Gender Asymmetries in Academic Collaboration Networks

Sophia Pink, sopink@stanford.edu **Mayuka Sarukkai,** mayuka@stanford.edu

Abstract

This paper investigates differences in structural roles of male and female academics within collaboration networks, in order to identify the structural roles that contribute to their success. We find that successful men and successful women exhibit very similar network characteristics: they work with many collaborators, have collaborators who are not already collaborating with each other, and work with successful researchers. This suggests that the gap in research success by gender is a pipeline problem – successful men and women exhibit similar network characteristics, but women are less likely to have those characteristics than men, and women are underrepresented among researchers with the longest careers. While structural differences explain differences in success in most academic disciplines, gender is still significantly predictive of success disparities between male and female researchers in the field of computer science even after accounting for collaboration network features, suggesting that female researchers in computer science may face systematic barriers to success despite achieving collaboration network patterns comparable to their male counterparts.

1 Introduction

Female researchers face an immense problem: a citation gap. On average, female authors publish fewer papers than male authors, and papers by female authors receive fewer citations than those by men. This disparity has been found in STEM fields [5], political science [14], international relations [8], and even in majority-female journals [3]. Because promotions in academia are based on publication and citation counts, women are at a disadvantage when it comes to advancing their careers in research. Furthermore, research benefits from diverse teams and ideas, so systematically excluding women will likely reduce the quality and impact of scientific research.

A key part of the research process is academic collaboration and paper co-authorship, which propels the spread of new ideas. Given that there are many women who are successful in academia, we wonder if these women exhibit different collaboration network structures that allow them to succeed despite gender biases. In order to better understand these structural roles, we apply graphical network analysis to collaboration networks. While most analyses focused on gender asymmetries in academic research do find evidence of gender differences in citation and collaboration, few apply graphical network analysis to characterize research networks.

In this paper, we investigate differences in structural roles of male and female academics within collaboration networks, and identify the structural characteristics of female academics in collaboration networks that contribute to their success across different academic fields. We address the following research questions:

1. What are differences in success outcomes between male and female authors across academic fields?

2. What are structural differences between collaboration networks of male and female authors?

Preprint. Work in progress.

3. Which structural differences in collaboration networks are associated with the success of an author?

4. Do we observe these differences between the collaboration networks of successful men and successful women? Do these patterns differ across academic field?

Additionally, this study is a first step towards building a robust collaboration network containing node-level information about gender and research discipline. We hope that the creation of this dataset will be a valuable contribution to ongoing research on human collaboration, gender bias, and scientific productivity.

2 Related Work

Several studies have investigated factors that contribute to the difference in outcomes between male and female researchers. Studies have found key differences in citation and co-authorship behavior between male and female researchers, such as higher self-citation rates among men [7] and lower rates of first or last authorship positions for women [16]. Furthermore, women experiencing lower productivity and research impact are often associated with departments receiving less institutional support and resources [4].

Given the important role of team structure in the output and success of research work, past studies have focused on collaboration behaviors to understand these gender asymmetries. Past work has found strong evidence of gender homophily among male researchers – while men are more likely to prefer collaborating with other men [2], hiring other men [9], and citing other men [3], women are more egalitarian in their collaborative approach. In fact, Abramo et al. [1] find that women are better equipped to be collaborative in a research setting, and Rhoten and Pherman [10] suggest that women may hold a greater preference towards interdisciplinary collaboration than their male counterparts.

However, few studies have applied network analysis techniques to study these trends or identify network-level sources of gender asymmetry. Sarigol et al. [11] use a collaboration network in computer science to investigate whether the success of computer science papers is due to the collaboration network of the authors. More specifically, they investigate whether the centrality of authors in a coauthorship network can predict whether a given paper will be highly cited in the future. The researchers find that there is no single measure of centrality that is a good predictor of future citation success. However, it does not consider other features of nodes, such as subfield, gender, institution, and the length of an author's career.

Jadidi et al. [6] take a more nuanced approach by studying collaboration network patterns of more than one million computer scientists over a period of 47 years in order to better understand graph structural properties and longitudinal network changes associated with success in the field, as well as how these patterns vary between male and female scientists. The analysis generates node and collaboration network features and uses these to model success as measured by the h-index and citation impact of the author. They find that successful women and successful men are characterized by the same collaboration patterns – they collaborate with more colleagues, maintain longer lasting collaborations, and act as "brokers" between disjoint research communities. However, women are also less likely to adopt these collaboration patterns; they tend to drop out earlier in their careers, and they are more likely to embed within smaller, tightly knit communities rather than to act as a bridge between clusters.

