[Northeast Journal of Complex Systems (NEJCS)](https://orb.binghamton.edu/nejcs)

Volume 5
Number 1 *Special Issue: NERCCS 2023 Papers*

Article 7

June 2023

She is an Expert in this Research Field: The Signal of Recent Publications' Relevance

Gil Zeevi
University of Haifa, gzeevi25@gmail.com

Osnat Mokryn
University of Haifa, omokryn@is.haifa.ac.il

She is an Expert in this Research Field: The Signal of Recent Publications' Relevance

Gil Zeevi and Osnat Mokryn*
Information Systems, University of Haifa, Haifa 3498838, Israel
* Corresponding author, e-mail: omokryn@is.haifa.ac.il

Abstract

Assessing the expertise of researchers has garnered increased interest recently. This heightened focus arises from the growing emphasis on interdisciplinary science and the subsequent need to form expert teams. When forming these teams, the coordinators need to assess expertise in fields that are often very different from theirs. The conventional reliance on signals of success, prestige, and academic impact can unintentionally perpetuate biases within the assessment process. This traditional approach favors senior researchers and those affiliated with prestigious institutions, potentially overlooking talented individuals from underrepresented backgrounds or institutions.

This paper addresses the challenge of determining expertise by proposing a methodology that leverages the relevance of a researcher's recent publication track to the proposed research as a "sensemaking" signal. We introduce a novel $\alpha-$*relevance* metric between the trained embedding over the titles and abstracts of a researcher's recent publications and the embedding of a call and show that high values of $\alpha-$*relevance* indicate expertise in the field of the call. By evaluating the $\alpha-$*relevance* threshold, we establish a robust framework for the assessment process. For the evaluation process, we use (1) NIH grant-winning records and researchers' publications obtained from Scopus and (2) grant submissions dataset from a research university and the corresponding researchers' publications. Additionally, we investigate the optimal time window required to capture the researcher's expertise based on their publication timeline. Considering the temporal relationship between grant winnings and publications, we identify the most informative time window reflecting the researcher's relevant contributions.

The data-driven methodology transcends traditional signals of success, promoting a fair evaluation process of the researcher's relevance to the proposed research. By leveraging objective indicators, we aim to facilitate the formation of expert teams across disciplines while mitigating biases in assessing expertise.

1 Introduction

Identifying expertise, in research, in particular, is a challenging task that influences careers, funding, and science itself. Often, expertise is equated with prestige. Yet, the known Mathew effect, implying that prestigious scientists receive larger recognition than their research alone merits, has been shown repeatedly [26, 36, 4, 15, 28].

Identifying experts and expertise is of interest at large. Expertise has been the focus of much research in the social sciences and is often established using peer evaluations and seniority, defined by years in the field [19, 8, 14]. Special recent attention was given to expertise in the context of human-computer interactions and collaborative work, as experts often require different design considerations [31, 38, 17, 29, 25], and identifying experts in organizations is important for collaborative tasks [29, 21, 18]. The distinction between novices and experts is well established. Yet, experts can also be differentiated and ranked. Measures to evaluate levels of expertise in the workforce vary and include peer evaluation [19] and seniority, measured in years [8].

In recent years, collaborative and interdisciplinary efforts are becoming popular in research and science. Growing demand from funding agencies for interdisciplinary research contributes to the formation of collaborative research teams [22]. The expertise of team members was found to have a profound effect on the performance and results of the team [7]. Hence, in forming such groups, it is important to find highly proficient researchers in their field and considered experts. Recent research evaluating the impact of millions of research teams found that teams composed of senior, prominent researchers had a substantially higher impact than that of heterogeneous teams [37]. However, following the seniority rule, or choosing the highest impact researchers for the team, is a self-reinforcing feedback mechanism: it will exacerbate the rich-get-richer phenomena, deprive young and minority researchers of opportunities [20], and deepen the inherent bias in academia of favoring researchers from prestigious institutions [35].

It is then imperative to devise a methodology to identify research experts that will not rely on known biased characteristics. Yet, the ability to assess others' quality of work decreased with the "intellectual distance" from it, even in one's own field [6]. Here, we suggest that when choosing a collaborator for a grant proposal, researchers' publication lists will be used to identify expertise in the field. We devise a methodology that creates an embedding of a researcher's publications and compare it to the suggested grant call. We devise a similarity measure, $\alpha-$*relevance*, that quantifies the similarity of the researcher's embedded publications to the embedding of the proposed grant. We then show that $\alpha-$*relevance* can be used for signaling a researcher's expertise.

Academic papers are research products. A researcher's productivity and success in publishing papers, i.e., *publish or perish*, is critical to their success [11]. We study the recent publication record to quantify signals of expertise. In most grant proposals, a PI submits their recent publication record. A strong recency effect for popularity has been previously found in many fields, among them in scientific publications [27]. We then further investigate the suitable time window for detecting the proposed signal.

To evaluate the $\alpha-$*relevance* metric and find the most relevant time window, we measure the similarity of the embedding of the recent publication track record of single-PI grant recipients with the grant proposal they won. We hypothesize that people that win grant proposals have highly relevant papers in that field before the grant submission. A grant is a signal of recognition awarded to a researcher whose grant proposal was deemed excellent. An additional necessary condition is that the committee recognizes the researcher as an expert in the field. When submitting grants, it is typical for the funding agency to require the researcher to submit, in addition to the proposal, a CV detailing the researcher's academic achievements, including a list of prior publications. The grant proposal reviewers are required to evaluate the researcher's expertise in the field *in addition* to evaluating the proposal.

EVALUATION CRITERIA

Criterion 1 - RESEARCH PROJECT
Ground-breaking nature and potential impact of the research project
To what extent does the proposed research address important challenges?
To what extent are the objectives ambitious and beyond the state of the art (e.g. novel concepts and approaches or development between or across disciplines)?
To what extent is the proposed research high risk/high gain?

