Relevance and Challenges of Altmetrics for Repositories – answers from the *metrics project

Maryam Mehrazar, Jan Weiland, Wolfgang Riese (ZBW), Julius Stropel (VZG), Astrid Orth (SUB)

Presentation at OR19, 13 June 2019, Hamburg
Outline

1. What can Altmetrics do for Repositories?
2. Answers from *metrics project
3. EconStor as a use case
1. What can Altmetrics do for Repositories?

**Bibliometrics:**
- **How often do users interact with a scientific output online?**

**Altmetrics/usage metrics (= web-based metrics):**
- **Timing:** availability very early after publication of scientific results
- **Access:** greater openness, transparency, wider re-use than costly citation databases
- **Reach:** broader impact (disciplines, geographies, and societal groups)
- **Objects:** go beyond articles, measure relevance of software, data, other formats
Reliability & Perception of *metrics

- How reliable are *metrics?
  - How are researchers using social media?
  - Are all platforms used similarly?
  - Are there differences in usage of functions?
  - What does this mean for the construction of metrics?
- How are *metrics perceived?
  - What do researchers think about them?
  - Are they understanding and using them?
- Practical challenges when crawling for *metrics data
2. *metrics Project

- DFG-funded
  01/2017 – 05/2019

- Partner
  - State and University Library Göttingen (SUB),
  - Leibniz-Informationszentrum Wirtschaft (ZBW)
  - Headoffice Gemeinsamer Bibliotheksverbund (VZG)
  - Leibniz Institute for the Social Sciences (GESIS)
Differences in use of platforms

Choice of platforms depending on career level

Survey of ~3400 participants
Patterns in use of functions

58 actions (used by at least 150 survey respondents)

Actions used more frequently by early-stage researchers:

Actions used more frequently by professors:

Writing  Commenting  Downloading  Liking  Sharing  Bookmarking  Other
FINDINGS I - Usage of social media platforms

- Users of different career stages:
  - Select different platforms for communication
  - Use different functions on the same platform

BEHAVIOUR ON SOCIAL MEDIA PLATFORMS IS COMPLEX AND DIVERSE WHICH PROHIBITS SIMPLE AGGREGATIONS
## Perception of *metrics

### Usefulness of different types of metrics for researchers

<table>
<thead>
<tr>
<th>Metric</th>
<th>Very useful</th>
<th>Useful</th>
<th>No answer/Don't know</th>
<th>Hard to use</th>
<th>Useless</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citation number of the scientific output</td>
<td>38%</td>
<td>48%</td>
<td>2%</td>
<td>9%</td>
<td>1%</td>
</tr>
<tr>
<td>The Impact Factor of the journal in which the output is published</td>
<td>38%</td>
<td>48%</td>
<td>2%</td>
<td>9%</td>
<td>1%</td>
</tr>
<tr>
<td>H-index of the author</td>
<td>16%</td>
<td>41%</td>
<td>17%</td>
<td>16%</td>
<td>11%</td>
</tr>
<tr>
<td>Number of downloads of the scientific output</td>
<td>14%</td>
<td>48%</td>
<td>4%</td>
<td>24%</td>
<td>10%</td>
</tr>
<tr>
<td>Number of mentions of the scientific output in media/news</td>
<td>9%</td>
<td>37%</td>
<td>6%</td>
<td>32%</td>
<td>17%</td>
</tr>
<tr>
<td>ResearchGate score of the author</td>
<td>8%</td>
<td>19%</td>
<td>25%</td>
<td>13%</td>
<td>18%</td>
</tr>
<tr>
<td>Number of followers/subscribers of the author</td>
<td>21%</td>
<td>14%</td>
<td>27%</td>
<td>25%</td>
<td>18%</td>
</tr>
<tr>
<td>Alternate attention score of the scientific output</td>
<td>21%</td>
<td>38%</td>
<td>21%</td>
<td>18%</td>
<td>16%</td>
</tr>
<tr>
<td>Number of views/visits of the scientific output’s page</td>
<td>23%</td>
<td>31%</td>
<td>29%</td>
<td>24%</td>
<td>16%</td>
</tr>
<tr>
<td>Number of posts mentioning the author/scientific output on social media</td>
<td>22%</td>
<td>14%</td>
<td>32%</td>
<td>30%</td>
<td>16%</td>
</tr>
<tr>
<td>Number of shares/retweets/ forwards of such posts</td>
<td>24%</td>
<td>13%</td>
<td>28%</td>
<td>31%</td>
<td>18%</td>
</tr>
<tr>
<td>Number of comments/replies to such posts</td>
<td>19%</td>
<td>12%</td>
<td>29%</td>
<td>33%</td>
<td>16%</td>
</tr>
<tr>
<td>Number of Likes/Favorites on such posts</td>
<td>13%</td>
<td>12%</td>
<td>25%</td>
<td>43%</td>
<td>47%</td>
</tr>
<tr>
<td>Number of online bookmarks (e.g. on Mendeley) of the scientific output</td>
<td>23%</td>
<td>4%</td>
<td>31%</td>
<td>22%</td>
<td>25%</td>
</tr>
</tbody>
</table>

- **Bibliometric indicators**
- **Alternative indicators**
Application of *metrics

You are doing literature research for a topic you are not yet familiar with. Your query in the scholarly search engine of your choice reveals 3 potentially relevant publications alongside their impact metrics. Please rank those publications in the order in which you would read them by dragging them to the area on the right. The publication you would read first should afterwards be at the top of the list, the publication you would read last at the bottom.