While Jadidi et al. present a robust analysis of network patterns of male and female researchers in computer science, no study to date has extended this research to investigate collaboration trends across other disciplines.

3 Methods

3.1 Data

3.1.1 Network Definition

We construct a collaboration graph by aggregating collaborations between authors across a dataset of academic papers, where a collaboration two authors is defined as the occurrence of at least one paper

Table 1: Field Inference Counts

Field	Count
Computer Science and Engineering	3635
Math and Physics	3236
Chemistry and Material Science	7284
Economics, Business, and Management	2599
Biology and Medicine	13449
Arts and Humanities	1196
Social Sciences	2438
Unclassified	8606

co-authored between the two authors. Each node represents an author, and each edge indicates that the start and end node authors have co-authored at least one paper. The full network contains over 400 million author nodes and over 2 billion collaboration edges.

3.1.2 Co-authorship Data

We use the subset of the Open Academic Graph dataset [13], [12] containing all papers from the Microsoft Academic Graph (MAG), a large network of academic citations. This dataset contains a list of papers. We subset this dataset to include only papers published after 1990, and only papers with more than one author, in order to capture collaboration patterns of recent papers.

3.1.3 Gender Inference

To infer gender, we use the package genderperformr, a high-performing LSTM-based model that infers gender from usernames [15]. Gender was assigned only if the probability of a given label prediction was greater than 90%.

3.1.4 Academic Field Inference

We employ a naive approach to identify the academic discipline of a given author, in order to perform preliminary analyses on variations in success metrics across diverse fields. Using the list of journals for which an author has published, we identify the number of keyword matches using field-related keywords sourced from top journals in each of seven fields: biology, economics, physics and math, computer science, chemistry and materials sciences, arts and humanities, and social sciences. Each author is assigned to the field associated with the greatest number of keyword matches. Table 1 indicates the distribution of inferred academic fields.

3.1.5 Data Preprocessing

We select a subset of 54,385 nodes for which gender could be inferred and all other success metrics were available, and generate a subgraph of these nodes and their neighbors. Since this study focuses on current collaboration patterns rather than longitudinal analysis, this dataset is limited to all collaborations occurring after 1990.

We then calculate node characteristics for a subset of the nodes based on the following conditions:

1. Authors for which genderperformr makes gender prediction with greater than 90% certainty
2. Authors with 5 or more publications
3. Authors whose most recent paper was published after 1990
4. Authors with less than 4000 total publications and less than 90 publications per year on average, to account for outliers that were likely due to mis-entered data).

3.2 Network Metrics

To define success of a given author node, we use two variables available in the MAG dataset for all authors in the sample: number of publications, and number of citations. This section provides a survey of network metrics used as additional node-level features that we calculate.

Degree The degree of a node indicates the number of people with whom an author has collaborated.

Clustering Coefficient The local clustering coefficient of a single node provides a measure of how much an author's collaborators collaborate with each other. This measure indicates whether certain types of authors embed themselves in already highly connected collaboration communities. The local clustering coefficient c_i of node i is defined as follows:

$$c_i = \frac{e_{ab} : a, b \in N_i, e_{ab} \in E}{k_i(k_i - 1)}$$

where the numerator indicates the number of triangles in which node i participates, and k_i is the degree of node i.

Neighborhood Success In order to measure the extent to which authors collaborate with successful neighbors, we measure the mean and median number of citations and number of publications of all nodes adjacent to a given author.

Other Controls Additionally, we assign each node to its inferred gender and inferred discipline. We also use career age (based on the number of years between the first and last publication) as a control variable, to account for the relationship between a longer career and a higher publication or citation count.

3.3 Models

We fit a linear regression model predicting each success outcome (number of publications, number of citations) using network-based node features, as well as other node attributes. In this paper, we report results using the number of citations as the outcome variable, although models with number of publications as the outcome variable produced similar results.