Scientific Approach
To what extent is the outlined scientific approach feasible bearing in mind the extent that the proposed research is high risk/high gain (based on the Extended Synopsis)?

Criterion 2 - PRINCIPAL INVESTIGATOR
Intellectual capacity, creativity and commitment
The questions below can have one of the following four responses: Outstanding/Excellent/Very good/Non-competitive
To what extent has the PI demonstrated the ability to propose and conduct ground-breaking research?
To what extent does the PI provide evidence of creative independent thinking?
To what extent have the achievements of the PI typically gone beyond the state of the art?

Figure 1: Evaluation criteria for grant proposals

Figure 1 shows the evaluation instructions of a prestigious funding agency. The last point in "Criterion 2 - Principle Investigator" relates to the researcher's prior achievements in the field. The reviewer is asked to estimate the quality of the researcher's previous achievements, among them her prior publications in the grant's field of research.

Our dataset comprises a database of winnings obtained from the National Institutes of Health (NIH) from 1992-2021 and a database of submissions to various funding agencies and their outcomes for all the researchers in a Research University (RU) from 2008 to 2022. Collecting the researchers' publication records from Scopus for the eight years before each grant submission, we perform a longitudinal study to quantify the relevance of the researcher's publication record in the years prior to a grant submission date to that of the proposed grant.

Our experiments show that the high relevance of a trained embedding over the titles and abstracts of the last four years of publications is a valid signal for expertise in a field.

2 Signals of expertise

Social communities, whether in nature or humans, use different mechanisms to signal each other their perspective rank and dominance [33, 39]. Like all life forms, humans are looking for signals that help them to assess the relative ranking and status of others in both their personal life, work environment, and, in essence, in all social encounters [32, 23, 10, 13, 24]. Recently, Signal Theory has been suggested as a mechanism to infer strangers' expertise from signals in digital artifacts [31]. Digital artifacts are a form of digital representations of real personas, such as their blog post, a self-description, or other information distilled, reported, or summarized in an online profile. They further demonstrated the use of signaling theory as a decision heuristic in the 'people sensemaking' process [31].

The use of signals to infer the status of scientists and their publications has been the focus of much research [16]. Within these signals are their overall publication record, awards, and prizes, as well as status signals such as the department or the university they belong to. The known Mathew effect relates to the effect of such signals, claiming that highly reputed scientists, whether because of their publication track record, prizes, or their university's prestige, receive disproportionately more citations than those with low reputations [26]. However, somewhat contradicting results were found in a more recent study that examined the signaling mechanisms for the quality of a publication [34]. They suggested that such attribution exists but might be more prominent during the review process. The importance of citation counts and their signaling power has also been the focus of much debate [5, 1].

Here, we suggest that a measure of the relevance of the recent publications of a scientist to a grant proposal is a signal that can be used to infer their expertise in the field.

3 Data

We collected several datasets of grant submissions, grant wins, and the grant authors' publication lists. Two grants were collected. One is of winning submissions made to NIH grant proposals, and the second is of submissions made by a research university's scientists and their outcomes. In addition, we have the publication record of all scientists in our datasets. The datasets differ in the information we have for the grants and the submissions: The Research University (RU) dataset contains precise submission dates but lacks information on grant call release and expiration dates. The NIH dataset provides grant call release and expiration dates but does not contain proposals' submission dates. Furthermore, the call for submission period varies across calls in the NIH dataset, ranging from several months to over three years.

We detail here the process of obtaining the various datasets.

3.1 NIH winnings related datasets

We have obtained expired funding opportunities from the NIH website and conducted scraping to retrieve the short description of the grant calls, as this information is not provided by default in the exported report of funding opportunities. After retrieving the expired funding opportunities, we used the NIH's reporter API to retrieve the winning applications for each opportunity. The dataset of winning applications comprises detailed information about the winning principal investigators (PIs). Each row in the dataset provides comprehensive details, including the full names of the winning team, the awarded amount, the affiliated organization, the city and country of affiliation, the proposal's abstract, and other pertinent information.

Following that, we employed Elsevier's Scopus API to access the publications of successful PIs. The first step in this procedure entailed retrieving the Scopus ID for each PI. For this purpose, we executed a query using the PI's first name, last name, and country of affiliation. The query response potentially yielded multiple Scopus ID candidates for each PI. To identify the most suitable candidate, we attempted to associate it with the affiliated organization name as provided in the NIH winning applications. This method proved successful in only 48% of the cases. To improve the results, we conducted a refined search by modifying the affiliation name. Specifically, we removed terms such as 'university' and 'of' from the affiliation name. This adjustment resulted in an additional 3% of successful retrievals. For the remaining 49% of Scopus IDs, we employed an alternative approach. We selected the ID corresponding to the city where the affiliated organization was located. This method allowed us to retrieve the Scopus IDs for nearly half of the cases where the previous affiliation-based method was unsuccessful. To validate the effectiveness of this alternative method, we randomly sampled some of the retrieved IDs and found that it yielded positive results. However, it is important to note that due to the inability to validate all of the IDs manually, bias may be introduced into the retrieved data through this approach.

In conclusion, the outcome of this process revealed a total of 21,641 distinct PIs associated with 42,304 winning submissions across 5,785 grant calls and a total of 2,740,592 publications published by the winning PIs up to the retrieval point. It is worth noting that a PI can secure multiple victories, and the distribution of repeated wins by PIs can be visualized in Figure 2. For the experiments conducted here forth, we used the details of 10,810 single-winning PIs and their first recorded win within the examined period.

