Evaluating the Impact of Community-Engaged Scholarship: Implications for Promotion and Tenure

ICHE Faculty Leadership Conference

February 15, 2019

Heather Coates, MLS, MS
IUPUI University Library Center for Digital Scholarship

- What are the core concepts of research evaluation, bibliometrics, and altmetrics?
- Are metrics fit for purpose for evaluating research?
- Are metrics fit for purpose for evaluating researchers?
- What are some common mistakes in using metrics?
- How can we use metrics responsibly?
- How can P&T processes reward community engaged research?

What are the core concepts of research evaluation?

activities – actions that people can take on scholarly products

direct measurement – observable actions, effects, etc.

indirect measurement – gathering information from other sources

indicators – a measure that must be sufficiently proven to correlate with the associated concept

normalization – how data are put into appropriate context for the evaluation purpose

What is rewarded in the P&T process?

Productivity

Quality

Prestige

Reputation

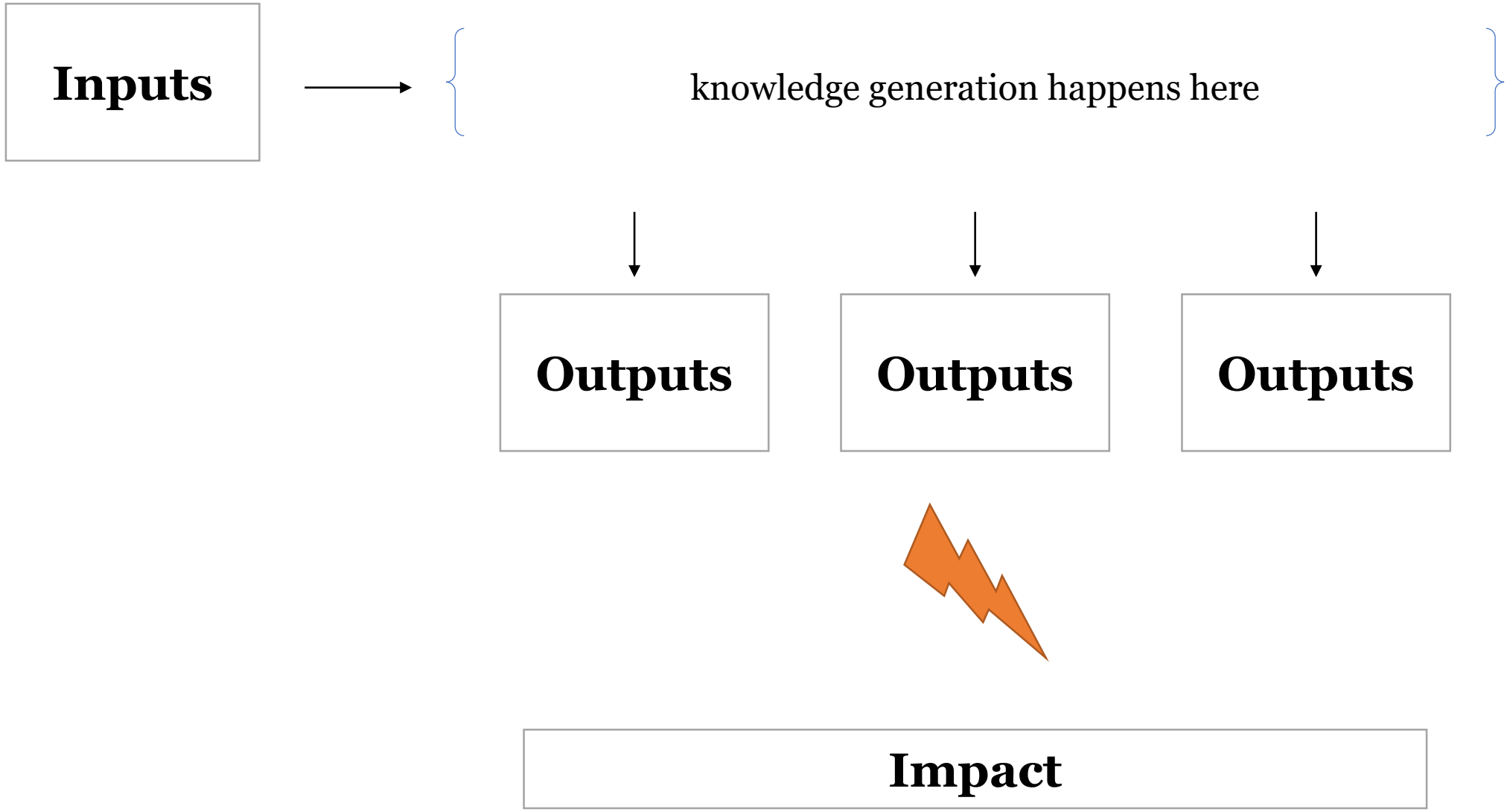
Impact

Professional growth

What is rewarded in the P&T process?

