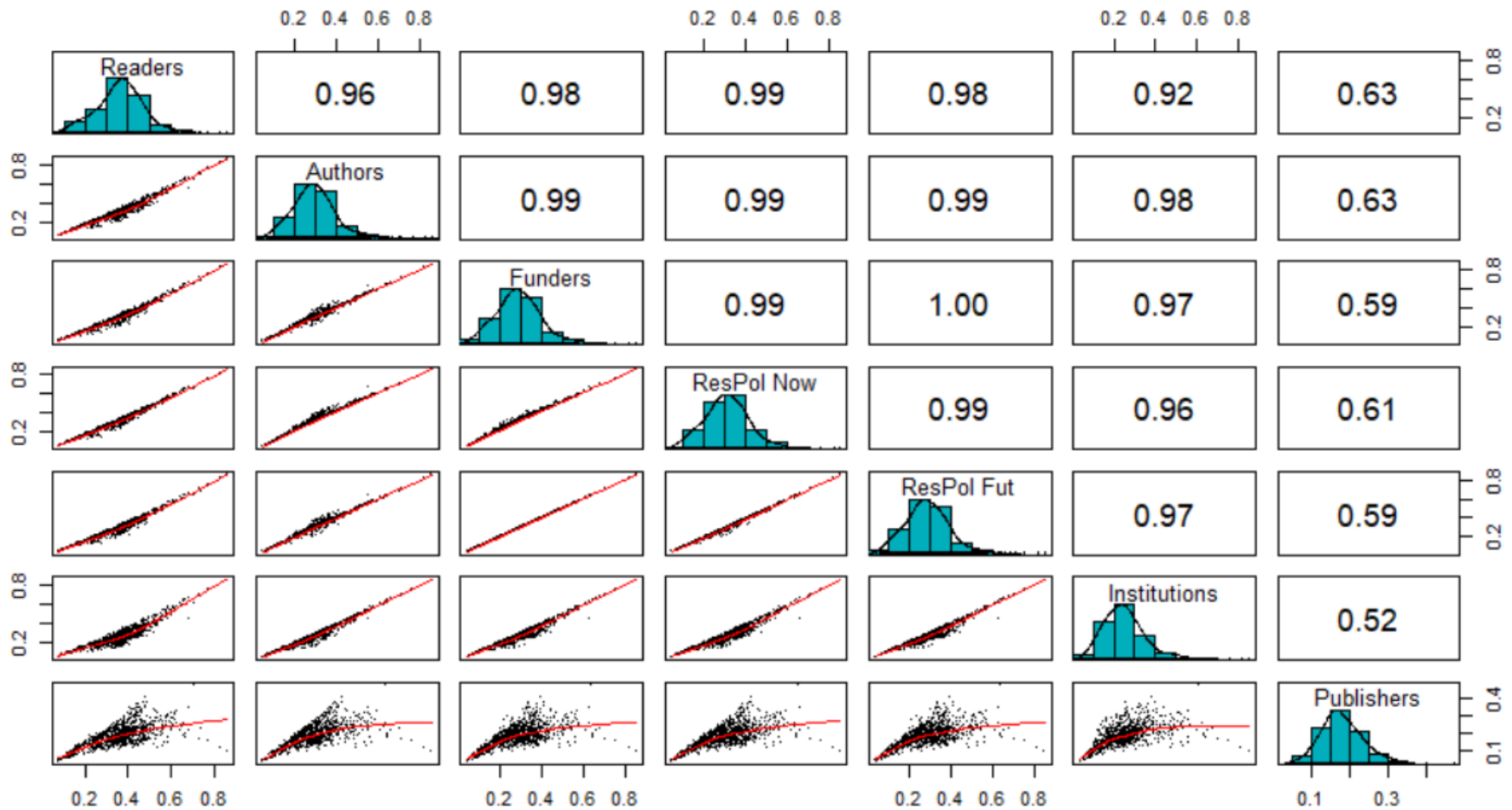
Why use perspectives on OA uptake?

- The use perspectives approach shows that, depending on where one is positioned, OA uptake has a different meaning, and hence, a different volume.

- This also shows that one single superior method of analyzing OA uptake does not exist.

- More practical, the gap between the two extremes (Authors and Publishers) covers some 300k publications

- Most other perspectives fluctuate somewhere in between that 'space'.

- Studies like this can help sorting out concepts and definitions, thereby clarify somewhat the discursive space in which OA is discussed

CWTS

14

What is next ?

- Try to get a more in-depth understanding of the data of Unpaywall.

- Try to get a more in-depth understanding of the archiving in the Green OA variation.

- Is it possible to use the preferences <u>within</u> each use perspective in some way ?

- A further exploration on the Green OA, as a crucial type of OA publishing (as it coincides with the other types)
  - Archiving function in view of the sustainability aspect
  - What is stored in institutional/disciplinary repositories?

Thank you for your attention !

For questions,
ask me now or mail me…

leeuwen@cwts.nl

Many thanks to:
Bianca Kramer & Jeroen Bosman (UU)
Jason Priem & Heather Piwowar (Unpaywall)
for fruitful discussion and clarification