Ranking of publications based on *metrics

Which indicator do you find most useful? n~250
FINDINGS II - Use and concerns towards metrics

- Alternative metrics are:
  - Not yet widely known among researchers
  - Used with considerable caution

HELP RESEARCHERS BECOME METRIC-WISE
*metrician – Gathering information from social media platforms

- Repositories
  - GoeScholar, EconStor, SSOAR
  - ~ 225k works
- Gathering methods
  - DOI, Handle
  - Landing page URL
  - Metadata
- Access
  - Software: [https://github.com/gbv/metrics-crawler](https://github.com/gbv/metrics-crawler)
  - Data dumps (ask us)
FINDINGS III - Technical Challenges

- Occurrences on platforms difficult to identify:
  - Different habits of identifying papers
  - Inconsistent or no use of identifier systems

QUALITY VARIES WITH DATA SOURCES
RECOMMENDATIONS

- Check for applicability
  - Select platforms depending on user behaviour
  - Weight functions according to exhibited patterns
- Avoid simple one-dimensional metrics
  - Check aggregations for their validity
- To gain trust and acceptance by researchers
  - Provide context and greater openness
  - Explain application areas, strengths and limitations
- Improve results from data sources
  - Develop standardizations in altmetric „citations“
  - Promote use of identifiers
More information

- *metrics in transition workshop
  - [https://metrics-project.net/en/events/workshop2019/](https://metrics-project.net/en/events/workshop2019/)

- DINI recommendations paper
  - To be published soon
  - Leave your contact details to be notified
as a Use Case
Overview EconStor

- Disciplinary repository for economics and business administration
- Mainly „secondary use“ of publications via
  - written agreements: complete series & journals
  - open content licenses: CC licensed OA journals
- Overall content: >175,000 full texts
  - Working papers (60%)
  - Journal articles (25%)
- Focus on dissemination services: RePEc & Google Scholar
- „Classical“ metrics available:
  - Citation data: CitEc (RePEc service)
  - Download statistics: COUNTER rules applied
- Software: DSpace 5
Initial Situation

- Predominantly pure DOI analysis in altmetrics services
- EconStor content:
  - Predominantly working & discussion papers (= preprints)
    - DOI assignment not common for preprints!
    - EconStor assigns Handle.net identifiers (prefix: 10419)
- Possible solution:
  - Including alternative (persistent) identifier systems in altmetrics!
    - Find mentions containing URLs like
      - hdl.handle.net/10419/*
      - econstor.eu/handle/10419/*
      - econstor.eu/bitstream/10419/*
Integration of *metrics Numbers

- Basis: Static SQL data dump from the *metrics project
  - Back end: Implemented as JSON REST web service
  - Front end: DSpace JSPUI
- Evaluation period:
  - March 2018 to January 2019
- *metrics data on EconStor landing pages visible for
  - Mendeley
  - Twitter
  - Wikipedia
- No differentiation according to subgroups like tweets/ retweets
- Backlinking to the corresponding Tweets & Wikipedia sites
Naturalization and citizenship: Who benefits?

Gathmann, Christina

2015

IZA World of Labor | ISSN: 2054-8571 | Year: 2015 | Issue: 125

Politicians, the media, and the public express concern that many immigrants fail to integrate economically. Research shows that the option to naturalize has considerable economic benefits for eligible immigrants, even in countries with a tradition of restrictive policies. First-generation immigrants who are naturalized have higher earnings and more stable jobs. The gains from citizenship are particularly apparent among immigrants from poorer countries. A key policy question is whether naturalization causes labor market success or is taken up by those immigrants who would anyway be most likely to succeed in the labor market.
Results (1): Mendeley

- Number of mentioned EconStor titles relatively high!
  - ~ 25,000 titles
- But: Due to a lack of interfaces an additional title matching used
  - Methodically problematic in economics, as working paper version and published version often have an identical title!
  - Backlinking to Mendeley has been waived
Results (2): Twitter

- Number of mentions very low!
  - ~ 460 titles overall
  - Only ~30 EconStor titles identified directly through a matching of Handle.net identifiers
  - ~430 EconStor titles identified through DOI matching
    - 60% originate from a single journal!
    - Primary publication platforms are more active in social media to promote their content!
Concluding Remarks

- *metrics project user studies confirmed:
  - Social media tools not widely used in economics!
- Primary publication platforms more likely used for referencing
- Collection of data from social media platforms still very „tricky“:
  - Inaccurate usage of persistent identifiers
  - Lack of interfaces
  - Entries might be deleted (Tweets) or changed (Wikipedia) over time!
- Very low dissemination effects through EconStor!
  - Referrer analysis: Only 0.2% downloads via Twitter sites
Thank you / Vielen Dank!

- Web [https://metrics-project.net/](https://metrics-project.net/)
- Email [metrics-project@sub.uni-goettingen.de](mailto:metrics-project@sub.uni-goettingen.de)
- Twitter [@metrics_project](https://twitter.com/metrics_project)
- Facebook [@metricsproject](https://www.facebook.com/metricsproject)