Count of publications, typically journal articles	Productivity
“High impact” journals, “top tier” presses	Quality
Exclusivity, longevity, imprimatur, eminent editors and/or reviewers	Prestige
Word of mouth, perception, being established	Reputation
Journal Impact Factor, citation counts per article	Impact
General increase or upward trajectory in quantity and or quality of work over time	Professional growth

High-level mental model for research



What are the core concepts of bibliometrics?

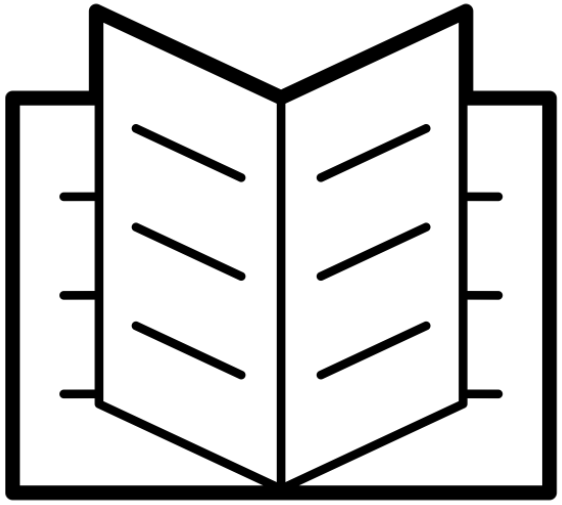
Bibliometrics is the statistical analysis of publications such as articles, books, and other sources. *It is not the quantitative measurement of all scientific activities.*

Units of measurement: inputs, outputs

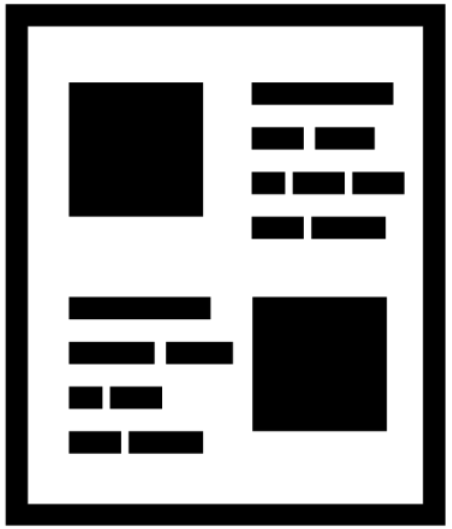
- inputs include people, instruments, money, space, etc.
- outputs include articles, books, data, code, models, algorithms, etc.

Concepts: use, visibility, impact, quality

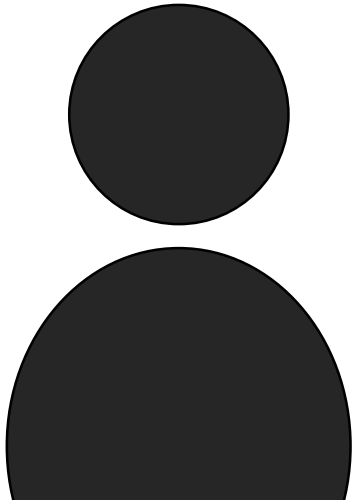
What is the metric about?



Journal/Venue Level Metrics



Output/Article Level Metrics



Author Level Metrics

What are the core concepts of bibliometrics?

Bibliometric indicators are generally derived from a core set of inputs and outputs

- Publications per author
- Publications in a field over time
- Citations per article
- Citations patterns between journals
- Co-authorship patterns between countries

What are the core concepts of altmetrics?

Altmetrics are indicators for activities taking place on social media platforms, such as Twitter, Facebook, CiteULike; typically considered to be complementary to bibliometrics, rather than a replacement

Activities include view/download, discuss, save, cite, recommend

By default, these are item-level metrics (about the article, book, data, etc.)

