Seminar: Meta-Science

TECHNISCHE UNIVERSITÄT DARMSTADT

Steffen Eger

eger@aiphes.tu-darmstadt.de

What is this seminar about?

TECHNISCHE UNIVERSITÄT DARMSTADT

- Seminar name is a bit of a misnomer. Better name would be **“Science of Science”**
  - Since we analyze science with statistical, mathematical, and algorithmic tools
  - rather than doing philosophical debates about science in the spirit of Thomas Kuhn and others (theory of science)
- We study among others:
  - Whether scientific trends can be predicted
  - How scientific articles are evaluated (metrics, peer review, ...)
  - The dark side of science: scientific misconduct, improper comparisons, etc.

Resources

TECHNISCHE UNIVERSITÄT DARMSTADT

- Github with literature list:
  - https://github.com/SteffenEger/MetaScience
- Moodle:
  - https://moodle.informatik.tu-darmstadt.de/course/view.php?id=914

Course Organization

- Supervisor: Steffen Eger
- Main research focus on **Natural Language Processing**, including **humanities applications**, and **Deep Learning**
- Other courses/seminars:
  - Deep Learning for Natural Language Processing (DL4NLP) - SoSe
  - Deep Learning and Digital Humanities - WiSe
- https://www.informatik.tu-darmstadt.de/aiphes/aiphes/irg_position/index.en.jsp

Course Organization

TECHNISCHE UNIVERSITÄT DARMSTADT

- Tutor: Thomas Haider
- https://www.aesthetics.mpg.de/institut/mitarbeiterinnen/thomas-haider.html

Organization

TECHNISCHE UNIVERSITÄT DARMSTADT

To pass the course, you need to

- Make a presentation / video recording of about 40min + 20min (the latter for Q&A)
- Submit 3 questions to half/all of the video recordings
- Write a term paper of between 4-12 pages (limits depend on the number of authors on a paper)
- Be present during the discussion talk

Organization

TECHNISCHE UNIVERSITÄT DARMSTADT

- For both the term paper and the video, working in groups of 1-3 people **may** be possible
  - Depending on the number of final participants
  - Rule of thumb: we have 8-10 topics ...
- Groups of 3 people have to do a small-scale experiment
- Other than that, experiments can get bonus points

Organization

TECHNISCHE UNIVERSITÄT DARMSTADT

- We cancel the seminar block on 29.05.
- Instead, all presentations will be discussed on 26.06.
- You record a video at home and then send it in
- On 26.06. we discuss these videos
  - 17min for discussion, 3min for “spotlight” presentation
  - Overall, this will be 2-3 hours of discussion

Deadlines

TECHNISCHE UNIVERSITÄT DARMSTADT

- 01.06.: first version of slides is due (about 40 slides)
- 08.06.: final version of slides is due (with our feedback incorporated)
- 15.06.: video is due
- 25.06.: questions for videos are due
- 26.06.: Seminar takes place
- 10.09: Term papers are due

Topics

TECHNISCHE UNIVERSITÄT DARMSTADT

- 1: Overview over Sci-Sci. *Literature: A1, A2*
- 2: Citations and Altmetrics: *Literature: B1,B2,B3*
- 3: Trend Prediction I. *Literature: C1, C3, C4, C8*
- 4: Trend Prediction II. *Literature: C5, C6, C7*
- 5: Resources for Scientific Data: *Lit: D1,D2,D3,D4,E3*
- 6: Analysis. *Literature: E1,E2,(E3)*
- 7: Unethical behavior I. *Lit: F1, F2, F4, F7*
- 8: Unethical behavior II. *Lit: F3, F5, F6*
- 9: Peer Review. *Lit: G1, G2, G4, G5*
- 10: (Im)Proper Evaluation of Models. *Lit: H1,H2,H4,H5,H6*

Topic assignment & Presentation content

TECHNISCHE UNIVERSITÄT DARMSTADT

- How are topics assigned to students?
- You can vote on each topic with a score of 0-5
- Deadline for this is Monday, 27.04.
- We will then assign students to topics based on their scores

- In presentations, students should present a summarizing survey of the respective papers for their topic
- It is possible to put 1 or 2 papers in the center

