Altmetrics: opportunity or risk?

Nouvelles formes de communication et d'évaluation scientifiques

perspectives et légitimités

5ème journée d'étude du réseau des URFIST

25 septembre 2014

Outline

- Introduction to altmetrics

- Tools and data sources for altmetrics

- What do we know about altmetrics?

- Conceptual aspects

CWTS

Introduction to *Altmetrics*

- *What is the 'impact' of my research?*
  - Limitations peer review and citation analysis

- Recently a new concept has emerged: 'Altmetrics'
  - No definition of altmetrics
  - 'Altmetrics manifesto' (http://altmetrics.org/manifesto/)
  - Altmetrics are seen as metrics about articles (article-level metrics)
  - Altmetrics cover other aspects of the impact of a work (e.g. readerships, mentions in social media and news media, etc.)
  - Possibilities for "societal impact"(?)

- From our perspective altmetrics refers to the mentions of scientific outputs in 'social media' (e.g. Twitter, Facebook) or crowdsourced tools (e.g. Mendeley)

CWTS

Tools and data sources for 'altmetrics'

Analysis of the most important tools for altmetrics [Wouters & Costas, 2012]

Main characteristics and typologies

3 important elements:

- *Scalability*

- *Transparency and data management*

- *Normalization of indicators*

3 main typologies

CWTS

Post-publication peer review tools

Faculty of 1000; *Paper Critic*; *Peer Evaluation*

- Review, comment and rating of publications
- Represent a direct measure of the 'quality'
- Limitations:
  - Conceptual limitations of 'peer review'
  - Most publications don't have any review
  - No controlled systems
  - No normalization possibilities
  - Low predictability of impact [http://arxiv.org/abs/1303.3875]

Web-based citations

Google Scholar; *Google Citations*; *MAS; Arnetminer*

- Collect publications and citations from the web
- Limitations:
  - Traditional limitations of citations
  - Coverage & update not always clear
  - Data management & download not possible
  - Difficult scalability
  - No normalization of indicators
  - Low level of data standardization

CWTS

Altmetrics

Mendeley; *Altmetric.com*; *ImpactStory; PLoSONE*

- Altmetrics: readerships, tweets, FB shares, comments, rates, blogging, etc.
- Limitations:
  - No clear meaning of these metrics
  - Manipulability
  - Difficult scalability and data collection (although API's are available)
  - No normalization of indicators
  - Low level of data standardization
  - *Low correlation with citations!!*

What do we know about altmetrics?

- New emerging research line
  - Research teams in US, UK, Canada, Spain, Finland, Denmark, Germany, the Netherlands, etc.