In order to investigate whether successful men and women have structurally different collaboration networks, we observe whether gender is a significant predictor of success even after controlling for graph structural properties of collaboration networks, using the following model.

$$log(succ_i) = \alpha(g) + \beta log(k_i) + \gamma(CC) + \psi log(succ_{N_i}) + \delta(age) + \zeta(field) + int_g$$

where
$succ_i$ = number of citations of ego
g = gender
k = degree of ego
CC = clustering coefficient of ego
$succ_{N_i}$ = mean number of citations of all neighbors of ego
age = career age (last publication year - first publication year) of ego
$field$ = inferred academic field of ego
int_g = interaction terms between gender of ego and other node attributes

4 Results

4.1 Summary Statistics

Our sample includes 11,724 (23.9%) authors identified as female, and 37,313 (76.1%) authors identified as men. Table 2 shows the mean, median, and maximum values of our covariates across the entire sample.

4.2 Disparities in success by gender

We first look at summary statistics for success metrics by gender. Table 3 shows the average success metrics by gender. Without controlling for career age or other variables, we find that men have on average a higher number of publications than women, as well as a higher number of total citations than women.

Table 2: Summary of Node Attribute Variables

Covariate	**Mean**	**Median**	**Max**
Num. publications	25.37	11	2747
Num. citations	568.8	96	185136
Degree	71.38	21	25765
Clustering coefficient	0.43	0.37	1.0
Career age	12.4	9	85
First publication year	2003	2000	2018
Mean num. citations of collaborators	2034	1138.1	69246
Mean num. publications of collaborators	72.13	61.12	1678

Table 3: Success Metrics by Gender

	Male			**Female**		
Metrics	Mean	Median	SD	Mean	Median	SD
Num. publications	27	11	56	19	10	34
Num. citations	615	97	2633	421	95	1312

We also find that the most successful men have more citations than the most successful women. The 90th percentile of citations for men is 1169, whereas that 90th percentile for women is 883. Although the medians are similar, there are large disparities in success among the top authors.

Next, we calculate success statistics by gender and career age to account for the generally higher number of publications and citations associated with researchers who have been working for more years. We find that male and female early-career researchers have a similar number of publications and citations. However, after 20-25 years there does exist a success gap between late-career female and male researchers. See Figures and 1 and 2.

4.3 Network differences by gender

Next, we look at differences in network characteristics by gender. We observe that females have a higher median degree then men across career ages, while males have a higher mean degree. This suggests that overall, female authors tend to have a slightly higher number of collaborators. However, a small subset of highly connected male authors may be further skewing the degree distribution of male authors.

While the average success of neighbors is comparable across gender for early-career authors, Figure 3 indicates that males tend to collaborate with more highly published authors as compared to women later in their careers. However, the median number of publications of collaborators is consistent across genders, suggesting that male authors tend to develop collaborations with a few highly successful individuals as their careers progress more than female authors do. On the other hand,

