DOI:10.1145/3430803

Under optimistic projection models, gender parity is forecast to be reached after 2100.

BY LUCY LU WANG, GABRIEL STANOVSKY, LUCA WEIHS, AND OREN ETZIONI

Gender Trends in Computer Science Authorship

THIS ARTICLE PRESENTS a large-scale automated analysis of gender trends in the authorship of Computer Science literature. Specifically, we aim to address the following questions:

- How is gender balance among authors changing over time?
- When might gender parity be reached among authors?
- How is gender associated with co-authorship?
- And how does Computer Science compare against other fields of study?

We answer these questions by performing an automated study of literature metadata from scientific conferences and journals, using data from the Semantic Scholar academic search engine.[a] Our study incorporates metadata from 11.8M Computer Science publications. To provide a basis for comparison, we also analyze more than 140M articles from other fields of study. Our results demonstrate that although progress has been made, there is still a significant gap in gender representation among Computer Science authors. Continued delay in addressing the gender gap may perpetuate imbalances for generations to come.

Data

Our analysis was performed over the Semantic Scholar literature corpus.[2] The corpus contains publications between 1940 and the end of November 2019, and associated metadata such as title, abstract, authors, publication venue, and year of publication. Metadata in Semantic Scholar are derived from academic publishers, as well as scientific repositories such as arXiv, DBLP, and PubMed. We use the 19 fields of study defined by Microsoft Academic,[25] which are integrated with Semantic Scholar data. Table 1 shows the distribution of articles used in our analysis by field of study.

The author list is extracted from all publications and compiled into a list of first names. We use Gender API[b] to perform gender lookup for each name. Gender API is a large online database of name-gender relationships derived by linking publicly available governmental data with social media profiles in

a https://www.semanticscholar.org/
b https://gender-api.com/

» key insights

- **If current trends hold, gender parity among Computer Science authors will not be reached in a century.**
- **Computer Science lags behind other fields of study in equal gender representation among authors.**
- **Given the magnitude and trends associated with this gender gap, policy changes may be necessary to address these disparities in the short term.**

IMAGE BY HALFPOINT

various countries. For each name, Gender API outputs the predicted binary gender (*female* or *male*), along with the accuracy associated with the prediction and the number of samples used to arrive at that determination. We exclude authors for whom first names are missing, and for whom only first initials are available. We also filter out first names that occur less than 10 times in our overall corpus, to reduce the number of API calls to manageable numbers.

Because many names are ambiguous with respect to gender, we use the accuracy returned by Gender API to represent the gender of each author as a distribution over male and female probabilities. For example, Gender API estimates the first name Matthew to be male with an accuracy score of 100, the maximum. The name Taylor, however, is estimated to be female but only receives an accuracy score of 55. These accuracies are used to generate two probabilities for each name, (m, f), where m is the probability of the associated author being perceived as male, and f is the probability of the associated author being perceived as female, where $m + f = 1$. In this example, each author named Matthew will be represented with the probability tuple (1.0, 0.0), and each author named Taylor will be represented as (0.45, 0.55).

We acknowledge that gender identity is fluid and nonbinary. However, for the sake of this large-scale study, we adopt a simplified view of gender as a probability distribution over two genders, relying on first names as a proxy for the author's perceived gender (as opposed to self-reported gender). We use Gender API's results as an estimation of authors' perceived binary gender, and use these estimates to generalize over our corpus. We are not making claims about any author's true self-reported gender.

Analyses

We perform two types of analysis on this data. First, we analyze publication trends, examining the number and proportion of female authors over time. To identify when gender parity may be reached, we project the proportion of female authors based on trends from the last 50 years (since 1970). In this article, we define parity as the proportion of female authors falling within 10% of 0.5, within the range of 0.45–0.55. We also study trends in co-authorship behavior as reflected in our data.

Authorship analysis. Most articles are authored by more than one individual. For the purposes of our analysis, each

Table 1. Corpus statistics for different fields of study.