  - Main topics discussed so far:
    - Coverage
    - Correlations (mainly with citations and other bib. elements)
    - Data problems and inconsistencies
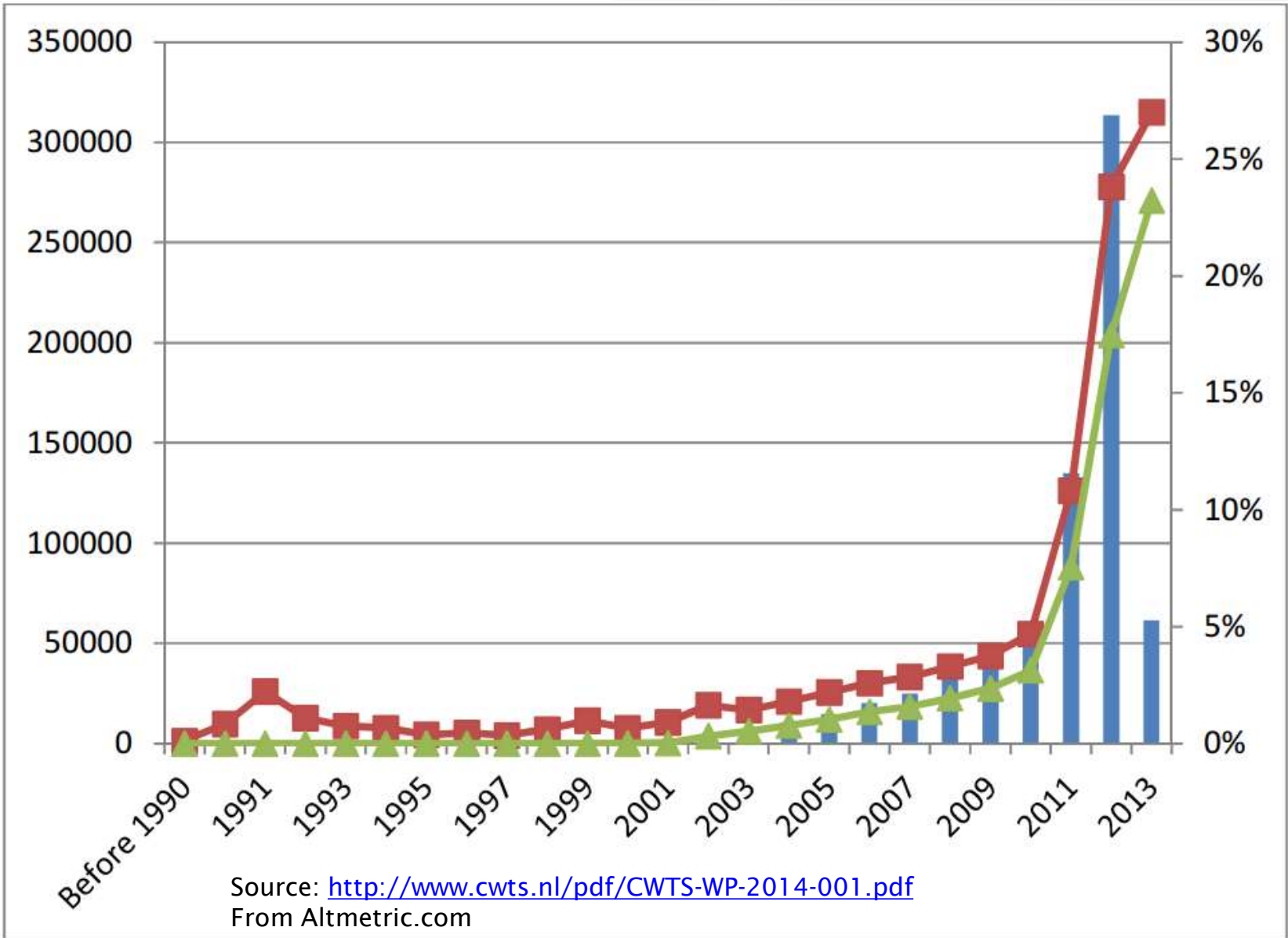   - Content & meaning

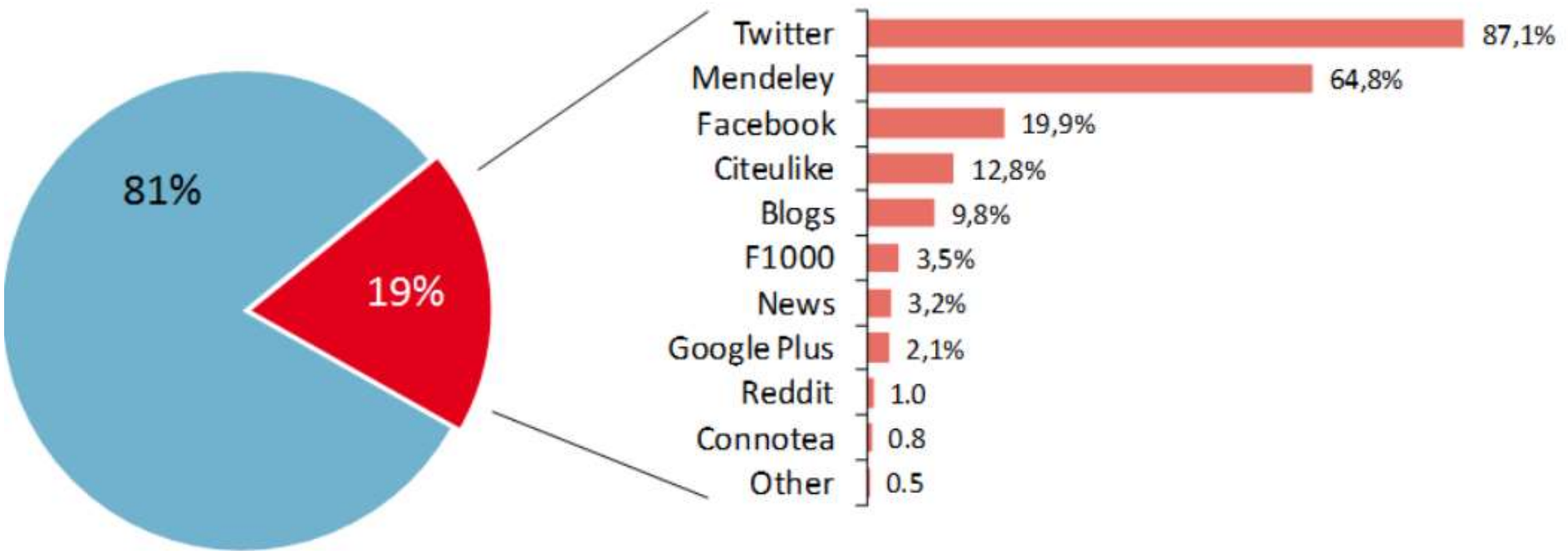 - Still many *unanswered questions*!

