

Measuring Scholarly Output

Amy Seipke, Engineering Librarian
Jeneen Jadallah, CIS and Mathematics Student
University of Michigan - Dearborn

Topics



1. Need
1. Problem
1. Hypothesis
1. Data Collection
Decisions
1. Results
1. Limitations
1. Future steps

1. Need

- University:
 - evaluate research output
 - student recruitment
 - inter-university comparisons
 - grants and funding sources

- Faculty:
 - simple and time efficient tracking of research output
 - eliminate redundant or unnecessary work

- Library:
 - get the best data
 - maximize efficiency

2. Problem

- Multiple systems of tracking research output within UM-Dearborn.
 - creates faculty overwhelm
 - poor levels of follow through results in bad data
 - not comparable across units or with other universities.

- Platforms being used at UM-Dearborn
 - Faculty 180
 - M Portfolio
 - Web of Science
 - SciVal
 - Altmetrics
 - Google Scholar
 - ORCiD
 - Dimensions
 - Michigan Experts

3. Background

- Scopus/Web of Science:
 - Articles only
 - Only those published in journals in that database/subscription
- Google Scholar:
 - Anything can be added by the user
 - Scrapes from ?
- Altmetrics:
 - Posts on social media
 - News outlets
 - Blogs
 - Academic forums
 - Policy docs
 - Citation management saves (Mendeley)

- ORCID:
 - Anything it can scrape with your ID on it
 - Scopus
 - Web of Science
 - Deep Blue
- Michigan Experts:
 - Anything it scrapes
 - Scopus
 - Web of Science
 - ORCID
 - Deep Blue
 - Anything the author adds

4. Hypotheses

1. The humanities are underrepresented in these systems

2. Michigan Experts (ME) will provide some mitigation of this field bias

-individuals can correct their scholarship record by inputting more correct information

3. Updating profiles in ME will have an impact on the faculty impact scores on Dimensions, ORCID, and Google Scholar.

-this will not have any measurable impact on the impact scores in Scopus or Web of Science

5. Methodology/Decisions

Research Methodology Overview

- Comparative analysis across 6 platforms
- Focus: accuracy, consistency, and coverage
- Mixed-method approach:
 - Quantitative → citation counts
 - Qualitative → missing/duplicate data

Platforms Analyzed:

- Web of Science (journal-focused)
- Scopus (citation-based analytics)
- Google Scholar (broad + user-controlled)
- Altmetrics (social/media impact)
- ORCID (ID-based aggregation)
- Michigan Experts (institutional system)

Faculty Sampling Strategy

- **Population:**
 - 222 faculty (full + senior associate professors)
- **Stratified Sampling Approach:**
 - Top 2 faculty (highest reported output)
 - Bottom 2 faculty (lowest reported output)
 - Remaining sample selected randomly
- **Why this approach?**
 - Capture extremes (overrepresented vs underrepresented)
 - Avoid bias from only high-performing faculty
 - Ensure representative sample across colleges
- **Final Sample Distribution:**
 - CEHHS: 5
 - COB: 9
 - CECS: 19
 - CASL: 38
- **Statistical Confidence:**
 - 95% confidence level $\pm 10\%$ margin of error

Methodology/Decisions 2

Data Collection Workflow

- Searched each faculty member across all platforms
- Recorded:
 - Publications (articles, books, etc.)
 - Author ID
 - Citation counts
 - Platform presence
- Disambiguation:
 - Removed confusion with UM-Ann Arbor faculty
 - Checked duplicate profiles

Data Organization & Tracking

- Spreadsheet structure:
 - Rows = individual scholarly works
 - Columns = platforms
- Binary tracking:
 - Present / Not present
- Additional fields:
 - Citation count
 - Duplicate flags
 - Time tracking, because its a very time intensive process

Analysis Approach

- Cross-platform comparison:
 - Which works appear where
- Citation comparison:
 - Compared platform values to individual averages
- Field-level aggregation:
 - Grouped by department
- Error analysis:
 - Missing items
 - Duplicate counts