Field of study	Total articles	Total author-article units	Average authors per article
Art	5.3M	7.4M	1.4
Biology	15.1M	55.2M	3.7
Business	3.7M	5.8M	1.6
Chemistry	14.7M	48.6M	3.3
Computer Science	**11.8M**	**27.3M**	**2.3**
Economics	3.8M	6.4M	1.7
Engineering	10.1M	20.9M	2.1
Environmental Science	2.0M	4.6M	2.3
Geography	4.0M	7.3M	1.8
Geology	3.2M	8.4M	2.6
History	6.0M	8.2M	14
Materials Science	7.4M	21.7M	2.9
Mathematics	5.5M	10.9M	2.0
Medicine	32.4M	111.9M	3.4
Philosophy	2.8M	3.9M	1.4
Physics	7.8M	31.0M	4.0
Political Science	4.9M	6.8M	1.4
Psychology	7.0M	14.7M	2.1
Sociology	4.6M	6.3M	1.4
Total	**152.1M**	**407.2M**	**2.7**

author-article pair is treated as one unit. An article with a single author yields one author-article unit; an article with three authors yields three author-article units, etc. In Computer Science, the average number of authors is approximately 2.3 per article. However, average authors per article have increased from approximately 1.5 per article in 1970 to approximately 3.0 in the past several years, which reflects patterns observed by other researchers.[11] Appendix B (available online at https://doi.org/10.1145/3430803) provides further discussion of this shift in relation to concurrent increases in author count in other fields.

The proportion of female authors over time is used to project the trend toward gender parity. The number of female authors in a given year is computed as the sum of probabilities f over the author-article units of that year, and the number of male authors is correspondingly generated as the sum of probabilities m. The proportion of female authors for each year F_t is computed as the number of female author-article units divided by the total number of author-article units for the corresponding year. We compute projections by performing an autoregressive integrated moving average (ARIMA) analysis, a widely used and established method for creating time series forecasting models.[4] ARIMA is an autoregressive forecasting technique, which means it uses historical values in a time series to predict current and future values. We use the auto ARIMA function in the R "forecast" package,[14] which automates the selection of ARIMA model order, with a preference for simple models with lower order.

We assume that the growth in female author proportion observes logistic behavior. The proportion of female authors is necessarily constrained between 0 and 1, and logistic growth assumes that a stable equilibrium will eventually be reached. We tested other fit functions (linear and exponential; see Appendix C at https://doi.org/10.1145/3430803 for details), but found them to be less suitable; the root-mean-squared-error (RMSE) of the logistic fit is lower than that of these other curve types when fitting to the growth curves of each field of study.

To perform the fit, we first apply σ_α^{-1}, the inverse of the α-scaled sigmoid (or logit) function $\sigma_\alpha(x) = \alpha/(1+\exp(-x))$, to map the gender proportion into the real number line so that the data is more amenable to linear approximation. We call α the expected equilibrium proportion parameter. This transform generates $y_t = \sigma_\alpha^{-1}(F_t)$, where F_t is the proportion of female authors per year. We then fit a nonseasonal ARIMA model with parameters p, d, and q for the transformed process y_t represented by the following equation:

$$\phi_p(B)(1-B^d)y_t = c + \theta_q(B)\varepsilon_t \qquad (1)$$

where B is the backshift operator, which shifts by one to the previous time point, and ε_t is zero-centered, normally distributed noise.[14]

Finally, we obtain the forecast in the original domain using a sigmoid transform over the projected values, applying σ_α to y_t for $t > 2019$. We sample α from the range [0.3, 1.0] so that σ_α has minimum and maximum values of 0 and α, respectively. This constrains the projected values to be between 0 and some expected equilibrium proportion defined by α. The 80% and 95% confidence intervals of the prediction are computed from averaging the projection confidence over 10000 iterations of model fitting.

The range for α is defined based on the space of likely equilibrium proportions, as estimated based on trends observed in various fields of study (see Figure 4). Note that α represents the proportion of female authors we expect in the long run. An equilibrium proportion of 0.5 indicates that we expect the authorship makeup to eventually stabilize at around 50% men and 50% women. An equilibrium proportion of 0.9 indicates that we expect the authorship makeup to eventually stabilize at around 10% men and 90% women. As we will elaborate later, we perform a sensitivity analysis to determine the effect of the selected α parameter on the year in which parity is expected to be reached.