CWTS

Coverage:

- Zahedi, Costas & Wouters (2013)

- Random sample (20,000 pubs) from Web of Science

- Period **2005-2011**

- Metrics from 'Impact Story'

Data Source	papers with metrics	%
Mendeley readers	12362	**62.7**
CiteULike bookmarks	1638	**8.3**
Wikipedia Mentions	289	1.4
Topsy Tweets	265	1.3
Facebook likes	142	0.7
Delicious bookmarks	72	0.3
Topsy influential tweets	59	0.3
Facebook shares	57	0.3
Facebook comments	42	0.2
Facebook clicks	16	0.01
PlosAlm_pmc_full_text	1	0
PlosAlm _pmc_abstract	1	0
PlosAlm_pubmed_central	1	0
PlosAlm _pmc_pdf	1	0
PlosAlm_pmc_supp_data	1	0
PlosAlm _pmc_unique_ip	1	0
PlosAlm _pmc_figure	1	0
PlosAlm _html_views	1	0
PlosAlm _pdf_views	1	0
PlosAlm _scopus	1	0
PlosAlm _crossref	1	0
PlosAlm	1	0

Increasing presence of altmetrics

Source: http://www.cwts.nl/pdf/CWTS-WP-2014-001.pdf
From Altmetric.com

Altmetric.com coverage

Twitter	87,1%
Mendeley	64,8%
Facebook	19,9%
Citeulike	12,8%
Blogs	9,8%
F1000	3,5%
News	3,2%
Google Plus	2,1%
Reddit	1.0
Connotea	0.8
Other	0.5

81%

19%

WoS publications with DOI not included in Altmetric.com

WoS publications with DOI included in Altmetric.com

CWTS

Source: Robison-García et al, 2014
(http://elprofesionaldelainformacion.metapress.com/link.asp?id=a734234878233531)

10

Correlations

- Moderate correlations for **Mendeley** [Mohammadi et al, 2014; Zahedi et al, 2013]

- Lower for other altmetrics [Haustein et al, 2013; Costas et al, 2014]