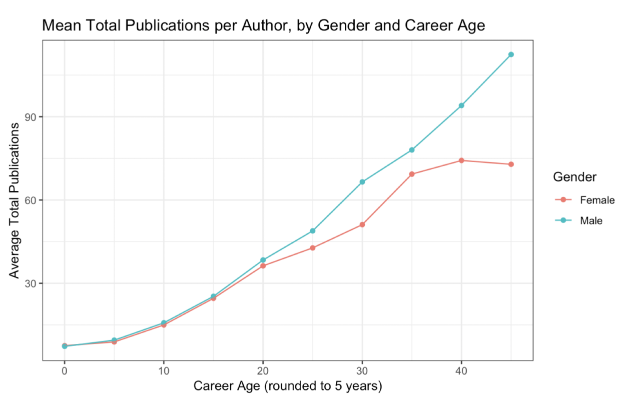

Figure 1: Total Publications per Author, by Gender and Career Age

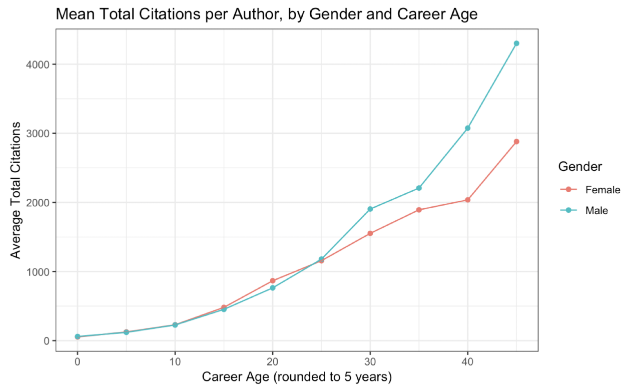

Figure 2: Total Citations per Author, by Gender and Career Age

Table 4: Network Features by Gender

Covariate	**Male**	**Female**
Degree (mean)	73.8	63.7
Degree (median)	20	23
Clustering coefficient (mean)	0.42	0.43
Neighborhood number of publications (mean)	70.9	76.0
Neighborhood number of publications (median)	40.6	43.0
Neighborhood number of citations (mean)	1940	2350
Neighborhood number of citations (median)	780	901

Figure 4 indicates that female researchers do tend to collaborate with more highly cited collaborators throughout their careers.

This result suggests that female authors are embedding themselves within larger networks and networks with more successful authors; however, despite this structural behavior, a disparity persists publication and citation counts for women, especially later in their careers.

4.4 Network predictors of success

Table 5 shows the correlations between success metrics (number of citations and number of publications) and network features. There is a significant correlation between all of the network characteristics and the success metrics.

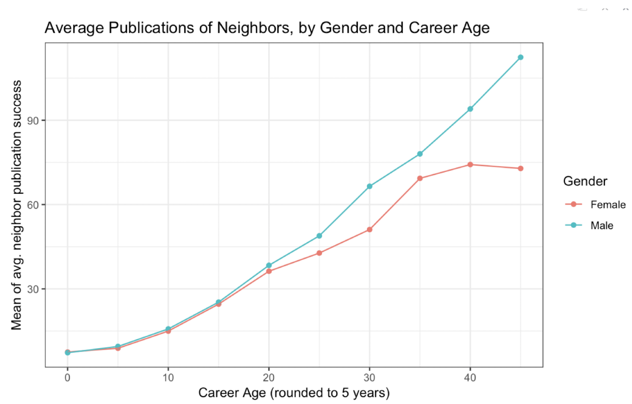

Figure 3: Average Publication of Neighbors, by Gender and Career Age

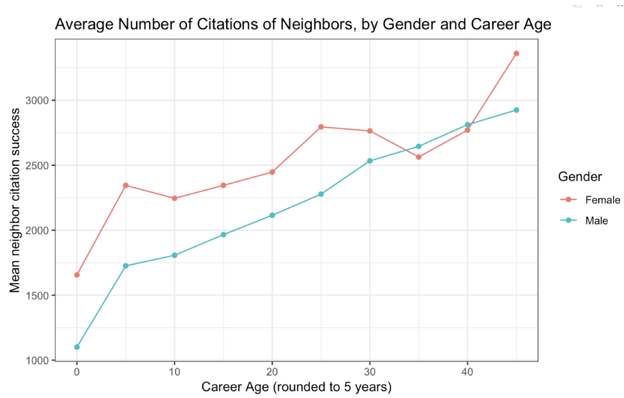

Figure 4: Average Citations of Neighbors, by Gender and Career Age

Table 5: Correlation Table of Network and Success Metrics

Covariate	Log(#pubs)	Log(#cit)	Log(degree)	Clust. coef.	Neigh. pub.	Neigh. cit.	Career age
Log(#pubs)	1.00	-	-	-	-	-	-
Log(#cit)	0.64	1.00	-	-	-	-	-
Log(degree)	0.65	0.56	1.00	-	-	-	-
Clust. coef.	-0.47	-0.35	-0.31	1.00	-	-	-
Mean # pubs of neigh.	0.14	0.28	0.27	-0.02	1.00	-	-
Mean # cit of neigh.	0.12	0.41	0.29	-0.06	0.71	1.00	-
Career age	0.59	0.48	0.31	-0.35	0.08	0.11	1.00

These results support empirical evidence that successful authors tend to have a greater number of collaborators, collaborate with more successful people, and often collaborate across authors and communities that may not be connected, as indicated by a lower clustering coefficient.

4.5 Differences in Gender and Success by Academic Field

In order to better capture variations in success by gender, we also investigate variations by academic field. Here, we perform a preliminary analysis of the distribution of gender and network features across academic disciplines.