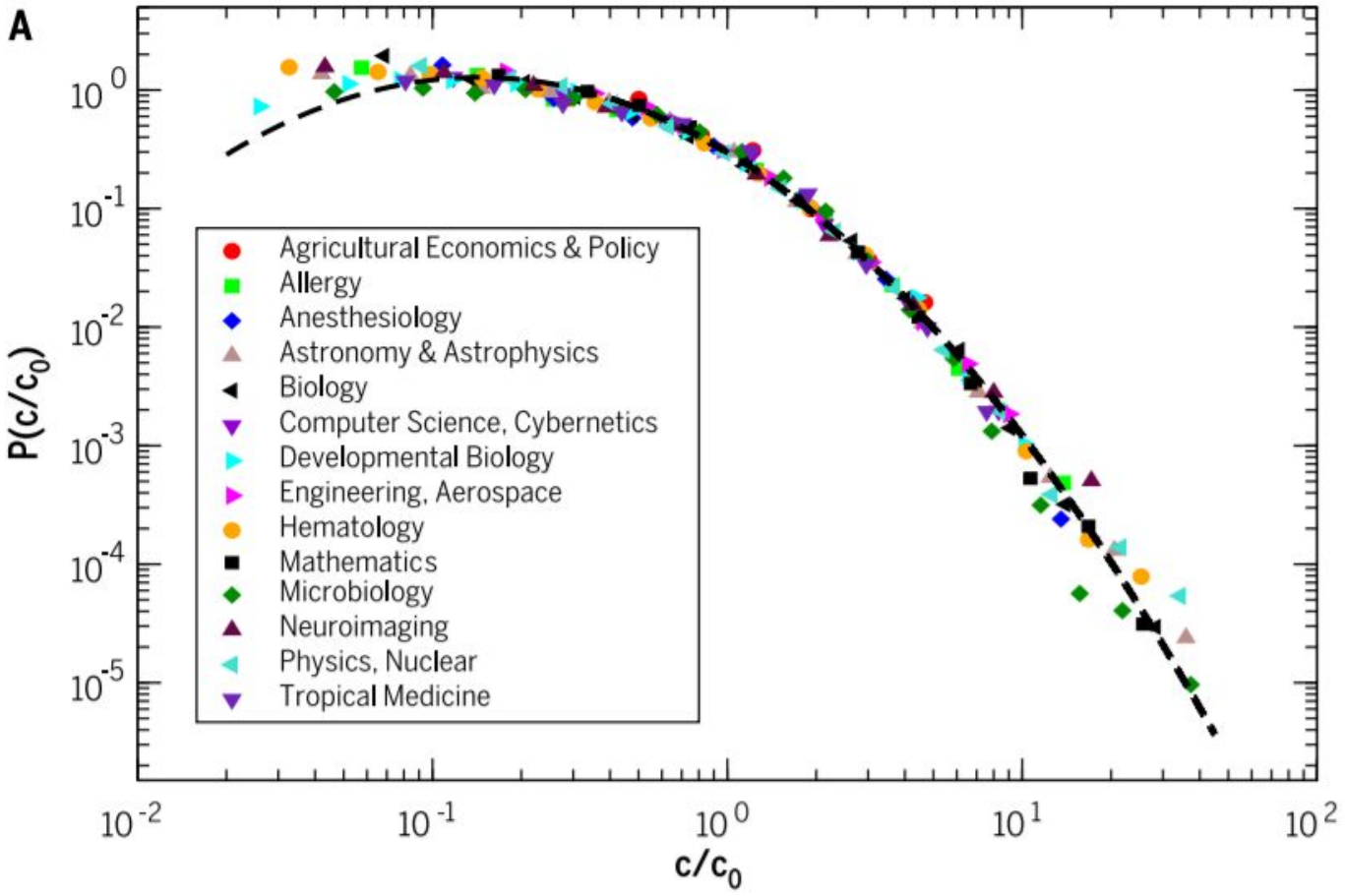
Topics

1. Overview

TECHNISCHE UNIVERSITÄT DARMSTADT

- When does a scientist have his/her largest impact paper?
- How does the number of co-authors evolve over time?
- Citation dynamics - do most papers get few citations?
- Growth of science - i.e., number of papers published within the last 100 years
- …..

1. Overview



From: Fortunato et al. Science of science

2. Citation and Altmetrics

TECHNISCHE UNIVERSITÄT DARMSTADT

- Can we trust citation counts as a measure of quality of research?
- What other metrics are there to evaluate scientific impact? (e.g. Twitter mentions/likes)
- How do authors use citations in papers?