Another mental model

- Journal articles
- Books
- Book chapters
- Conference presentations & papers
- Abstracts
- White papers
- Reports
- Data
- Code, algorithms
- Models, simulations

- View
- Download
- Bookmark
- Like
- Share
- Discuss
- Engage
- Review
- Cite
- Adopt/Implement

- Citation counts
- Views & downloads
- h-index
- Relative Citation Ratio
- Media coverage
- Reviews
- Eigenfactor Article Influence Scored
- Altmetric Attention Score

Are metrics fit for purpose for evaluating research?

Is the metric a valid indicator for the variable being evaluated?

Example: Journal Impact Factor (JIF)

In any given year, the impact factor of a journal is the number of citations received in that year by articles published in that journal during the two preceding years, divided by the total number of [research, proceedings, and review] articles published in that journal during the two preceding years.

$$\frac{\text{\# of citations to } \textbf{all items} \text{ published in 2011-2012}}{\text{\# of } \textbf{articles} \text{ published in 2011-2012}}$$

Note: the JIF is not a good predictor of whether an individual article will be highly cited, which is how many people interpret it

Are metrics fit for purpose for evaluating research?

Is the metric a valid indicator for the variable being evaluated?

Example: raw citation counts

- Coverage for journal articles is much greater than for other types of products
- Raw counts are meaningless without context – what is the primary research product of the field, what is the citation half-life for the field

Example: h-index

- Typically available only for journal articles, favoring articles over other products
- Favors more senior scholars – increases with time
- Favors productivity over quality, but not measure either well
- Does not behave predictably – does not necessarily increase as citations or publications increase