Co-authorship analysis. Co-authorship is computed for each unique pair of author-article pairs for each article. If an article has n authors, $\binom{n}{2}$ co-author pairs are generated. Given one co-author pair (n_1, n_2) and associated gender probabilities $n_1 \rightarrow (m_1, f_1)$ and $n_2 \rightarrow (m_2, f_2)$, we compute three probabilities, p_{mm}, p_{mf}, and p_{ff}, corresponding to the gender combinations, that is., between two male authors, a male and a female author, and two female authors, respectively:

$$\begin{aligned} p_{mm} &= m_1 m_2 \\ p_{mf} &= m_1 f_2 + f_1 m_2 \\ p_{ff} &= f_1 f_2 \end{aligned} \qquad (2)$$

where $p_{mm} + p_{mf} + p_{ff} = 1$. The numbers of each type of co-author pair for each year are computed by summing over the above probabilities over all co-authorship pairs of that year.

We then assess the number of same-gender and different-gender collaborations over time. The results are measured as a deviation from the expected, where the expected co-authorships are determined by sampling from the numbers of female and male authors active in a given year, assuming the same number of collaborations per year as observed in our data. The total number of extra or missing collaborations is computed as the difference between the observed counts of each type of collaboration and the expected value. To show

rates of change, we also compute the ratio between observed and expected collaborations (O/E) of each type.

Results

Here, we discuss the main findings of our study.

Gender API results. The 152.1M articles in our corpus resulted in 407.2M author-article units. Of these author units, 14.5M lack first names, 110.0M have only a first initial, and 5.7M have a first name that occurs less than 10 times in the corpus. These author units are removed from further analysis. The remaining 277.0M author units are associated with 521K unique first names. We query these 521K names in Gender API, and acquire gender information for 351K; 170K names have insufficient information and are excluded from analysis. Of the 11.8M articles in Computer Science and the 27.3M author-article units therein, 24.1M authors have valid first names, and 16.9M author-article units (61.8%) resulted in associated gender information, which is higher coverage compared to authors in other fields (we acquire gender information for approximately 50.4% of authors across all fields).

Gender trends among authors. Figure 1 shows that the overall author count in Computer Science has increased substantially over the last several decades, as the field has experienced significant growth. The total number of author-article units in 2018 is above 1.2M. The proportion of female authors has also increased during this time.

Figure 2 shows the projected proportion of female authors in Computer Science. Residuals of the ARIMA fit line over the logit-transformed data appear normally distributed and are not significant under the Shapiro-Wilk Normality Test.[24] The proportion of female authors in Computer Science is predicted to reach 0.45 around 2124, more than 100 years from now. The upper bound of the 95% CI reaches 0.45 in 2065, and the lower bound of the 95% CI reaches 0.45 beyond the range of our projection. Appendix A (available online at https://doi.org/10.1145/3430803) provides further discussion on model choice and the sensitivity of ARIMA projections to the choice of the equilibrium parameter.

We also make the somewhat concerning observation that the rate of growth in female author proportion has slowed in recent years, visible in Figures 2 and 4. Our projection makes the optimistic assumption that the proportion will continue to grow towards or beyond parity, but the data may suggest otherwise. It remains to be seen whether a new trend is emerging that exhibits not an increase, but rather a leveling off or decrease in the proportion of female authors.

Association of gender and co-authorship. The numbers of same- (*male-male* or *female-female*) and cross-gender (*male-female*) co-authorships in Computer Science are computed for each year. Figure 3 shows the difference between the number of observed and expected collaborations of each type since 1990.[c] In this time period, there are more same-gender co-authorships than would be expected, and fewer cross-gender co-authorships than would be expected. In recent years, around 50000 cross-gender co-authorships per year were missing when compared to expected numbers.

The observed to expected ratio shows both optimistic and pessimistic collaboration trends. Although both men and women are more likely to co-author with authors of their own gender (positive O/E), the degree of same-gender bias is declining

c We show collaboration counts after 1990 because there is higher data volume in this period of time.

Figure 1. Gender of Computer Science authors over time, computed by averaging across gender probabilities in our dataset.