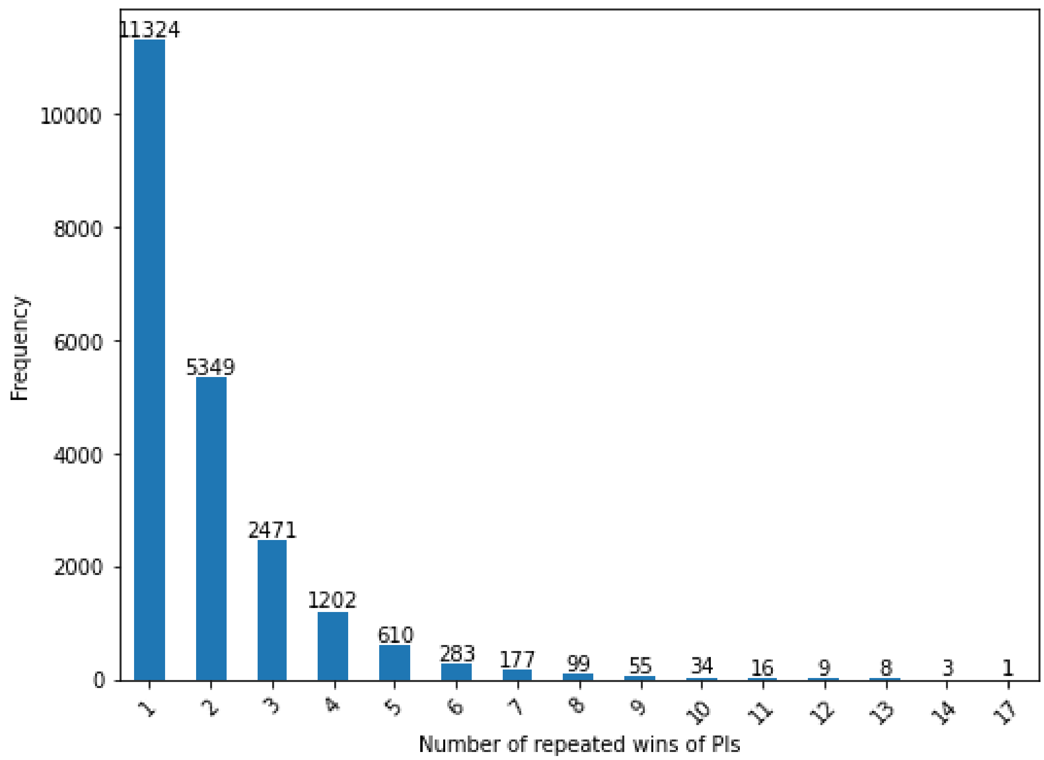


Figure 2: The frequency distribution of the number of repeated wins among PIs in the acquired winning submissions from the NIH

3.2 RU datasets

The second dataset was collected at a research university. It contains the university's scientists' grant submissions to various funds between 2008-2022 and their outcomes. The dataset contains 7258 proposals that 783 individual researchers submitted. Out of these, 2402 proposals were approved, and 4856 were rejected. The number of collaborative teams was relatively low, with only 86 being awarded a grant and 149 being denied. Similar to the NIH datasets construction, we also obtained for each PI their publications from 1999 to 2021 using Scopus, a total of 21799 publications for all these researchers. The researchers' publication record was obtained to its fullest.

4 Methodology

The methodology used in this study aims to investigate the similarity between researchers' papers and the grant calls to which they submitted proposals. We conduct a longitudinal study that evaluates for each researcher the similarity between her publication record in the years before submitting the grant proposal and the grant itself.

To that end, we employ a cross-referencing approach by comparing papers' titles, abstracts, and keywords (when available) with the corresponding elements of a grant call. Each paper's vector representation, defined below, is compared to the corresponding grant's representation using cosine similarity.

To represent the papers and the grants, we used two Transformers Embedding methods: SPECTER - Document-level Representation Learning using Citation-informed Transformers [9] and Sentence-BERT [30].

SPECTER builds on SciBERT [2], an adaptation of the original BERT [12] architecture to the scientific domain. SPECTER adds a relatedness signal into the embedding representation by designing a loss function that trains the Transformer model to learn closer representations for papers based on their citation patterns. That is, a paper will be closer to the papers it cites or that are citing it than to papers with which it has no citing relations. SPECTER's advantage is that it only requires the title and abstract of the given input paper at the inference stage and does not need the relevant citation information about it. SPECTER produces word embedding also for new papers that have yet to be cited, which allows us to work with it as we consider recent scientific papers.

As we use the title and abstract of each paper, we also use Sentence-BERT [30], a popular embedding commonly used for short texts and embedding sentences of up to 128 tokens.

We evaluated these two transformers against a classical method - Latent Dirichlet Allocation (LDA) [3]. When using transformers, we encoded the grant calls and papers text without preprocessing but with a separation token ('[SEP]') between the title and abstract. In contrast, when using LDA, we preprocessed the data by removing stop words and punctuation and performing lemmatization to reduce the dimensionality and focus on the keywords in the text.

To identify topics with LDA, we assembled all publications and grant calls into one corpus. We then used LDA to identify the latent topics in the corpus by optimizing the topic coherence measure. For each document in the corpus, we created a sparse vector representation that held for each identified topic the probability that the paper belonged to that topic.

By generating a vector representation using either transformer models or LDA, we could analyze similarity measures between the researcher's publications before the submission date and the grant.

In our suggested work, we hypothesize that a researcher's recent publication record in a field can be used as a signal for their expertise in the field. To that end, we examine their track record before being awarded a grant and compare the similarity of their publications in the years before being awarded the grant with the grant itself, i.e., the grant call for proposals.