Total Publications

9,609 Analyze

1989 2008

h-index

153

Average citations per item

30.85

Sum of Times Cited

296,439

Without self citations

294,993

Citing articles

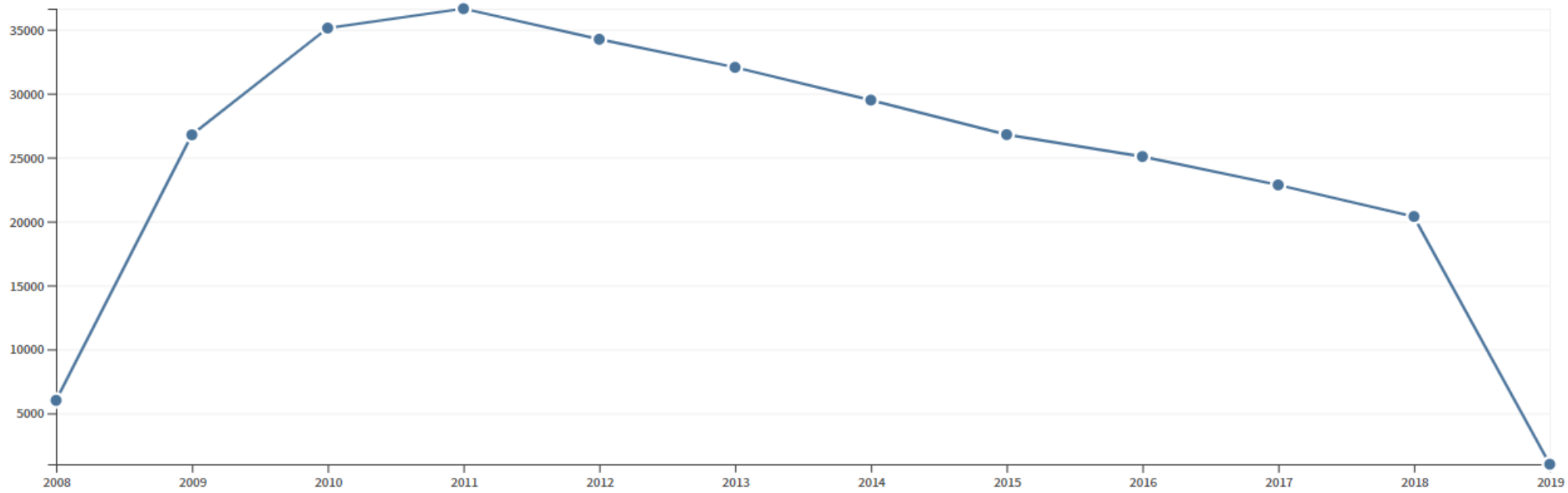
205,466 Analyze

Without self citations

204,489 Analyze

Sum of Times Cited per Year

Subject = Infectious Diseases



Total Publications

6,544 Analyze

1989 2008

h-index

52

Average citations per item

4.39

Sum of Times Cited

28,757

Without self citations

28,592

Citing articles

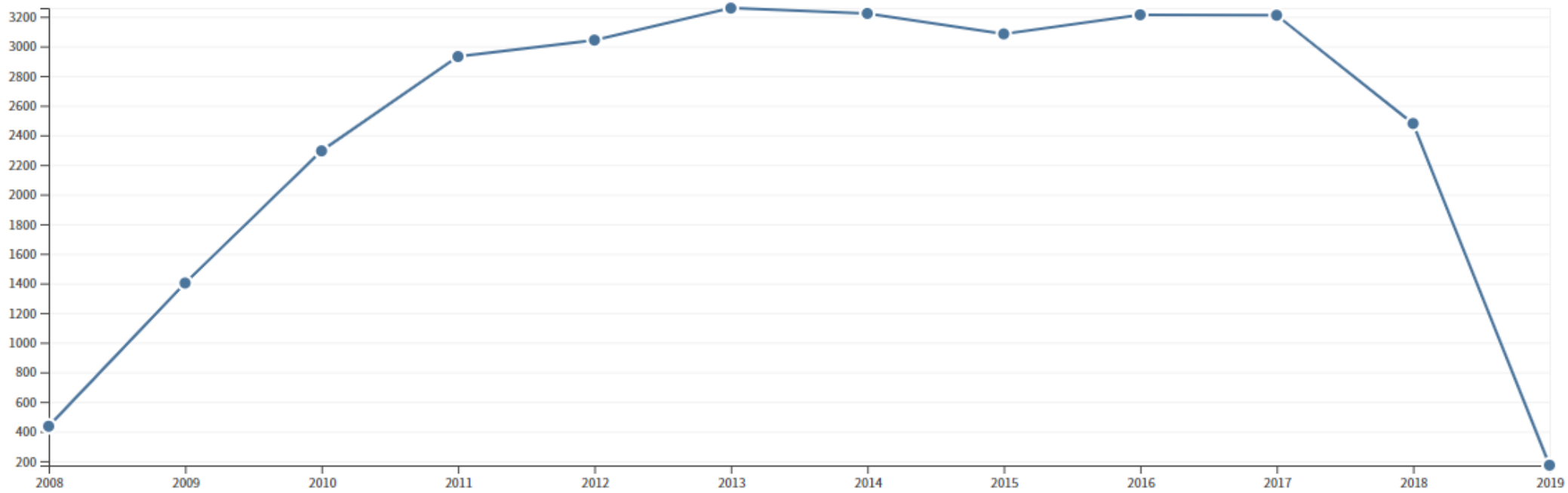
25,133 Analyze

Without self citations

25,017 Analyze

Sum of Times Cited per Year

Subject = History



Total Publications

3,007 Analyze

1989 2008

h-index

94

Average citations per item

19.49

Sum of Times Cited

58,618

Without self citations

58,378

Citing articles

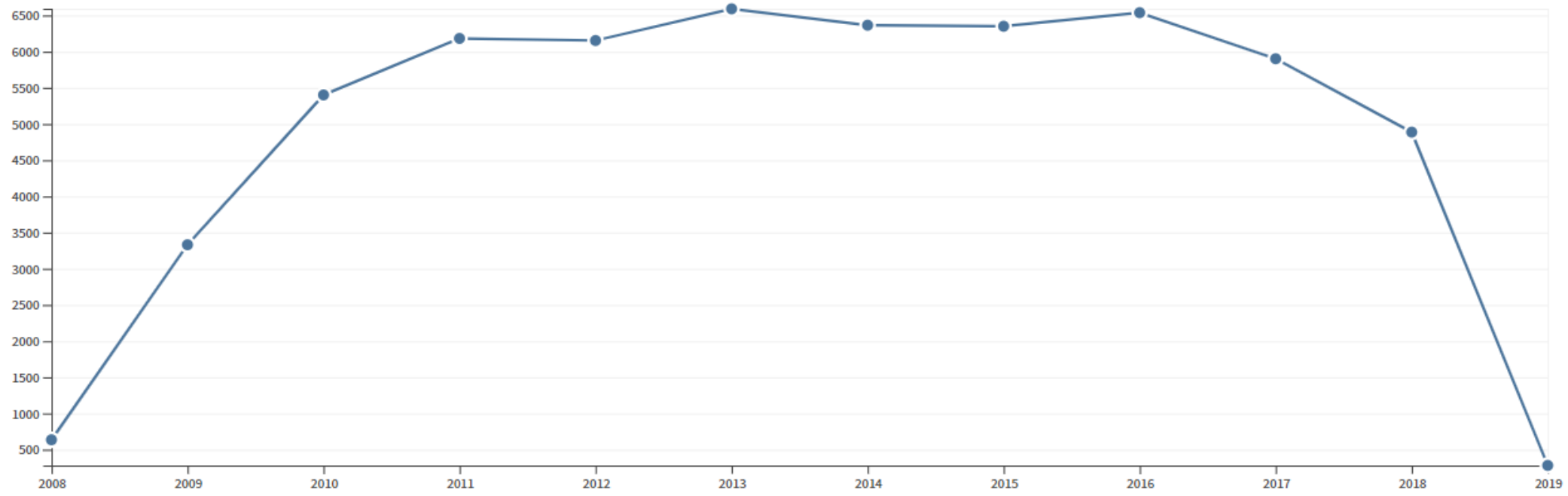
44,742 Analyze

Without self citations

44,603 Analyze

Sum of Times Cited per Year

Subject = Information Science & Library Science



Total Publications

325 Analyze

2008 articles = 24,149

h-index

270

Average citations per item

468.4

Sum of Times Cited

152,231

Without self citations

152,189

Citing articles

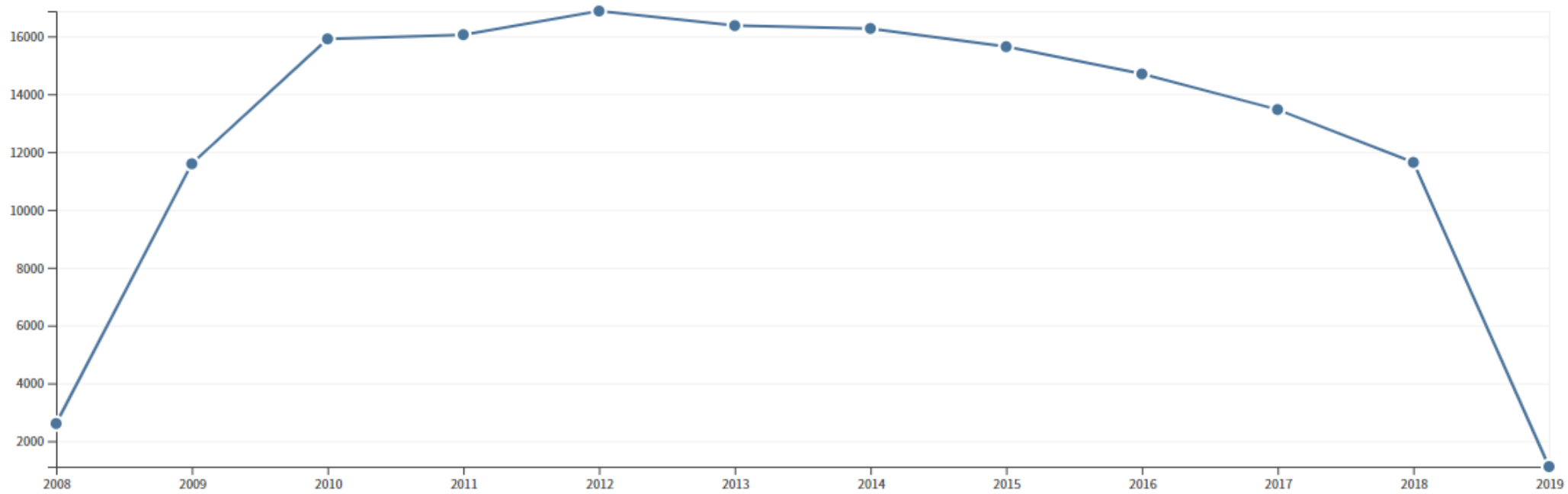
118,378 Analyze

Without self citations

118,349 Analyze

Sum of Times Cited per Year

Subject = Oncology