Table 6 indicates that computer science, math/physics, and economics are the most male-dominated fields. Humanities and social sciences are more gender-balanced, although the majority of authors in these fields are still male.

Table 6: Percent of Authors who are Female, by Field

Field	**% of Authors who are Female**
Computer Science and Engineering	11.9
Math and Physics	14.6
Chemistry and Material Science	23.2
Economics, Business, and Management	24.6
Biology and Medicine	32.0
Arts and Humanities	32.9
Social Sciences	37.8

There also exist significant differences in network collaboration structure across fields. Authors in biology tend to co-author with more highly cited collaborators and collaborate with a significantly higher number of co-authors (mean degree = 128) compared to fields like CS (mean degree = 41) or the humanities (mean degree = 27). This difference is likely due to differences in co-authorship

Table 7: Model Results Predicting outcome variable log(number of citations)

Covariate	**Baseline**	**Network features**	**Network feat. + interactions**
Gender [reference = women] men	-0.05	0.13*	0.57*
Career age	0.058*	0.037*	0.036*
Log(degree)		0.44*	0.44*
Clustering coef.		-1.13*	-1.21*
Log(mean neigh cit.)		0.51*	0.56*
Career age*gender	-0.001	0.00	0.001
Log(degree)*gender			-0.003
Clustering coef.*gender			0.088*
Log(mean neigh cit.)*gender			-0.069*
Intercept	4.23*	-0.16*	-0.52*
R2	0.09	0.62	0.65

Significance: * $p < .0001$

Table 8: Model Results Predicting outcome variable log(number of citations) for authors with careers of at least 25 years

Covariate	**Network feat. + interactions**
Gender [reference = women] men	0.062
Career age	0.014**
Log(degree)	0.456***
Log(mean neigh cit.)	0.59***
Clustering coef.	-1.47***
Career age*gender	0.011*
Log(degree)*gender	0.017
Clustering coef.*gender	0.10
Log(mean neigh cit.)*gender	-0.049
Intercept	-.003

Significance: * $p < .05$, ** $p < .01$, *** $p < .001$

norms across fields, and suggests that academic field may be an important variable to include when drawing inferences about the relationship between gender, network structure, and success.

4.6 Structural differences between successful male and female authors

Next, we look at whether successful women exhibit different structural properties than successful men. We use a linear model with the log of total citations as the outcome variable. We use a binary gender variable as a covariate in our models. To determine whether there are differences by gender, we look a the coefficient on this binary variable as well as the coefficient on the interaction between gender and other variables.

First, we first create a naive model that uses only gender, career age, and the interaction between gender and career age as the covariates. This model shows no significant effect of gender on success.

Second, we create a model that adds the following network characteristics: log of total degree, clustering coefficient, and log of mean number of citations of neighbors. This model shows that there is an effect of gender. This suggests that after controlling for network effects, there is a significant boost to being male. However, this model does not look at interactions between gender and other variables.

The third model we create includes interactions between gender and all other covariates. Here, we see a significant effect of gender on success, but we also see a negative interaction term between male gender and neighbor citations. This indicates that having more successful collaborations is more strongly associated with female author's success than male author's success. In other words,

Table 9: Model Results Predicting outcome variable log(number of citations), by field

Covariate	**Biology**	**Comp. Sci.**	**Econ.**	**Human.**	**Math/Phys.**	**Soc. Sci.**
Gender [ref. = women] men	.34	1.23*	.50	.32	.99	-0.07
Career age	.03***	.034*	0.043***	.042***	.045***	.033***
Log(degree)	.51***	.59***	0.35***	.38***	.56***	.52***
Clustering coef.	-.99***	-1.25***	-1.04***	-1.56***	-1.76***	-1.41***
Log(mean neigh cit.)	0.58***	0.62***	.63***	.53***	.51***	.52***
Career age*gender	0.007*	.006	-.003	.01	-.006	0.02**
Log(degree)*gender	-.02	-.067	.056	.07	-.11	-0.08
Clustering coef.*gender	-.242	-.24	-.35	.30	.01	0.36
Log(mean neigh cit.)*gender	-.04	-.11	-.06	-.076	-.04	-.02
Intercept	-.943***	1.34*	-0.70*	-.06	-.71	-.07
R2	0.64	0.61	0.60	0.51	0.57	0.63

Significance: *p < .05, ** p < 0.01, *** p < .0001

either collaborating with successful academics is slightly more beneficial for women than for men, or women with the most citations have more collaborations with other successful researchers than the men with the most citations.