- Also low correlations with other bibliographic elements: pages, title length, n. co-authors, n. references, etc.
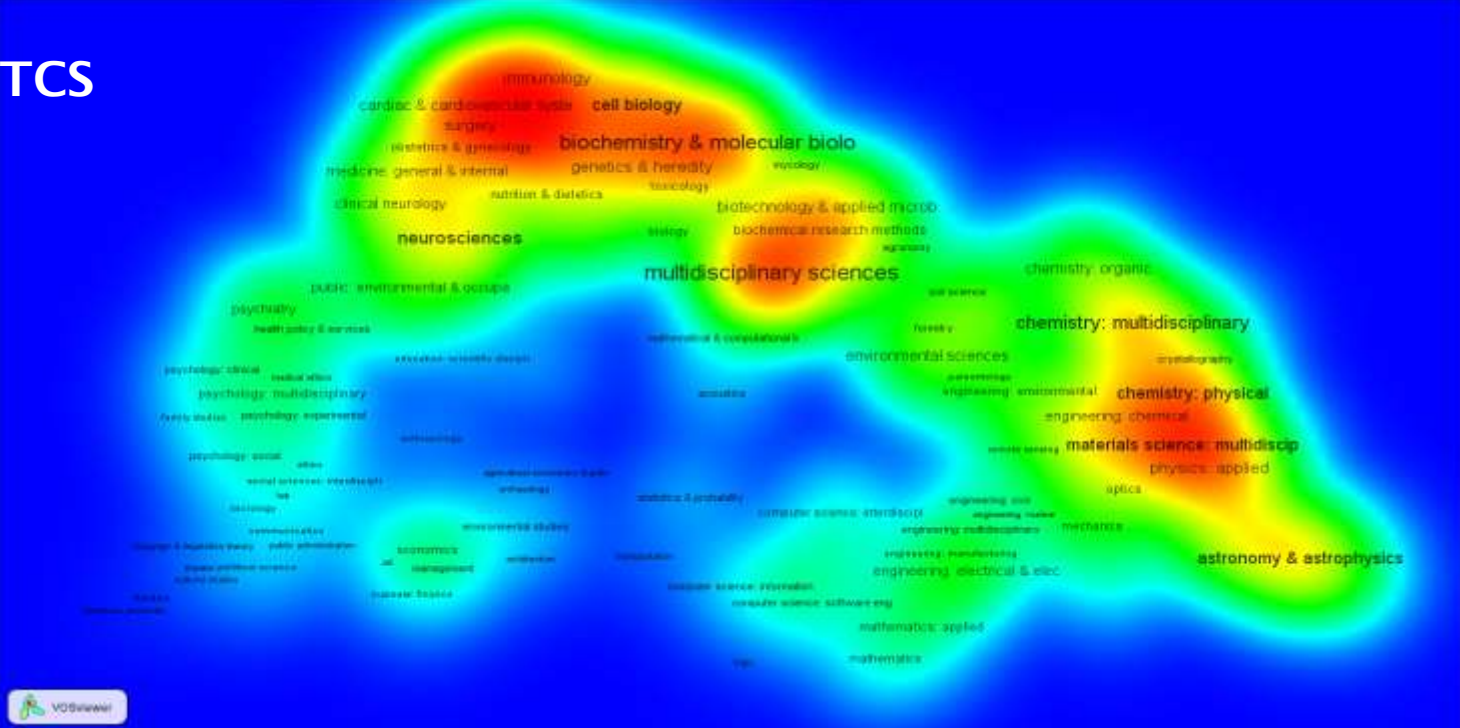
CWTS

Data problems

- Inconsistencies across data sources
  - Differences in counts between Altmetric sources (PLoS ONE, Altmetric.com & Mendeley) controlling by time variables

    [Zahedi, Fenner & Costas, 2014]
  - Between Mendeley itself at different points in time [Bar-Ilan, 2014]