2. Citation and Altmetrics

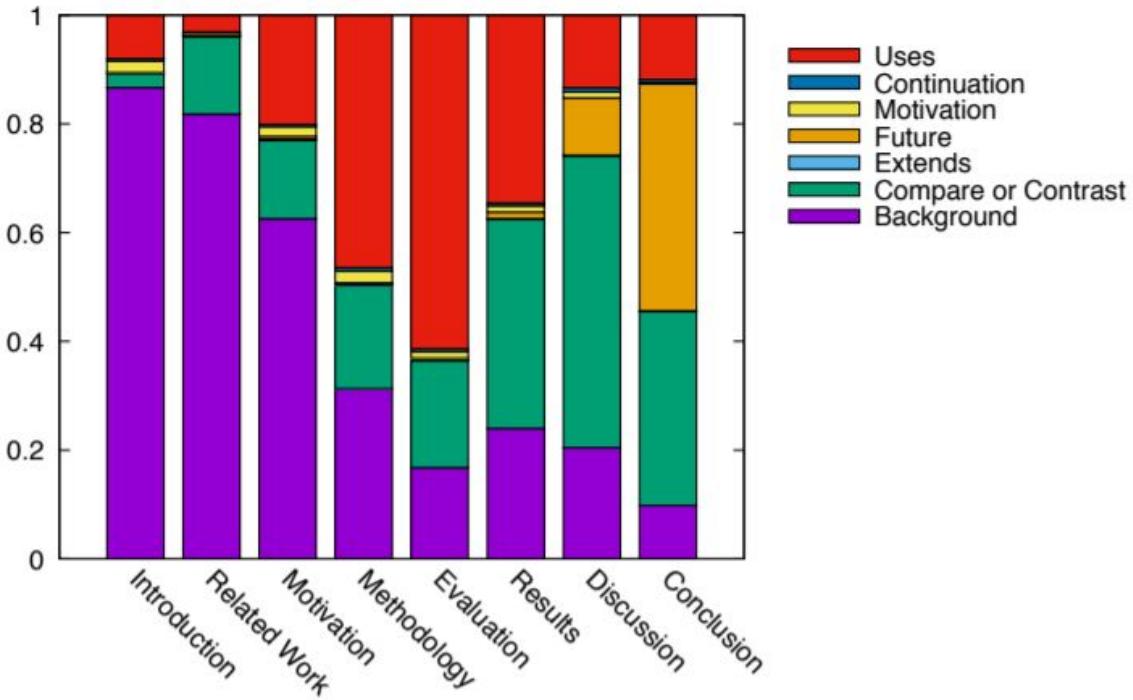


Figure 2: Expected percentage of citation functions per section shows a clear narrative trajectory across sections.

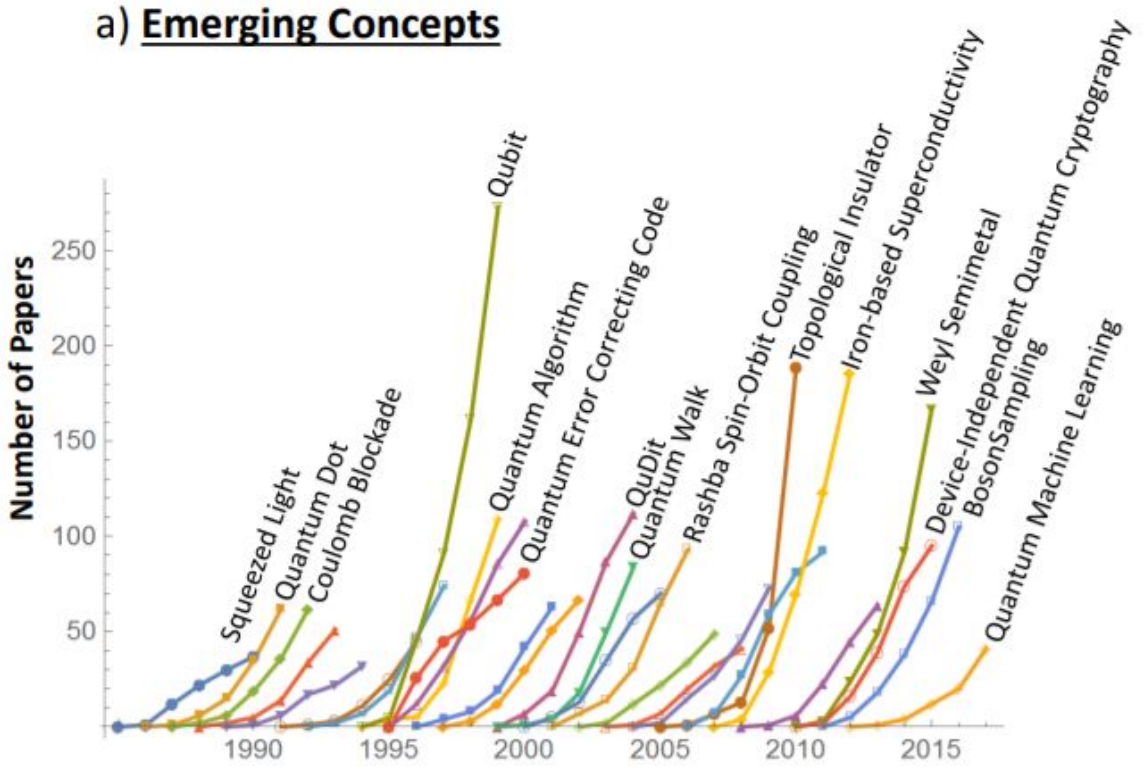
From: Jurgens et al. **Citation Classification for Behavioral Analysis of a Scientific Field**

3. Trend prediction I

TECHNISCHE UNIVERSITÄT DARMSTADT

- Can we predict which papers will have large citation counts in the future?
- Can we predict which research trends will be popular in 5 years? (could the Deep Learning revolution have been predicted?)