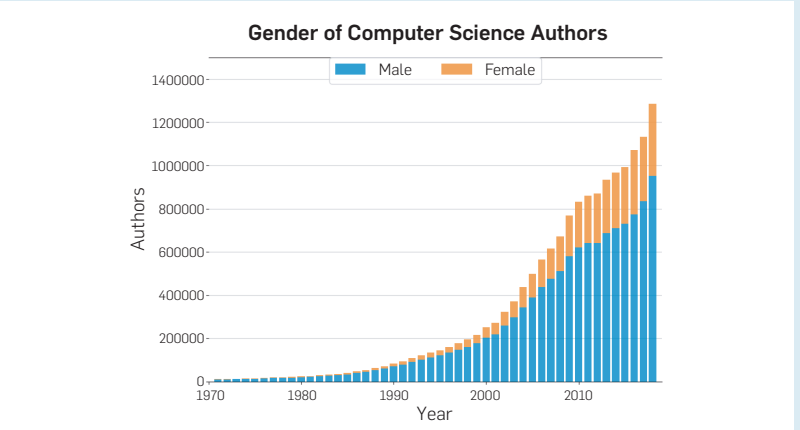


Figure 2. The proportion of female authors is projected using an ARIMA model assuming logistic growth toward equilibrium proportions in the range [0.3, 1.0]. Confidence intervals at 80% and 95% are shown.

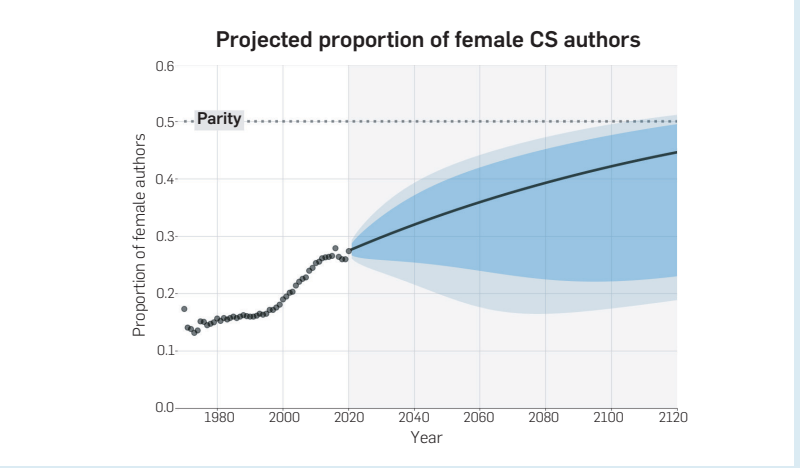
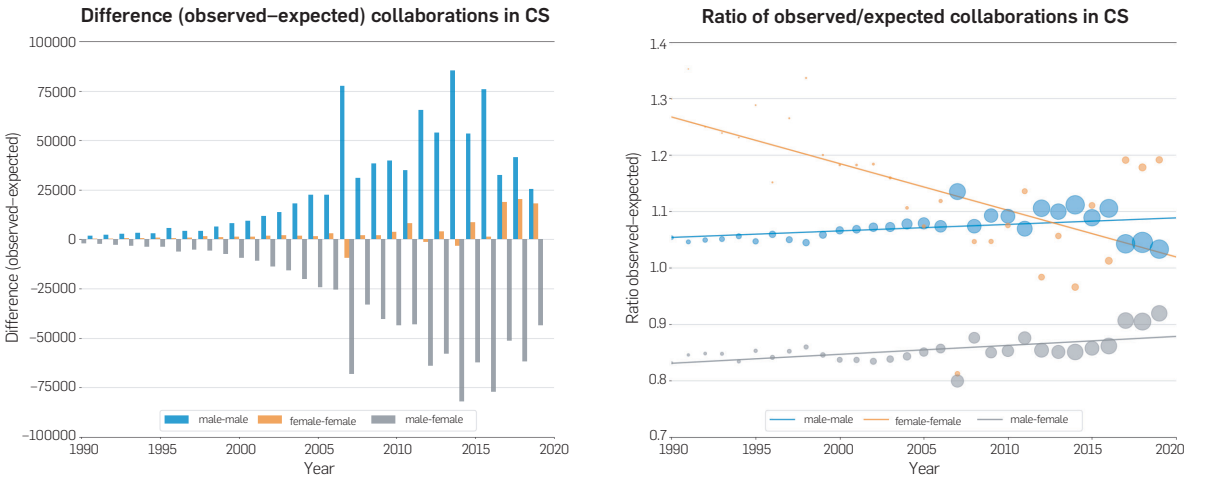