Total Publications

670 Analyze

2008 articles = 76,291

h-index

337

Average citations per item

487.1

Sum of Times Cited

326,356

Without self citations

326,098

Citing articles

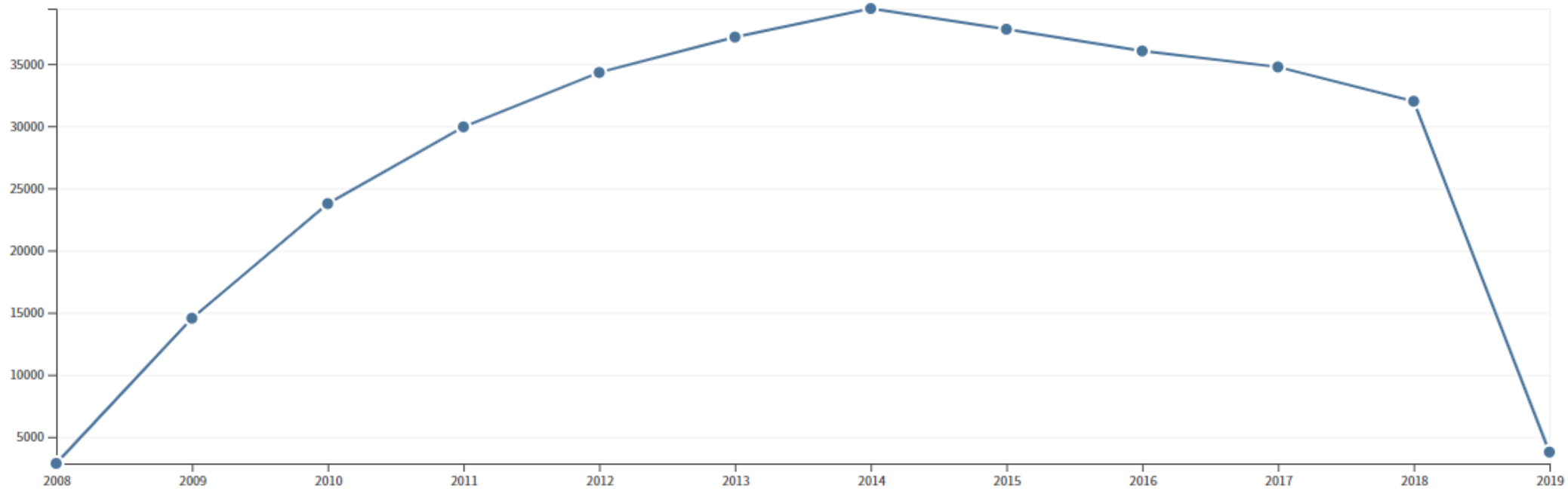
239,806 Analyze

Without self citations

239,679 Analyze

Sum of Times Cited per Year

Subject = Materials Science



Are metrics fit for purpose for evaluating researchers?

Is the data relevant to the variable being evaluated?

Is the metric a reliable indicator for the variable being evaluated?

- Citations are a reasonable indicator for scholarly impact
- The Journal Impact Factor is not an indicator of the quality or impact for a particular article

Are program, school, or country level data being used to evaluate individuals?

- It may not be feasible to adequately normalize and contextualize data for an individual scholar in a particular field
- Increasingly, organizational schools, departments, and programs cross traditional disciplinary boundaries

Are metrics fit for purpose for evaluating researchers?

Are the available data representative of all publication activity across the globe? Are the data available? Are the available data comprehensive? reliable?

> Web of Science & Scopus do not contain representative samples; significant unevenness by country of author, language of the publication, and discipline (Sugimoto & Lariviere, 2018)
>
> Field classifications in citation indices are more granular for the natural and medical sciences than the social sciences and humanities, so normalization for the latter is less precise (Sugimoto & Lariviere, 2018)
>
> “It is one thing to identify indicators for an assessment; it is quite another to ensure that the data to construct them are available, reliable, and accessible at an affordable price.” (Gringas, 2016, Ch 4)

What are some common mistakes in using metrics?

Using metrics as indicators of quality

Forgetting the limitations

- inherent in the bibliometric approach
- the properties of the data
- of the available data
- the ways in which the sociocultural context influences what research is done, by whom, and for what rewards.

What are some common mistakes in using metrics?

Using the Journal Impact Factor uncritically

- As a measure of quality for an individual article
- As a measure of quality for an individual scholar's work
- Without context – quartile rankings, disciplinary comparison, citation half-life
- Exclusively, without other evidence

- Using the h-index
- Presenting a metric or raw count without context and normalization
- Disconnect between the metrics presented and the story
- Using rankings in place of evaluation

How can we use metrics responsibly?