3. Trend prediction I



From: Krenn and Zeilinger. Predicting Research Trends with Semantic and Neural Networks with an application in Quantum Physics

4. Trend prediction II

TECHNISCHE UNIVERSITÄT DARMSTADT

Can we predict trends or citation counts from

- Social influence (rather than content of the paper)?
- The bibliography of papers?
- Rhetorical framing?

4. Trend prediction II

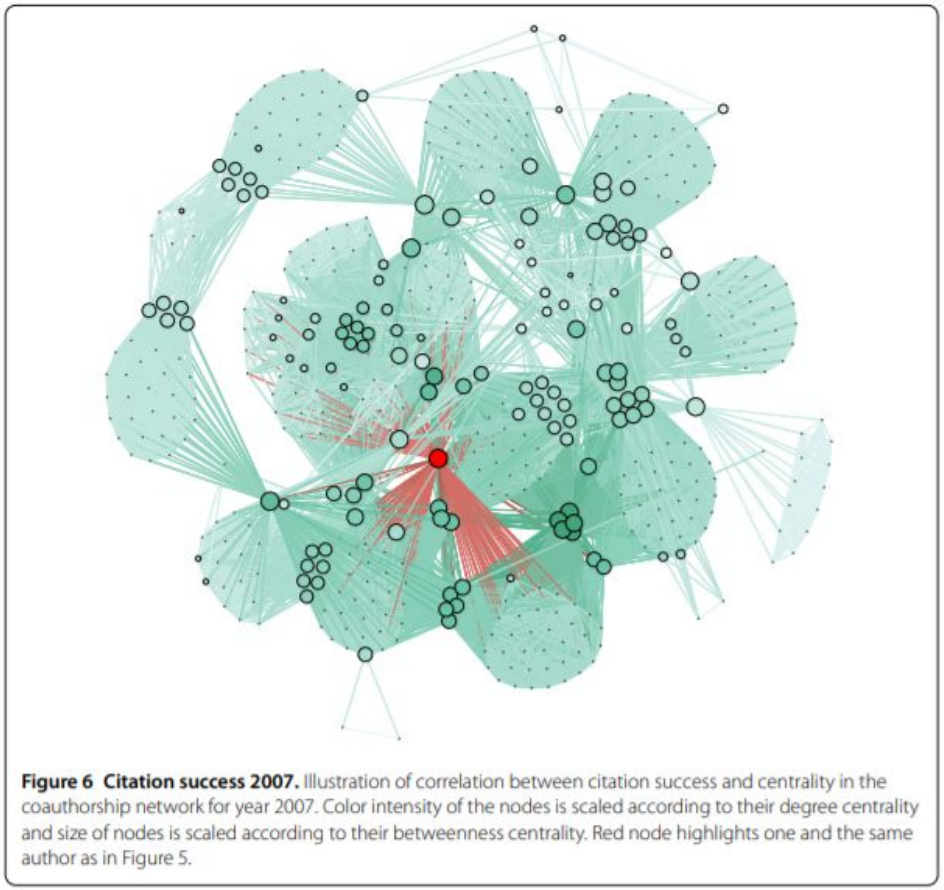


Figure 6 Citation success 2007. Illustration of correlation between citation success and centrality in the coauthorship network for year 2007. Color intensity of the nodes is scaled according to their degree centrality and size of nodes is scaled according to their betweenness centrality. Red node highlights one and the same author as in Figure 5.