5 Results

Here, we explore the resemblance between researchers' recent publications and the grant they won. We hypothesize that in the part in which the reviewer evaluates the researcher's expertise in the field, *after* finding the researcher's proposal exciting and relevant, the reviewer uses the researcher's recent publication list as a signal of established expertise in the field. Other expertise signals might be used, such as the institution's prestige, the researcher's citation track and seniority, prizes and previous wins, and more. Here, we suggest that the relevance of the recent works is also a 'makesense' signal of expertise and can be used in cases where other signals are missing.

As a preliminary step, we examine the degree of similarity between the researcher's proposal for the grant and their previous publications over the NIH dataset, where the proposal is available. Figure 3 shows the α-correlation between researchers' publications in the years before the submission and their winning grant proposal. Here, we find for each researcher the publication *most similar* to their grant proposal in the years prior to the grant release date. This indicates whether the proposal is based on any of their recent works. We see that, on average, a researcher's proposal is similar to some of the papers in her recent publication record. The longitudinal study shows that, on average, the proposal is more similar to one of the researcher's *very recent* papers. The SPECTER transformer gives better results,

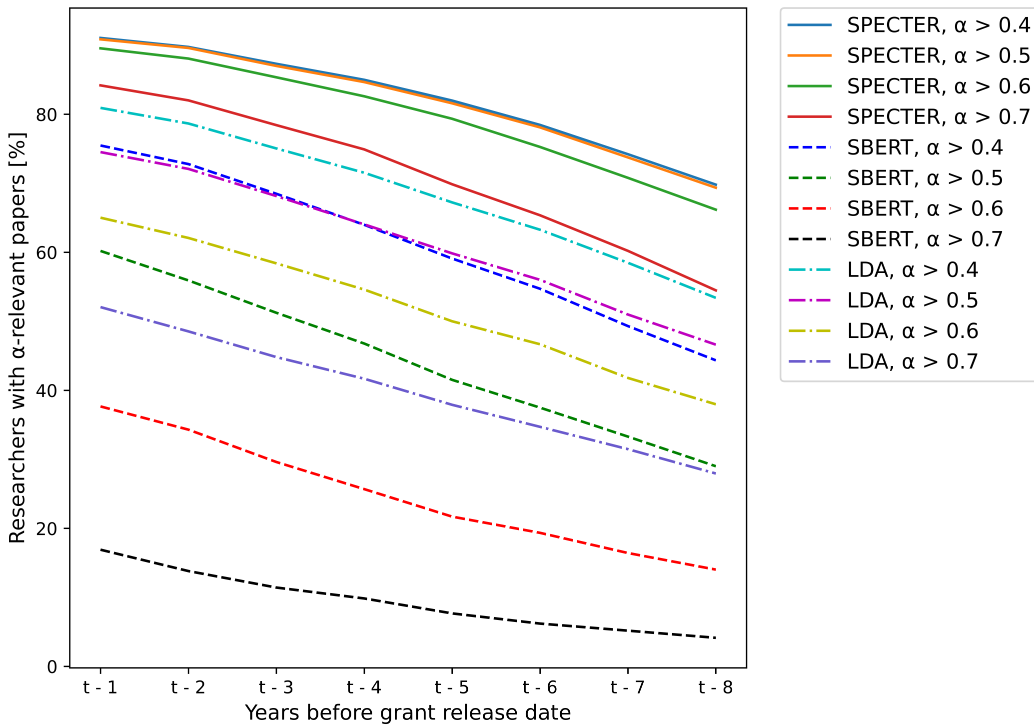


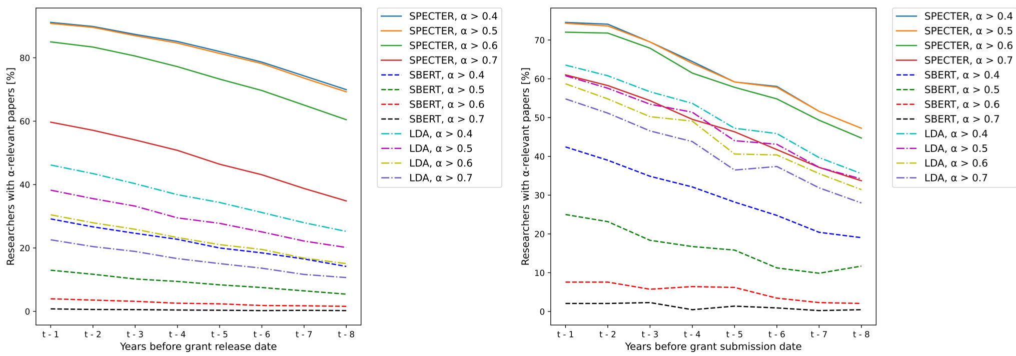
Figure 3: The α-relevance between a winning NIH grant proposal and the yearly most similar paper of the PIs who won that grant, in each of the years before the grant release date, denoted with 't'. Calculated separately for each researcher and the NIH grant they won. Y-axis denotes the percentage of winning researchers with that similarity.

even when looking for cosine similarity as high as $\alpha = 0.7$.

We then disregard the researcher's proposal and will examine how her publication record correlates with the grant call itself.

We find similar results when performing the longitudinal study against the grant's call for proposal rather than the researcher's proposal. Figure 4 depicts the results for the most similar paper for both datasets. In the RU dataset, we included only winnings of grants indicated as *competitive* by the University's research authority. We find that, on average, more than 50% of the researchers that won a grant had at least one publication that was highly relevant to the grant call in the four years before the grant release date. This finding shows that recent relevance can be used as a signal of expertise in the field and can be accepted as a signal in the 'sensemaking' process when looking for experts in the field.