Methodology/Decisions 3

Key Methodological Decisions & Justifications

Inclusion Decisions

- Included faculty with **zero reported outputs**
→ captures underrepresentation and missing data
- Included **self-edited profiles (flagged)**
→ analyzed separately to assess impact of manual updates

Exclusion Decisions

- Excluded **peer reviews (Web of Science)**
→ not standardized or comparable across platforms
- Excluded **Mendeley saves (Altmetrics)**
→ reflects intent to read, not actual impact

Metric Standardization

- Altmetrics included:
 - Citations (Dimensions)
 - Social/media mentions
- Focused on **comparable indicators of impact across platforms**

Challenges in Data Collection

- Name ambiguity across platforms
- Duplicate profiles
- Missing publications
- Platform-specific biases
- Time-intensive manual validation

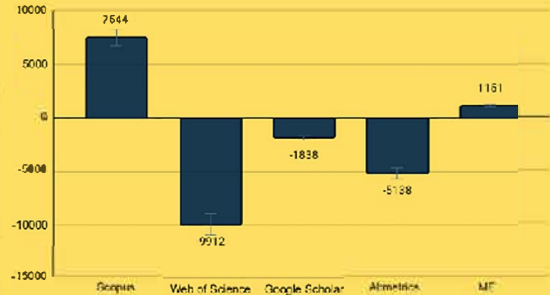
6. Results

What have we been able to analyze so far:

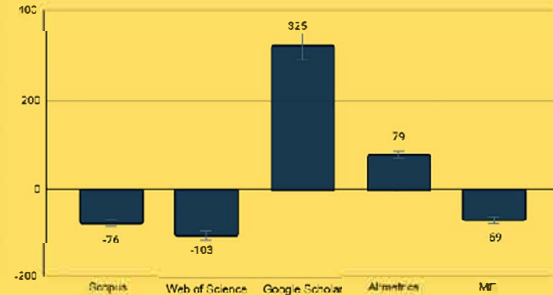
- We found the average number of citations across all the platforms for each faculty member
- Then we compared the number of citations reported in each platform to the average across all the platforms for each faculty member
- Then we collected and added up all those differences for the individual faculty members within each field of study.
- These graphs show the total of the individual differences from the individual average reported by each platform and grouped by field of study.

Results 2

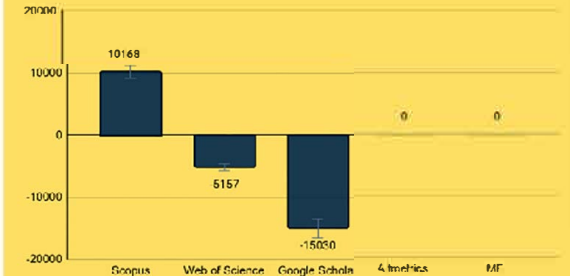
COB Average of individual differences from their mean



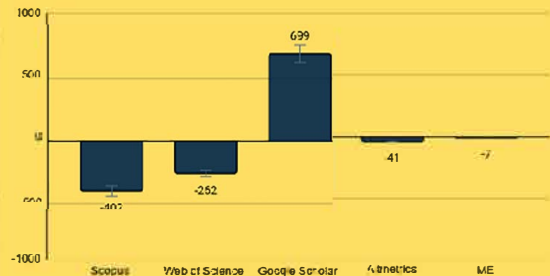
CEHHS Average of individual differences from their mean



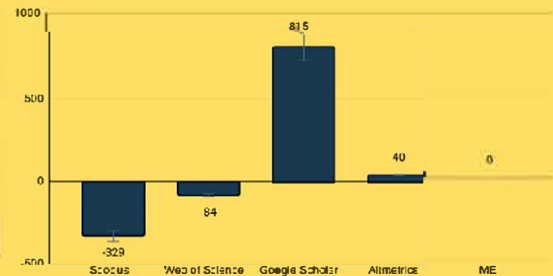
Language, Arts & Culture: Average of individual differences from their mean