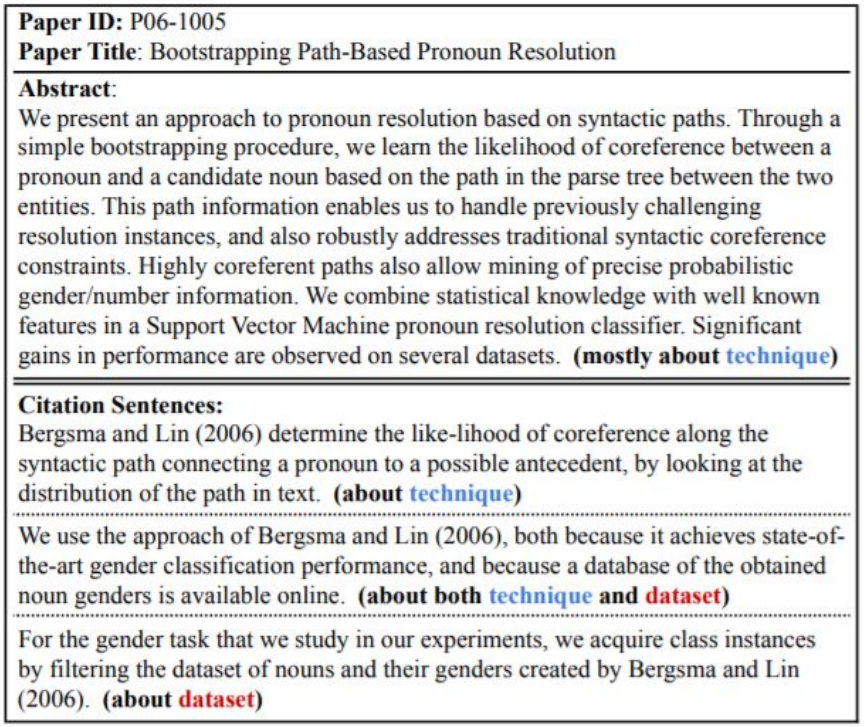
From: Sarigöl et al., Predicting scientific success based on coauthorship networks

5. Resources

TECHNISCHE UNIVERSITÄT DARMSTADT

- How to summarize and annotate scientific data?
- How can we extract information from scientific articles?

5. Resources



Paper ID: P06-1005
Paper Title: Bootstrapping Path-Based Pronoun Resolution

Abstract:
We present an approach to pronoun resolution based on syntactic paths. Through a simple bootstrapping procedure, we learn the likelihood of coreference between a pronoun and a candidate noun based on the path in the parse tree between the two entities. This path information enables us to handle previously challenging resolution instances, and also robustly addresses traditional syntactic coreference constraints. Highly coreferent paths also allow mining of precise probabilistic gender/number information. We combine statistical knowledge with well known features in a Support Vector Machine pronoun resolution classifier. Significant gains in performance are observed on several datasets. **(mostly about technique)**

Citation Sentences:
Bergsma and Lin (2006) determine the like-lihood of coreference along the syntactic path connecting a pronoun to a possible antecedent, by looking at the distribution of the path in text. **(about technique)**

We use the approach of Bergsma and Lin (2006), both because it achieves state-of-the-art gender classification performance, and because a database of the obtained noun genders is available online. **(about both technique and dataset)**

For the gender task that we study in our experiments, we acquire class instances by filtering the dataset of nouns and their genders created by Bergsma and Lin (2006). **(about dataset)**

Figure 1: Abstract and citations of (Bergsma and Lin 2006). The abstract emphasizes their pronoun resolution techniques and improved performance; the citation sentences reveal that their noun gender dataset is also a major contribution to the research community, but it is not covered in the abstract.