We also run the same model on the subset of researchers with career ages of over 25 years, given that we observe a larger success gap among the longest tenured researchers (Table 8). Here, we see no significant effects of gender on success, but we do see an interaction between gender and career age, with an additional year of research associated with a higher number of citations for men than for women. This suggests that among researchers with long tenures, successful men and successful women exhibit similar network characteristics, but men experience greater success as their careers advance.

4.7 Structural differences between successful male and female authors, by field

Finally, we investigate whether network structural differences can explain variation in success across genders for different academic disciplines. We run the same model with interactions as earlier, but on authors of each field separately. See Table 9 for model results.

We find that after controlling for network features, there is no significant association between gender and success metrics for authors in economics and the humanities. For authors in biology and social sciences, gender is not a significant predictor of success after controlling for network features. However, a significant positive interaction between male gender and career age does exist, suggests that male authors attain a greater increase in success with advancing career age than their female counterparts, in these fields. For authors in the computer science field, we find that even after controlling for network features and interactions, being male is still significantly positively associated with citation success. These findings indicate that despite engaging in collaboration patterns that are strongly predictive of success across fields, such as high degree of collaborators and high success of collaborators, women are still achieving fewer citations and writing fewer publications in the field of computer science than their male counterparts with comparable network structure. Computer science is the only academic field that exhibits such a pronounced effect.

5 Discussion

5.1 Conclusion

We find that successful men and successful women exhibit very similar network characteristics – they work with many collaborators (high degree), have collaborators who are not already collaborating with each other (low clustering coefficient), and work with successful collaborators (collaborators with a high average number of citations). However, we do find that working with successful collaborators is more strongly associated with success for women than for men.

We see the greatest gender gap in citations and publications among researchers with careers that have lasted more than 20 years. However, we find that among these researchers, there are no significant differences in network characteristics that determine success between successful men and successful women.

This suggests that the gender gap is a pipeline problem – successful men and women exhibit similar network characteristics, but women are less likely to have those characteristics than men. Only 14.5% of the researchers in our sample with at least 25 years of experience are female. In contrast, 25.3% of researchers with less than 25 years of experience are female. We also find that among late-career researchers, an additional year of research is associated with a greater increase in citations for men than for women. This suggests that late-career men may be more likely to occupy leadership roles where they are supervising research and therefore producing more research output.

Finally, a preliminary analysis of success predictors across academic field are mostly consistent with this finding that successful men and women have similar network patterns. However, a notable exception can be observed in the field of computer science: here, gender is still significantly predictive of success disparities between male and female researchers, even after accounting for collaboration network features. This finding suggests that in addition to a gender gap in the career pipeline, female researchers in computer science may be experiencing larger systematic barriers to success that persist even when women achieve network patterns associated with success.

Overall, we find that gender, career age, and a small number of network characteristics can explain 61% of the variance in success. This suggests that network properties are highly predictive of success, and network analysis offers a valuable approach to better understanding properties of successful academic research.

5.2 Limitations and Future Work

This study employs methods for gender inference and network measures on only a small subset of nodes from the graph – our next steps will involve analyzing network properties of all nodes in the graph.

This paper only uses collaboration data of papers published in the last twenty years, and does not capture larger trends in network properties and collaboration behaviors that may have changed over time. We have yet to understand whether disparities observed in late-career authors are due to career age, or due to historic differences in gender bias. Future models may consider capturing year of publication as an edge attribute in order to observe longitudinal changes in collaboration patterns.

Other key node and edge attributes should be included in future iterations of this graph. For instance, using unweighted edges fails to capture important variations in collaboration behavior such as the duration (number of years) of a collaboration, as well as the strength (number of co-authored papers). Other measures, such as betweenness centrality and more robust measures of structural holes, as well as gender homophily, will be valuable network-based characteristics to include in future iterations.

This study also employs gender prediction on a small subset of nodes, favoring Western names. Future studies may investigate more robust gender inference techniques that do not exclude non-Western names.