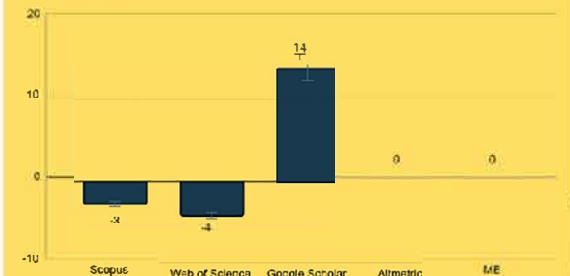
Natural Science: Average of individual differences from their mean



Social Science: Average of individual differences from their mean



Mathematics & Statistics: Average of individual differences from their mean



Results 3

What our preliminary results show::

- **Language, Arts, and Culture** as well as **College of Business** are have a much higher than average impact when measured in *Scopus*
- All the other departments (**College of Engineering, Health Human Services, Natural Science, and Social Science**) have a much higher than average impact when measured in *Google Scholar*
- All the fields are under reported in Web of Science but not equally so (COB is much lower than the others)
- **Social Sciences and Health and Human Services** are the only fields that have a higher than average representation in *Altmetrics*
- For all departments Michigan Experts shows the citation count that is nearest to the average.

6. Limitations

- Only people at UM can use Michigan Experts (but everyone can use ORCID and Google Scholar so that part can be transferable)
- We selected a smaller sample size than we originally wanted to. We settled on having a CI of 95% +/- 10% because to have a CI 95% +/- 5% we would have had to double our sample/double our work.

- We calculated these results compared to the average number of citations for each individuals as they appeared with minimal data cleaning...this is not necessarily the most accurate measurement of the impact
 - We plan two steps to improve this:
 1. We will clean our data for all
 2. We will get a smaller sample of CVs to get a more accurate “true” measure of which items should be included

7. Future steps

Part 1:

- Finish our sample (we didn't get all the way through CASL or CECS)
- Clean our data better (remove false hits and deduplicate the redundant articles)

Part 2:

- Collect CVs from a small number of faculty and manually edit their ME profiles for them.
- Run these calculations again using the “true” number of citations as the baseline (instead of the average like we did here)

Part 3:

- Run the method again on those faculty with the edited profiles to see how that process impacts their representation in the other platforms esp. Google Scholar and ORCID

Feel free to reach out with any questions or suggestions



Amy Seipke:

aseipke@umich.edu

Jeneen Jadallah:

jeneenj@umich.edu

Mardigian Library:

library.umd.umich.edu

Platforms Used:

Clarivate. *Web of Science*. (2025). Clarivate. Accessed: Apr. 14, 2026. [Online]. Available: <https://www.webofscience.com>

Digital Science & Research Solutions Ltd. *Altmetrics Explorer*. (2026). Digital Science & Research Solutions Ltd. Accessed: Apr. 14, 2026. [Online]. Available: <https://www.altmetric.com/explorer>

Digital Science & Research Solutions Ltd. *Michigan Experts*. (2026). Symplectic Elements. Accessed Apr. 14, 2026. [Online]. Available: <https://experts.umich.edu>

Elsevier. *Scopus*. (2026). Elsevier. Accessed: Apr.14, 2026. [Online]. Available: <https://www.scopus.com>

Google LLC. *Google Scholar*. (2024). Google LLC. Accessed:Apr. 14, 2026. [Online]. Available: <https://scholar.google.com>

OpenAI. *ChatGPT-5.3 Instant*. (2026). OpenAI. Accessed: Apr. 14, 2026. [Online]. Available: <https://chatgpt.com>

ORCID, Inc. *ORCID*.(2026). ORCID, Inc. Accessed: Apr. 14, 2026. [Online]. Available: <https://orcid.org>