Metrics do not constitute evaluation. They are indicators to be used and considered as evidence for specific claims.

~~We value what we can measure~~

We measure what we value

San Francisco Declaration on Research Assessment

General Recommendation

1. Do not use journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist's contributions, or in hiring, promotion, or funding decisions.

For institutions

4. Be explicit about the criteria used to reach hiring, tenure, and promotion decisions, clearly highlighting, especially for early-stage investigators, that the scientific content of a paper is much more important than publication metrics or the identity of the journal in which it was published.

5. For the purposes of research assessment, consider the value and impact of all research outputs (including datasets and software) in addition to research publications, and consider a broad range of impact measures including qualitative indicators of research impact, such as influence on policy and practice.

San Francisco Declaration on Research Assessment

For researchers

15. When involved in committees making decisions about funding, hiring, tenure, or promotion, make assessments based on scientific content rather than publication metrics.

16. Wherever appropriate, cite primary literature in which observations are first reported rather than reviews in order to give credit where credit is due.

17. Use a range of article metrics and indicators on personal/supporting statements, as evidence of the impact of individual published articles and other research outputs [11].

18. Challenge research assessment practices that rely inappropriately on Journal Impact Factors and promote and teach best practice that focuses on the value and influence of specific research outputs.

The Leiden Manifesto

1. Quantitative evaluation should support qualitative, expert assessment.
2. Measure performance against the research missions of the institution, group, or researcher.
3. Protect excellence in locally relevant research.
4. Keep data collection and analytical processes open, transparent, and simple.
5. Allow those evaluated to verify data and analysis.
6. Account for variation by field in publication and citation practices.
7. Base assessment of individual researchers on a qualitative judgement of their portfolio.
8. Avoid misplaced concreteness and false precision.
9. Recognize the systemic effects of assessment and indicators.
10. Scrutinize indicators regularly and update them.

http://www.leidenmanifesto.org/

How can we use metrics responsibly?

Remember the role of cumulative advantage: “Researchers who are affiliated with prestigious institutions are more likely to receive citations (even when controlling for author and document characteristics), articles in journals of high reputation receive more citations than those in lower regarded journals (controlling again for confounding factors), and researchers who have more citations (as well as publications) are more likely to gain additional citations in a nonlinear manner compared with those who have fewer citations. **In short, science, like other social activities, is one where the rich get richer.**” (Sugimoto & Lariviere, 2018)

How can P&T processes reward community engaged research?

- Evaluation should be driven by values, rather than what is easy to measure.
- Evaluation should be driven by the goals of the funder, institution, the department or school, and the researchers.
- Structure the review process as one of qualitative expert judgement of a candidate's portfolio.
- Recognize and value many methods of inquiry, rather than valuing the scientific process as conducted by academic experts over all else.

How can P&T processes reward community engaged research?