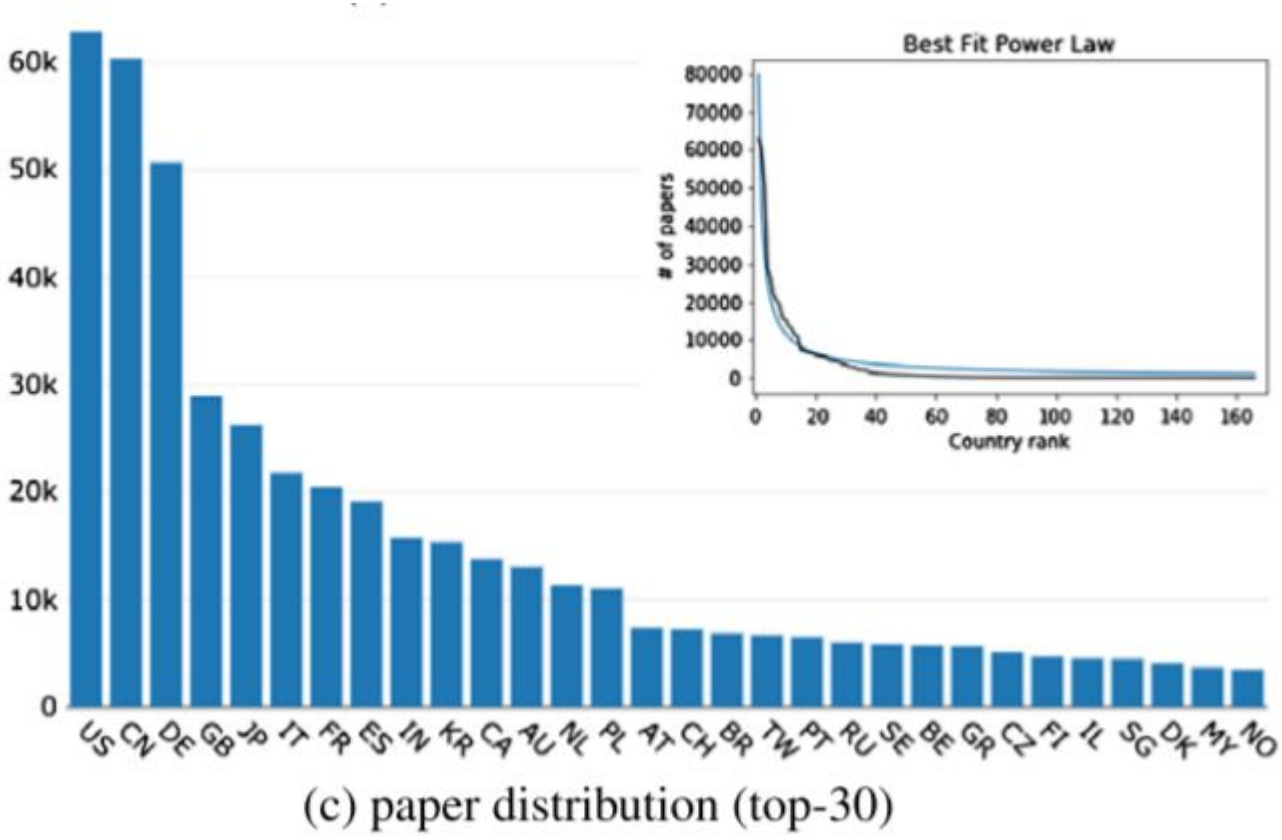
From: Yasunaga et al. ScisummNet: A Large Annotated Corpus and Content-Impact Models for Scientific Paper Summarization with Citation Networks

6. Analysis

TECHNISCHE UNIVERSITÄT DARMSTADT

- Geographic trends in publications: Do all countries contribute equally to the scientific process?
- How (fast) do new technologies spread across research communities?

6. Analysis



(c) paper distribution (top-30)

From: Mannocci et al. Geographical trends in academic conferences: An analysis of authors' affiliations