- Manipulation and validity
  - Easy manipulation of Google Scholar [Delgado López-Cozar, Robinson-García, Torres-Salinas, 2013]
  - Also possible with other Altmetric sources (e.g. Twitter, Facebook)!
  - Presence of 'bots' and 'cyborgs' in Twitter [Haustein et al, 2014]
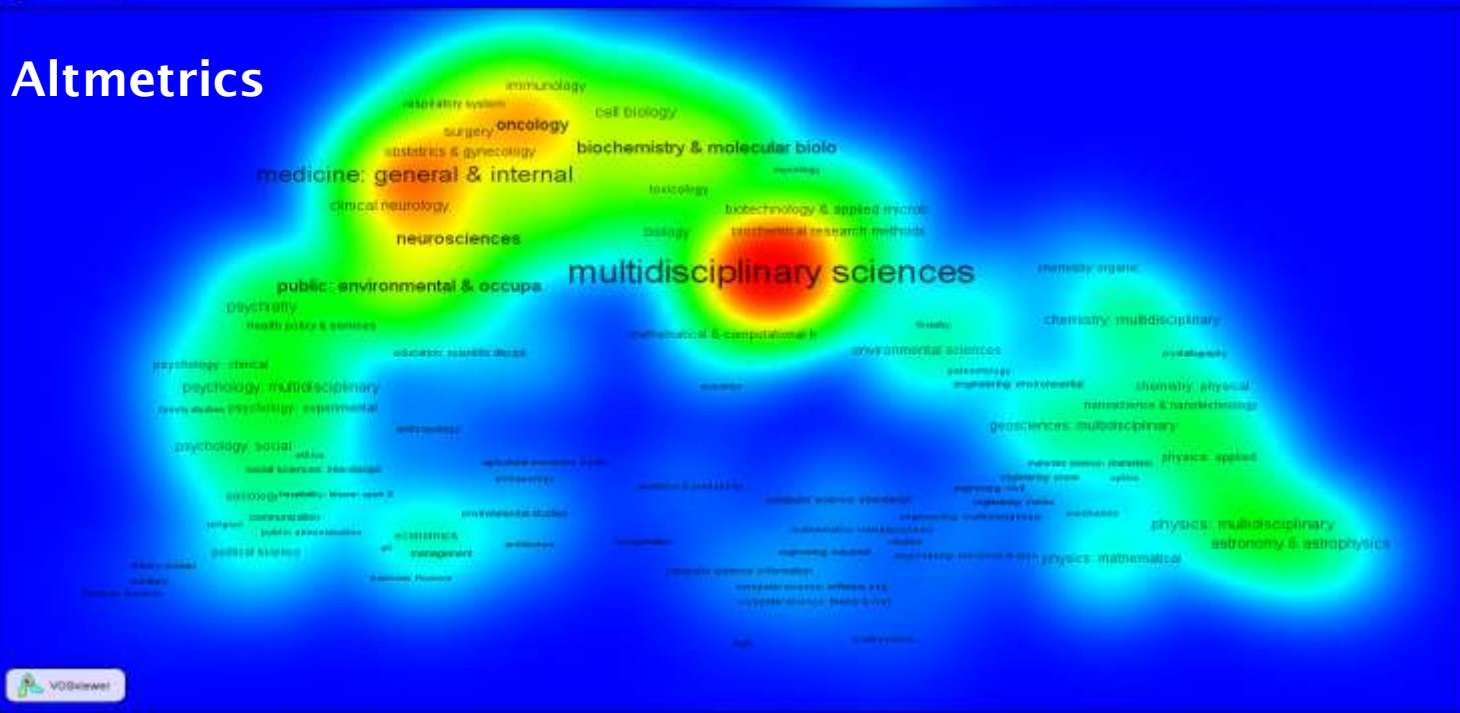
CWTS

Content

- Analysis of the titles of publications with altmetrics [Costas et al, 2014]

- 500,229 WoS publications, citations up to 2012 and altmetrics up to October 2013

- Degree of 'citedness' or 'altmetricness'(*Altmetric.com*) by
  - Disciplines (Subject Categories)
  - Topics (terms in the titles)

- VOS viewer (www.vosviewer.com)

CWTS

TCS

Altmetrics

TCS

Altmetrics

Conceptual aspects, challenges & possibilities of altmetrics

Four conceptual aspects ('promises')

- Diversity (of scientific publishing)
  - *Strongest argument*
  - *More research is still necessary*

- Broadness of scientific performance (new dimensions)
  - *Do we agree with and understand these new dimensions?*
  - *What does a 'tweet' or a 'Mendeley reader ' mean from a performance point of view?*

- Speed
  - *Is faster always better? Superficiality?*
  - *'Sleeping beauties' or 'Mendel syndrome'*

- Openness
  - *Openness is good (easy & free)*
  - *Transparency and consistency are more important*

CWTS

Challenges

They must be used and accepted by the scientific community!

- Conceptualization of these new (alt)metrics
- Standardization of tools and data
- Solution to problems in data quality and indicator construction
  - Care with data is important!
- Scalability
  - Need of using the APIs (*and still slow!*)
- Manipulability
- Normalization & reliability
- Production and use of the new metrics
- Potential trivialization due to their limited use ("narcissism")

Possibilities of altmetrics

- Most recent publications

- Complement to bibliometric indicators

- Possibilities for libraries: picking up highly discussed publications, suggesting relevant social media users, finding relevant blogs about a topic, etc.

- Support for assessment of some fields (e.g. Soc. Sciences & Humanities)

- Other types of impact:

  - Detection of papers with different types of impact not covered by citations (social/cultural impact, educative impact, professional impact)

  - Controversial papers? Controversial topics? Socially sensitive topics?

Thank you very much for your attention!

Questions?

CWTS