This work could also benefit from additional analyses of the roles of female authors within their academic communities, in order to better understand the persistence of publication gaps for women despite embedding within high-achieving networks. For instance, data on the position of women on a list of author names, or position within a lab or other research entity, may inform this analysis.

While this study presents a preliminary analysis of differences in network structure by discipline, we hope to build scalable representations of an author's academic field using deep embedding methods such as BERT, and by incorporating both journal and tag names for discipline prediction to identify fine-grained academic subfields with more precision. Our end goal is to construct a complete collaboration network with gender and discipline information, in order to contribute critical data to ongoing efforts to understand gender gaps in academic research.

References

[1] Giovanni Abramo, Ciriaco Andrea D'Angelo, and Gianluca Murgia. Gender differences in research collaboration. *Journal of Informetrics*, 7(4):811 – 822, 2013.

[2] Eduardo B. Araújo, Nuno A. M. Araújo, André A. Moreira, Hans J. Herrmann, and José S. Andrade, Jr. Gender differences in scientific collaborations: Women are more egalitarian than men. *PLOS ONE*, 12(5):1–10, 05 2017.

[3] Michelle L. Dion, Jane Lawrence Sumner, and Sara McLaughlin Mitchell. Gendered citation patterns across political science and social science methodology fields. *Political Analysis*, 26(3):312–327, 2018.

[4] Jordi Duch, Xiao Han T. Zeng, Marta Sales-Pardo, Filippo Radicchi, Shayna Otis, Teresa K. Woodruff, and Luís A. Nunes Amaral. The possible role of resource requirements and academic career-choice risk on gender differences in publication rate and impact. *PLOS ONE*, 7(12):1–11, 12 2012.

[5] Luke Holman, Devi Stuart-Fox, and Cindy E. Hauser. The gender gap in science: How long until women are equally represented? *PLOS Biology*, 16(4):1–20, 04 2018.

[6] Mohsen Jadidi, Fariba Karimi, Haiko Lietz, and Claudia Wagner. Gender disparities in science? dropout, productivity, collaborations and success of male and female computer scientists. *Advances in Complex Systems*, 21(03n04):1750011, 2018.

[7] Molly M. King, Carl T. Bergstrom, Shelley J. Correll, Jennifer Jacquet, and Jevin D. West. Men set their own cites high: Gender and self-citation across fields and over time. *Socius*, 3:2378023117738903, 2017.

[8] Daniel Maliniak, Ryan Powers, and Barbara F. Walter. The gender citation gap in international relations. *International Organization*, 67(4):889–922, 2013.

[9] Corinne A. Moss-Racusin, John F. Dovidio, Victoria L. Brescoll, Mark J. Graham, and Jo Handelsman. Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, 109(41):16474–16479, 2012.

[10] Diana Rhoten and Stephanie Pfirman. Women in interdisciplinary science: Exploring preferences and consequences. *Research Policy*, 36(1):56 – 75, 2007.

[11] Emre Sarigöl, René Pfitzner, Ingo Scholtes, Antonios Garas, and Frank Schweitzer. Predicting scientific success based on coauthorship networks. *EPJ Data Science*, 3(1):9, Sep 2014.

[12] Arnab Sinha, Zhihong Shen, Yang Song, Hao Ma, Darrin Eide, Bo-June (Paul) Hsu, and Kuansan Wang. An overview of microsoft academic service (mas) and applications. In *Proceedings of the 24th International Conference on World Wide Web*, WWW '15 Companion, pages 243–246, New York, NY, USA, 2015. ACM.

[13] Jie Tang, Jing Zhang, Limin Yao, Juanzi Li, Li Zhang, and Zhong Su. Arnetminer: Extraction and mining of academic social networks. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '08, pages 990–998, New York, NY, USA, 2008. ACM.

[14] Dawn Langan Teele and Kathleen Thelen. Gender in the journals: Publication patterns in political science. *PS: Political Science and Politics*, 50(2):433–447, 2017.

[15] Zijian Wang and David Jurgens. It's going to be okay: Measuring access to support in online communities. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2018.

[16] Jevin D. West, Jennifer Jacquet, Molly M. King, Shelley J. Correll, and Carl T. Bergstrom. The role of gender in scholarly authorship. *PLOS ONE*, 8(7):1–6, 07 2013.