We examine the correlations when considering the grant release and expiration date. Some grant calls are open for up to six years. For all grants that were open for three or more years, i.e., their expiration date is at least three years after their release



(a) NIH dataset (b) RU dataset

Figure 4: The α-relevance between a grant call for papers and the winning PIs' most similar paper in each of the years before the grant release date, denoted with 't'. Calculated separately for each researcher and the grant they won. Y-axis denotes the percentage of winning researchers with that similarity. Panel (a) depicts the findings for the NIH dataset and panel (b) for the 438 researchers who won competitive grants in the RU dataset.

date, we examined the correlation between a researcher's most relevant paper in each year and the grant. However, here we also examine the years between the grant release date and the expiration date. We do not have the winning time in the NIH dataset, and the review period differs greatly across grants, ranging between six to 20 months. Figure 5 shows the similarity between grants and the winning PIs' most relevant paper each year, from eight years before the release date up to the release year and the three years the call was still open. There is a clear trend of higher similarity in the years *since* a grant is released, indicating that researchers are more focused on that field during these years. This might indicate that grants drive interest, although more evidence is needed to determine that.

5.1 Relevance and productivity

We continue to examine the percentage of relevant publications for each researcher before winning the grant against their productivity. Here, we follow an aggregated view. We use the SPECTER transformer with various similarity measures in all the experiments reported here.

In a preprocessing step described in Algorithm 1 we calculate the cumulative cosine similarity counts and the total number of publications for each researcher who has won a grant, starting from the grant's release date. The information is accumulated in a *reverse order*, as seen in Algorithm 1 lines 6 and 7. For example, the calculation for the year prior to the release contains the percentage of α-relevant papers of the researcher in that year, and the calculation for the two years before the

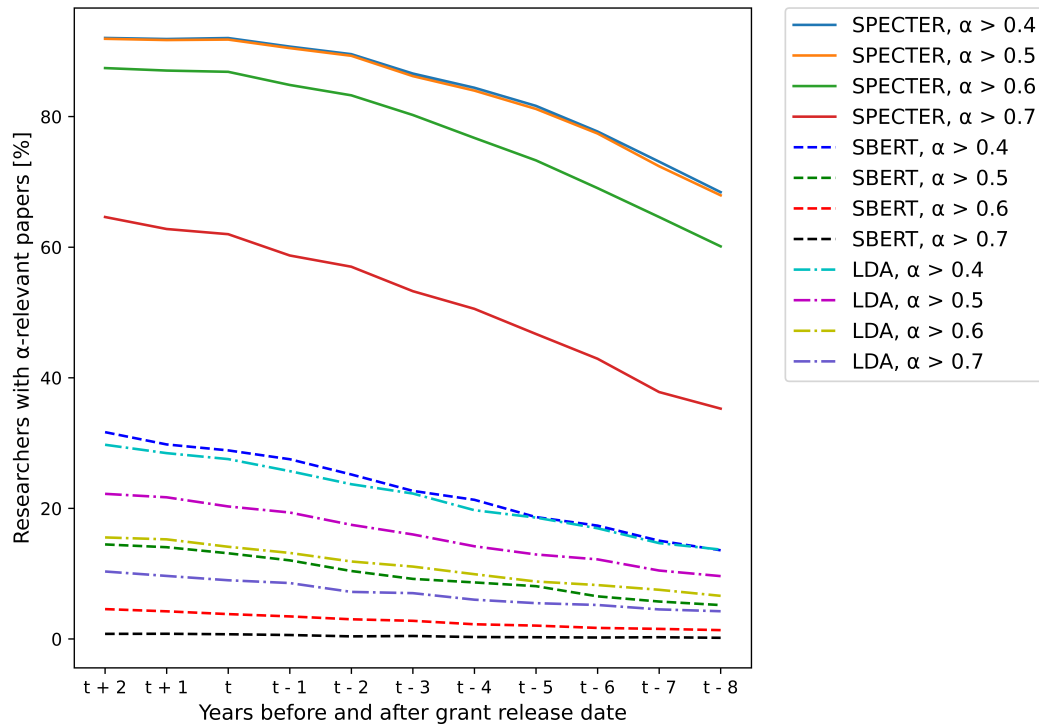


Figure 5: The α-relevance between the NIH grant call and the most similar paper each year of the PIs who won that grant, for 6805 single PIs. The longitudinal study examined this similarity for the eight years before the grant release date, denoted with t, and the three consecutive years in which the call is still open. Calculated separately for each researcher and the NIH grant they won. Y-axis denotes the percentage of winning researchers with that similarity.

grant release date corresponds to the aggregated percentage of α-relevant papers in these two years.

Figure 6 shows the accumulated $\alpha = 0.7$ relevance ratio of researchers' publications to the grant in the years before the grant's release date. That is, the percentage of papers the researchers published in the x-years before the grant release date that are very relevant to the call they won. As the information was accumulated over time, we could plot histograms for up to six years before the selected date and observe changes in the distribution. Here, we are looking for high-relevance values. On average, we see that only a small fraction of the winning researchers focus more than half of their research papers specifically on the subject of the grant. Thirty percent of the researchers do not have any paper that is $\alpha = 0.7$ relevant to the grant they have won during any of the six years prior to the grant. We then relax the requirement to having a similarity of $\alpha = 0.5$ and calculate the aggregated number

Algorithm 1 Alpha-relevance vs. productivity

$\alpha[K][6] = 0$
$published[K][6] = 0$
$y_0 \leftarrow$ Grant release date
for j in range $[0, 6]$ **do**
  for each researcher k in K **do**
    $y \leftarrow y_0 - j$
    if $j \neq 0$ **then**
      $\alpha[k][j] \leftarrow \alpha_k[j-1] + AlphaSim(p_{k,y}, g_k)$
      $published[k][j] \leftarrow published[k][j-1] + count(p_{k,y})$
    else
      $\alpha[k][j] \leftarrow AlphaSim(p_{k,y_0}, g_k)$
      $published[k][j] \leftarrow count(p_{k,y_0})$
    end if
  end for
end for

of $\alpha = 0.5$ relevant papers winning scholars published in the topic of the grant they won in previous years. Figure 7 depicts the percentage of PIs with that number of relevant publications in the x-years prior to the grant release date, aggregated the same way as before. For visual clarity, outliers with more than 40 relevant publications were omitted from the graph. We can see that less than 1% of the researchers do not have any $\alpha = 0.5$ relevant papers in the four years before the grant release date, and it is typical to have between two to 25 relevant publications in the four years before winning the grant.