7. Unethical Behavior I

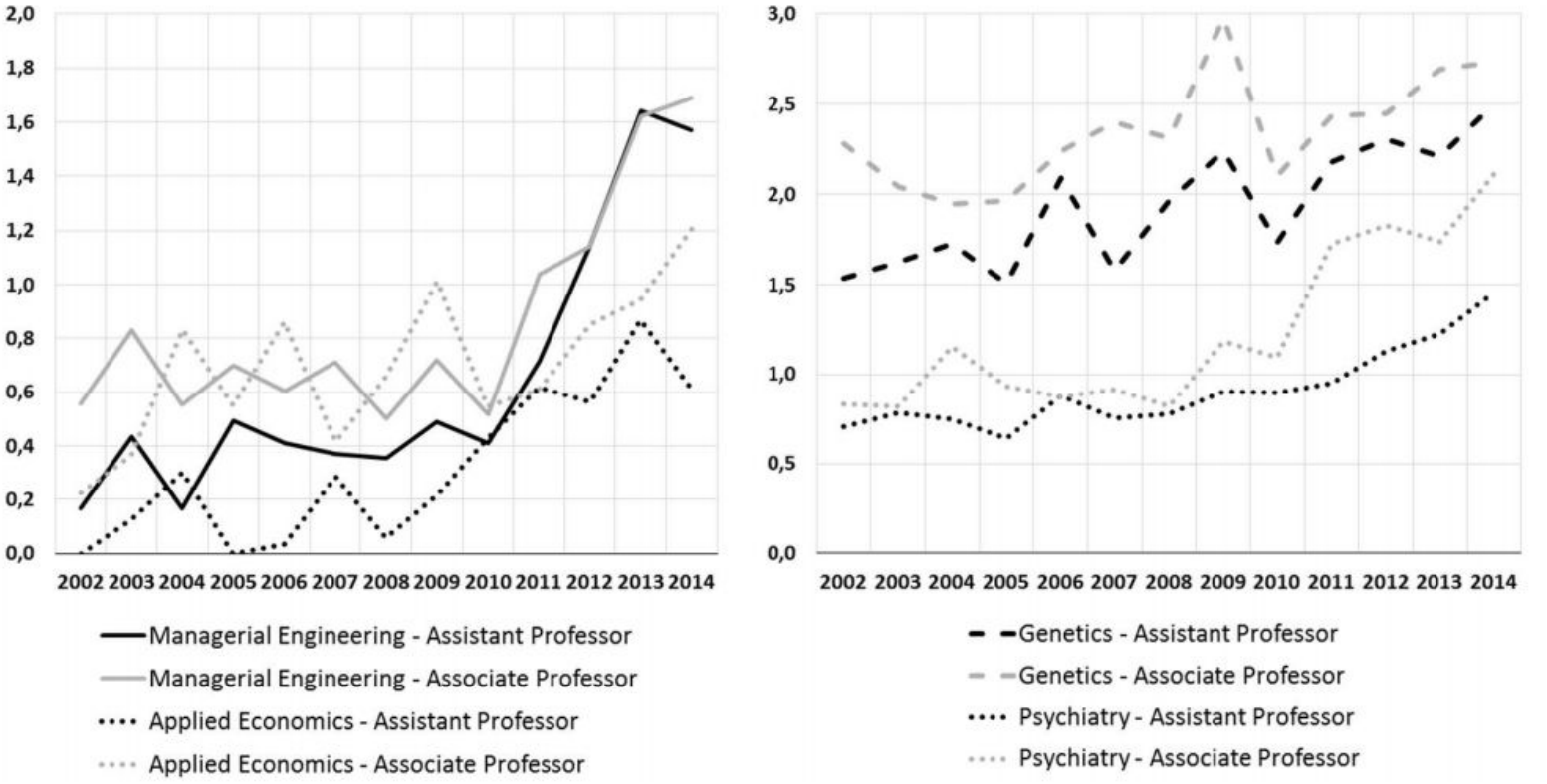
TECHNISCHE UNIVERSITÄT DARMSTADT

- How and why do authors “tune” their citation scores
  - By self-citations?
  - Establishing “citation cartels” with others?

7. Unethical Behavior I

TECHNISCHE UNIVERSITÄT DARMSTADT

Figure 1 – Evolution of the number of self-citations per article



From: Seeber et al., Self-citations as strategic response to the use of metrics for career decisions

8. Unethical Behavior II

TECHNISCHE UNIVERSITÄT DARMSTADT

- (Why) are scientists cheating - and how?
- Did the problem get bigger recently?
- What biases are there in scientific research?

8. Unethical Behavior II

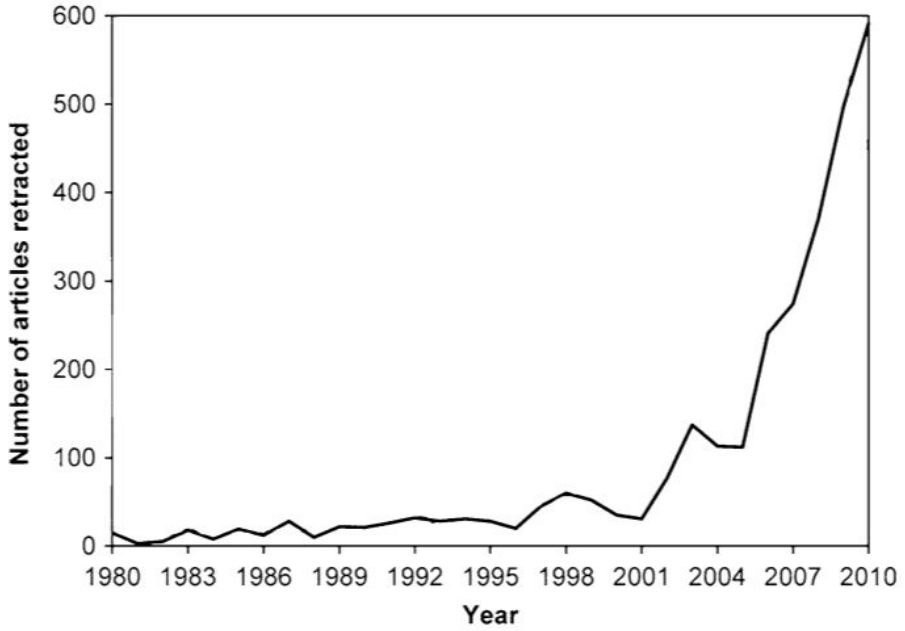


Fig. 4 Number of scientific articles in the years indicated between 1980 and 2010. The retractions have increased nearly 20-fold in the last 10 years. The rate at which the number of retractions have increased significantly exceeds that of the number of published articles (about 44 % in the last decade, see also Fig. 5). Modified from Grieneisen and Zhang (2012)

From: Carafoli, Scientific misconduct: the dark side of science

9. Peer Review

TECHNISCHE UNIVERSITÄT DARMSTADT

- Is the reviewing process random?
- How do reviewers react to the authors' rebuttal?
- How can we improve peer reviewing?