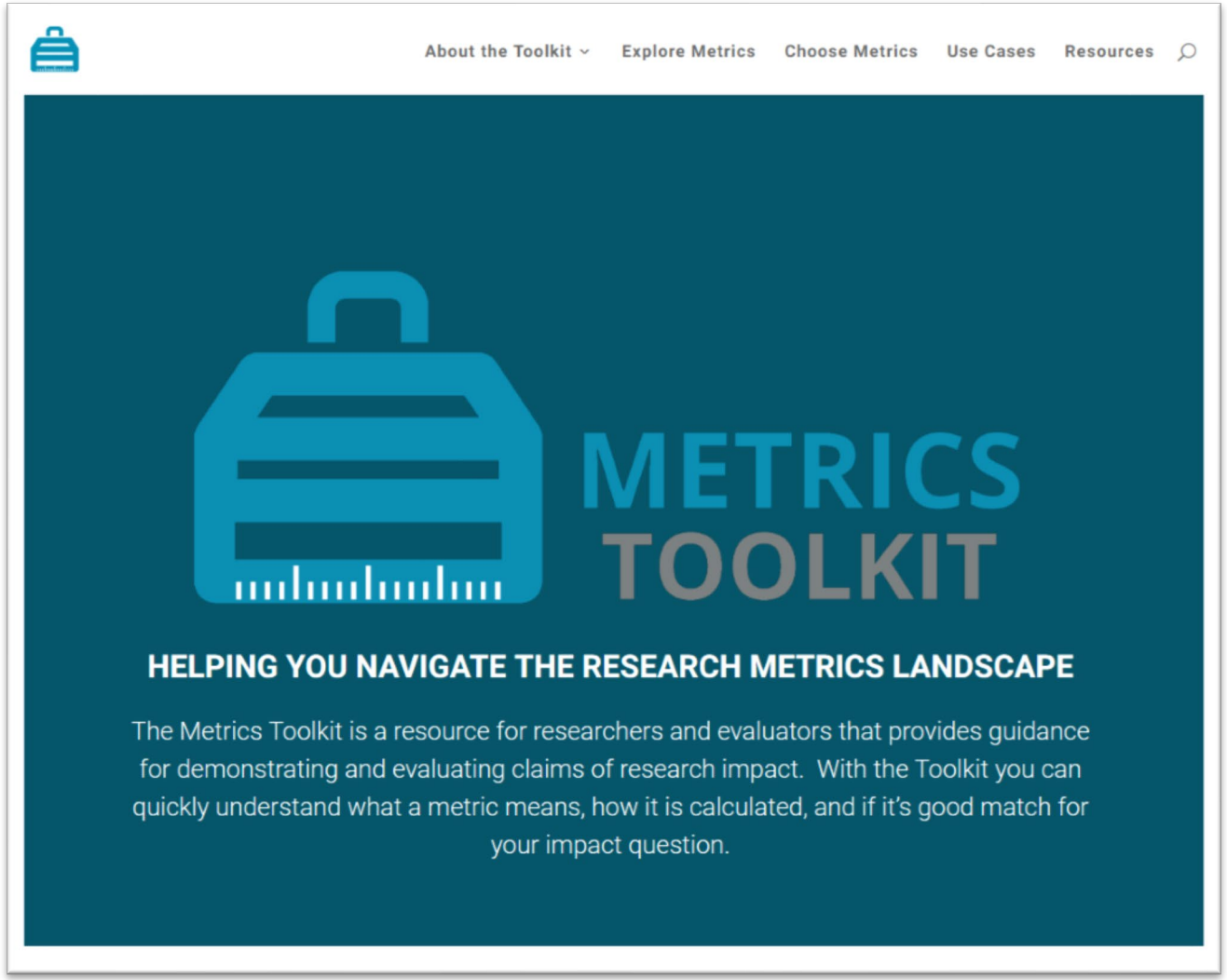
- Value and reward the work that goes into knowledge generation rather than centering rewards on publications.
- Explicitly recognize community, professional, and others as peers in the co-creation of new knowledge. Allow them to serve as both external reviewers and providers of recommendations.
- Explicitly acknowledge that the timeframe for many types of research may extend beyond the pre-tenure period. Support thoughtful consideration of intermediate outcomes based on goals set forth by the researchers.

How can P&T processes reward community engaged research?

- Institutions should do more to support administrators and faculty in conducting evaluations in a transparent and responsible manner.
  - Develop or adopt a set of guiding principles for using metrics and other indicators in evaluation processes (e.g., DORA, Leiden Manifesto).
  - Provide adequate training for department chairs, Associate Deans, and administrators in evaluating portfolios outside their field.
  - When external data sources will be used for evaluation purposes, demand better and more transparent documentation regarding the sources and limitations of data licensed from vendors like Clarivate Analytics (Web of Science, Journal Citation Reports), Elsevier (Scopus), Digital Measures (Academic Insight), Academic Analytics.

References

- Gingras, Y. (2016). *Bibliometrics and research evaluation: Uses and abuses*. MIT Press.
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature News*, *520*(7548), 429. https://doi.org/10.1038/520429a
- San Francisco Declaration on Research Assessment. https://sfdora.org/read/
- Sugimoto, C. R., & Larivière, V. (2018). *Measuring research: What everyone needs to know*. Oxford University Press.
- Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., …Johnson, B. (2015). *The Metric Tide*. https://doi.org/10.13140/RG.2.1.4929.1363



http://www.metrics-toolkit.org/

http://www.worldcat.org/oclc/1027811013

Heather L. Coates

Digital Scholarship & Data Management Librarian

Interim Co-Director, IUPUI University Library Center for Digital Scholarship

University Data Steward for Research Data

ORCID: 0000-0003-4290-6997

hcoates@iupui.edu