Table 1 summarizes the percentage of researchers with α-relevance of 0.7 and 0.5, respectively, at year $i, i \in [1 \dots 5]$ before their grant submission.

Table 1: The percentage of researchers with α-relevance of 0.7 and 0.5

	One year prior to release date	Two years prior to release date	Three years prior to release date	Four years prior to release date	Five years prior to release date
$\alpha-$relevance = 0.7	59.84%	70.24%	74.68%	77.20%	78.73%
$\alpha-$relevance = 0.5	91.07%	97.08%	98.61%	99.18%	99.63%

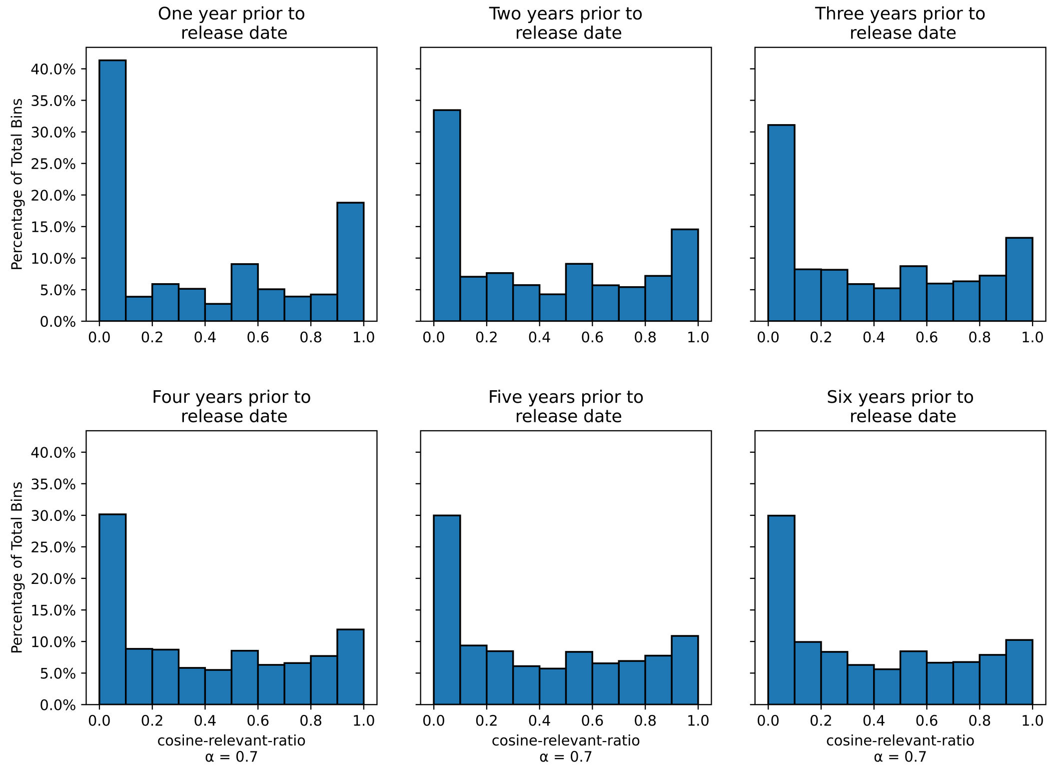


Figure 6

5.2 Is $\alpha-$relevance a fair signal?

We have established that people that win grant proposals have highly relevant papers in their recent publication track. However, is it a fair signal? it is plausible that grant winnings are biased. Would it bias this signal as well? We hypothesize that the signal is fair, as it does not predict winnings, but relevance to a field. To check whether the signal is fair, we examine the recent $\alpha-$relevance of PIs whose grant proposals were rejected. We hypothesize that PIs would submit to opportunities that match their expertise, hence that we would identify high expertise also in the population of PIs whose grant proposals were rejected.

Figure 8 shows the α-relevance between a grant call for papers and the PIs' most similar paper each year before the grant release date, for PIs whose proposals won and PIs whose proposals were rejected. The analysis was performed using PIs who submitted proposals to a list of competitive grants. The relevance of the signal of the recent publications is high in both cases, showing that our hypotheses hold, and the α-relevance is a fair signal of relevance and expertise.