9. Peer Review

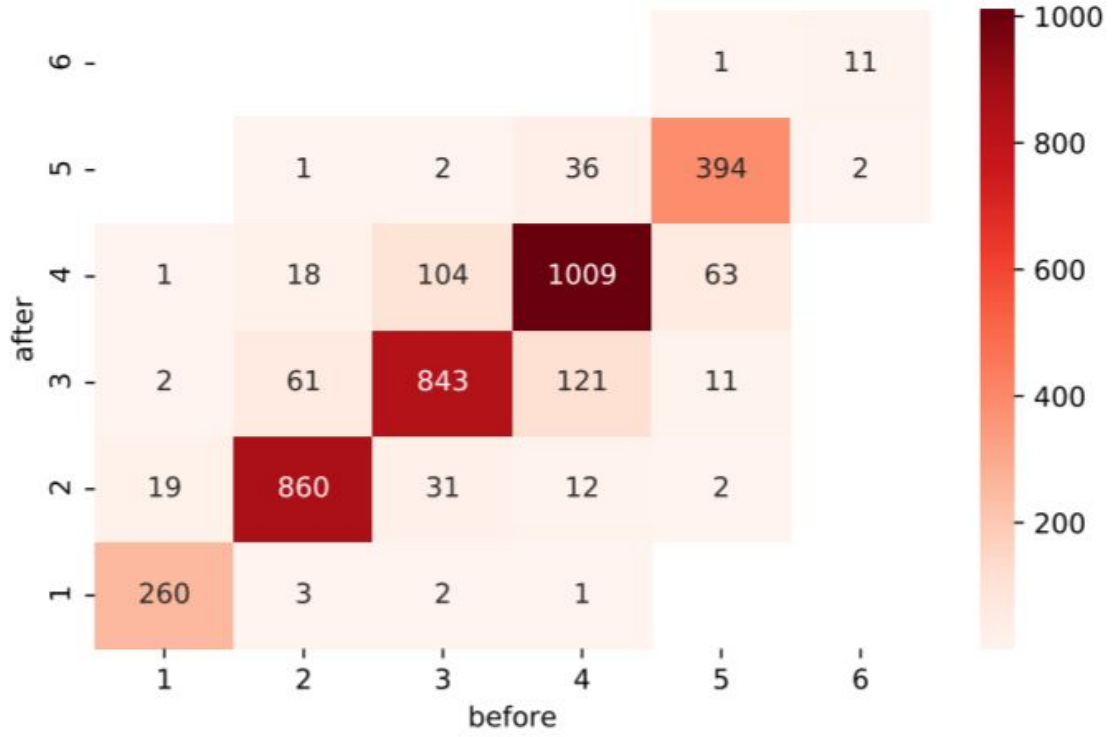


Figure 3: Before vs after rebuttal OVAL.

From: Gao et al., Does my rebuttal matter?

10. (Im)Proper Evaluation of Models

TECHNISCHE UNIVERSITÄT DARMSTADT

- Can we trust current evaluations of scientific models (in machine learning)?
- How can we improve evaluation methodology?

10. (Im)Proper Evaluation of Models

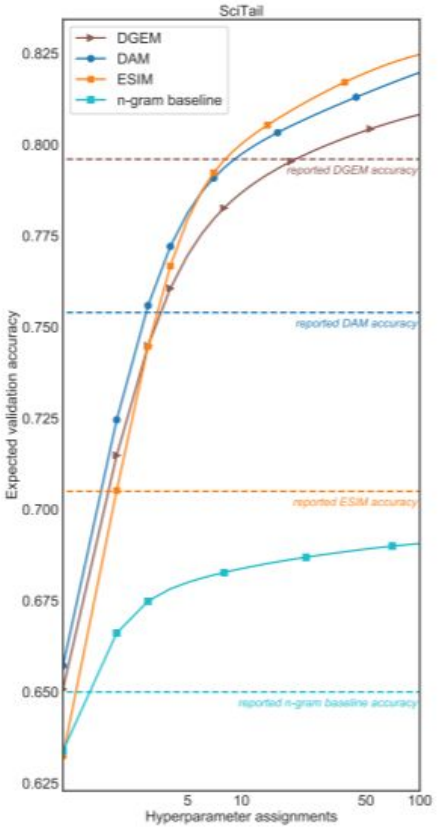


Figure 3: Comparing reported accuracies (dashed lines) on SciTail to expected validation performance under varying levels of compute (solid lines). The estimated budget required for expected performance to match the reported result differs substantially across models, and the relative ordering varies with budget. We omit variance for visual clarity.

From: Dodge et al., Show Your Work: Improved Reporting of Experimental Results