Interestingly, we see that on average, the relevance of papers of PIs that won

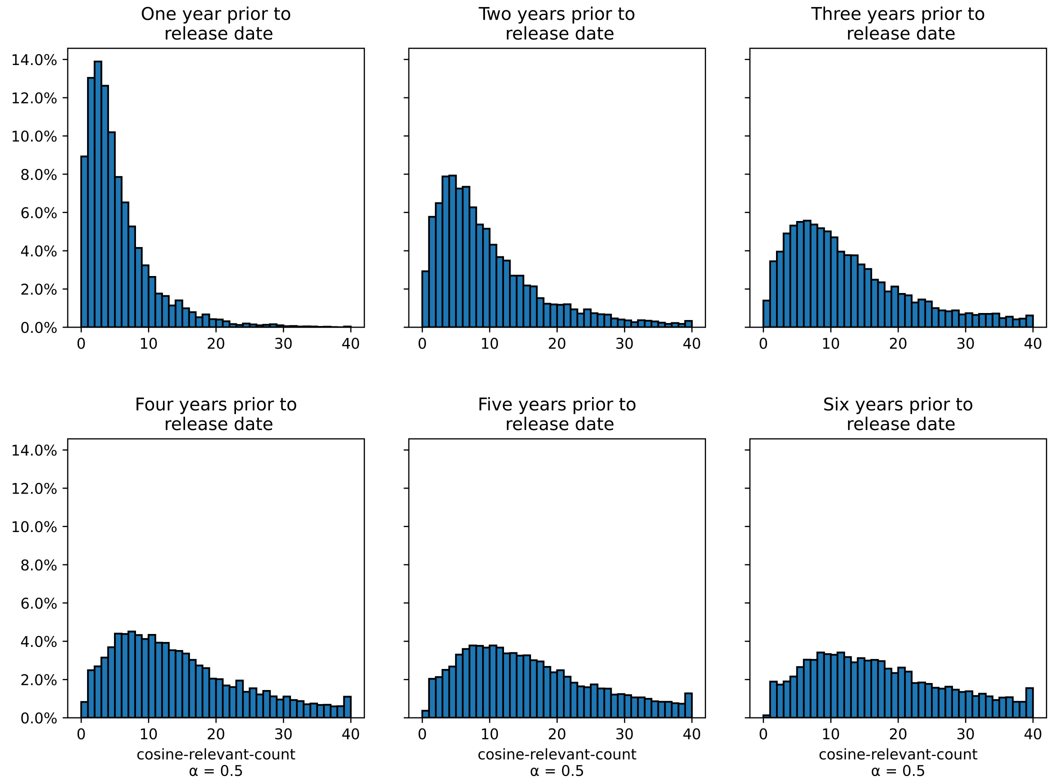
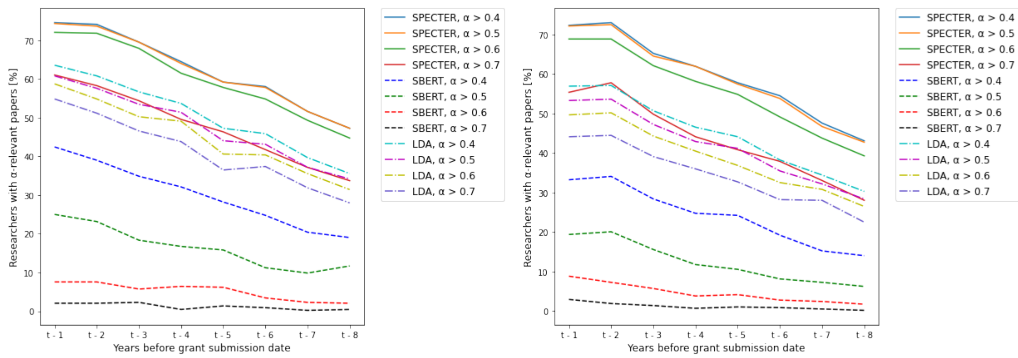


Figure 7

the grant is a bit higher. The relevance declines faster, on average, for PIs whose proposals were rejected.

6 Discussion and Conclusions

Departing from seniority, impact, and prestige signals of expertise, we explored the problem of determining expertise using 'makesense' signals as part of the decision-making process. This decision-making process is part of the job of PIs constructing multidisciplinary teams and of research authorities and institutes who are asked to form these teams or recommend participants. We suggest using a researcher's recent publications' relevance as a signal for this process. To that end, we investigated the recent publication relevance of PIs who won grants to the grant's call. Our longitudinal study established that the researcher's very recent history suffices for establishing the signal. We devised a metric, α-relevance, between a PI's recent publications and a grant call. A high α-relevance measure is an indication of expertise in the area of the grant. A sensitivity test showed that when the SPECTER



(a) Relevance of papers of PIs whose proposals won

(b) Relevance of papers of PIs whose proposals were rejected

Figure 8: RU dataset: The α-relevance between a grant call for papers and the PIs' most similar paper each year before the grant release date, denoted with 't'. Calculated separately for each researcher and the grant they won. Y-axis denotes the percentage of winning researchers with that similarity. Panel (a) depicts the α-relevance for PIs whose proposals won and Panel (b) depicts the α-relevance for PIs whose proposals were rejected.

transformer trained over scholar data was used to create the embedding, a cosine-similarity of $\alpha = 0.5$ over the recent four years of a scientist's publications establishes a strong signal, and it holds for scholars of different productivity levels. A team leader can use this measure to find collaborators in unknown fields.

One of the ways to depart from the current hierarchical structures of academia is to enlarge the set of signals used to determine expertise. In this work, we explore the relevance of recent publications as a signal in this process and show that it is valid and can be used successfully to signal expertise for either prolific or less prolific scientists. We showed that the signal is strong also for PIs whose proposals were rejected, although on average, their papers were slightly less relevant and they had highly relevant papers for less years before the grant proposal date. In our work, we did not consider traditional signals, such as citation counts, institutions, awards, or other known impact measures, known for increasing the inherent bias in academia.

In summary, we showed that an $\alpha-$*relevance* metric of over 0.5 between the trained SPECTER embedding over the titles and abstracts of a researcher's recent publications and the embedding of a call is a fair signal of expertise in the area described by the call.

References

[1] D. W. Aksnes, L. Langfeldt, and P. Wouters. Citations, citation indicators, and research quality: An overview of basic concepts and theories. *Sage Open*, 9(1):2158244019829575, 2019.

[2] I. Beltagy, K. Lo, and A. Cohan. Scibert: A pretrained language model for scientific text. *arXiv preprint arXiv:1903.10676*, 2019.

[3] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.

[4] T. Bol, M. de Vaan, and A. van de Rijt. The matthew effect in science funding. *Proceedings of the National Academy of Sciences*, 115(19):4887–4890, 2018.

[5] L. Bornmann and H.-D. Daniel. What do citation counts measure? a review of studies on citing behavior. *Journal of documentation*, 2008.

[6] K. J. Boudreau, E. C. Guinan, K. R. Lakhani, and C. Riedl. Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. *Management science*, 62(10):2765–2783, 2016.

[7] M. A. Burgman, M. McBride, R. Ashton, A. Speirs-Bridge, L. Flander, B. Wintle, F. Fidler, L. Rumpff, and C. Twardy. Expert status and performance. *PloS one*, 6(7):e22998, 2011.

[8] M. T. Chi. Two approaches to the study of experts' characteristics. *The Cambridge handbook of expertise and expert performance*, pages 21–30, 2006.

[9] A. Cohan, S. Feldman, I. Beltagy, D. Downey, and D. S. Weld. SPECTER: Document-level Representation Learning using Citation-informed Transformers. In *ACL*, 2020.

[10] L. Cronk. The application of animal signaling theory to human phenomena: some thoughts and clarifications. *Social science information*, 44(4):603–620, 2005.

[11] M. De Rond and A. N. Miller. Publish or perish: Bane or boon of academic life? *Journal of management inquiry*, 14(4):321–329, 2005.

[12] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

[13] J. Donath. Signals in social supernets. *Journal of computer-mediated communication*, 13(1):231–251, 2007.

[14] K. A. Ericsson and T. J. Towne. Expertise. *WIREs Cognitive Science*, 2010.

[15] M. Fire and C. Guestrin. Over-optimization of academic publishing metrics: observing goodhart's law in action. *GigaScience*, 8(6):giz053, 2019.

[16] S. Fortunato, C. T. Bergstrom, K. Börner, J. A. Evans, D. Helbing, S. Milojević, A. M. Petersen, F. Radicchi, R. Sinatra, B. Uzzi, et al. Science of science. *Science*, 359(6379):eaao0185, 2018.

[17] A. M. Harris, D. Gómez-Zará, L. A. DeChurch, and N. S. Contractor. Joining together online: the trajectory of cscw scholarship on group formation. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–27, 2019.

[18] M. Hertzum. Expertise seeking: A review. *Information processing & management*, 50(5):775–795, 2014.

[19] R. R. Hoffman, B. Crandall, and N. Shadbolt. Use of the critical decision method to elicit expert knowledge: A case study in the methodology of cognitive task analysis. *Human factors*, 40(2):254–276, 1998.

[20] T. A. Hoppe, A. Litovitz, K. A. Willis, R. A. Meseroll, M. J. Perkins, B. I. Hutchins, A. F. Davis, M. S. Lauer, H. A. Valantine, J. M. Anderson, et al. Topic choice contributes to the lower rate of nih awards to african-american/black scientists. *Science advances*, 5(10):eaaw7238, 2019.

[21] O. Husain, N. Salim, R. A. Alias, S. Abdelsalam, and A. Hassan. Expert finding systems: A systematic review. *Applied Sciences*, 9(20):4250, 2019.

[22] B. F. Jones. The rise of research teams: Benefits and costs in economics. *Journal of Economic Perspectives*, 35(2):191–216, 2021.

[23] C. Jones. Signaling expertise: How signals shape careers in creative industries. *Career creativity: Explorations in the remaking of work*, pages 209–228, 2002.

[24] K. Lenz. The presentation of self in everyday life. In *Goffman-Handbuch: Leben–Werk–Wirkung*, pages 267–274. Springer, 2022.

[25] A. Levy, M. Agrawal, A. Satyanarayan, and D. Sontag. Assessing the impact of automated suggestions on decision making: Domain experts mediate model errors but take less initiative. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2021.

[26] R. K. Merton. The matthew effect in science: The reward and communication systems of science are considered. *Science*, 159(3810):56–63, 1968.

[27] O. Mokryn, A. Wagner, M. Blattner, E. Ruppin, and Y. Shavitt. The role of temporal trends in growing networks. *PloS one*, 11(8):e0156505, 2016.

[28] M. W. Nielsen and J. P. Andersen. Global citation inequality is on the rise. *Proceedings of the National Academy of Sciences*, 118(7):e2012208118, 2021.

[29] S. A. Paul. Find an expert: Designing expert selection interfaces for formal help-giving. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 3038–3048, 2016.

[30] N. Reimers and I. Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*, 2019.

[31] N. S. Shami, K. Ehrlich, G. Gay, and J. T. Hancock. Making sense of strangers' expertise from signals in digital artifacts. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 69–78, 2009.

[32] M. Spence. Job market signaling. In *Uncertainty in economics*, pages 281–306. Elsevier, 1978.

[33] E. A. Tibbetts, J. Pardo-Sanchez, and C. Weise. The establishment and maintenance of dominance hierarchies. *Philosophical Transactions of the Royal Society B*, 377(1845):20200450, 2022.

[34] H. P. Van Dalen and K. n. Henkens. Signals in science-on the importance of signaling in gaining attention in science. *Scientometrics*, 64:209–233, 2005.

[35] K. H. Wapman, S. Zhang, A. Clauset, and D. B. Larremore. Quantifying hierarchy and dynamics in us faculty hiring and retention. *Nature*, 610(7930):120–127, 2022.

[36] Y. Xie. "undemocracy": inequalities in science. *Science*, 344(6186):809–810, 2014.

[37] H. Xu, Y. Bu, M. Liu, C. Zhang, M. Sun, Y. Zhang, E. Meyer, E. Salas, and Y. Ding. Team power and hierarchy: Understanding team success. *arXiv preprint arXiv:2108.04108*, 2021.

[38] S. Yarosh, T. Matthews, M. Zhou, and K. Ehrlich. I need someone to help! a taxonomy of helper-finding activities in the enterprise. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 1375–1386, 2013.

[39] A. Zahavi. Mate selection—a selection for a handicap. *Journal of theoretical Biology*, 53(1):205–214, 1975.