Figure 3. The difference (left) and ratio (right) between observed and expected same- and cross-gender co-authorships in Computer Science since 1990. Marker size for the O/E ratio is proportional to the number of expected collaborations of that type in each year.



among female authors but potentially increasing among male authors. At the same time, the cross-gender collaboration gap (O/E < 1.0) is still rather large, such that in recent years, only around 90% of expected cross-gender collaborations are observed. In other words, although there are more opportunities for cross-gender collaboration in recent years (due to an increase in the number of female scientists working in the field), the observed number of cross-gender collaborations is still below what would be expected. Optimistically, these trends may be shifting in the recent past, with numbers from the last three years showing a shift toward more cross-gender co-authorship; although it is too early to say whether this tendency will preserve itself in the future.

Comparison of CS with other fields of study. Figure 4 shows the proportion of female authors in 19 fields of study over the last 80 years. Computer Science is among the fields with the lowest female representation in recent years despite having relatively higher female representation in the middle of the 20th century.

Discussion

Our analysis of the Computer Science literature reveals the persistent patterns of inequality in gender and academic authorship. Although gender balance among authors is improving, progress is slower than we had hoped.

Limitations. Inferring gender from first names is imperfect, and all gender-inference tools are subject to biases. Several studies have described and measured the differences between these services.[15,22] Based on results in Santamaría and Mihaljević,[22] Gender API has the lowest overall error rate but was slightly biased toward underrepresentation of females in their evaluation; in other words, the number of women estimated may be slightly lower than in reality. However, this bias may be offset by our sampling bias, because the population of CS authors is unlikely to be an unbiased sample of the general population, or the population whose names were used to construct the database behind Gender API. We attempt to mitigate some of these biases by treating the perceived gender as a probability distribution. One way to compute a more precise estimate is to weight the probabilities assigned by Gender API to each name using the prior probabilities of being a female or male CS author; this would likely produce a more pessimistic projection.

The proportion of authors in our corpus with high uncertainty in Gender API results has also grown over time. The average confidence of our gender predictions decreased from around 95% in 1970–2000 to 90% since 2005. We show and discuss this change in confidence in Appendix D (available online at https://doi.org/10.1145/3430803). Although Gender API's average prediction confidence in our corpus is still high, this trend may pose a challenge for similar analysis in the future. Upon inspection of the data, we attribute this to the growing number of East Asian authors publishing in recent years. East Asian first names, when romanized, are more gender ambiguous. Gender API outperforms other gender lookup services, but still has lower overall confidence on names of East Asian origin.[22] In Mattauch et al.,[18] the authors explicitly exclude all authors with East Asian names from their name list during analysis, yet this accounts for the removal of more than 35% of their dataset. Rather than removing an entire group of authors from our data, we believe that representing each author name as a distribution of gender probabilities offsets some of the issues of increasing gender ambiguity in our corpus over time.

We also recognize the limitations of using author-article pairs as our units of measure. We do not distinguish between a person who is a single author on an article, and a person who co-authors with many others. This biases our data by overweighting articles with more authors. Similarly, in our analysis of collaboration, we take each combination of authors for an article as a collaborating pair, which again overweights

articles with more authors. In the Computer Science corpus, we observe an increase in the average authors per article over time, growing to approximately 3.0 authors per article in the last two years. However, Computer Science articles are still generally authored by smaller groups of individuals in the lower single digits, and we believe the bias introduced by our usage of author-article pairs or collaborating author pairs to be minimal.

Each author on a publication is also weighted equivalently in our analysis. We acknowledge that this discounts the special recognition extended to first authors, last authors, and single authors; we point readers to previous studies that have already demonstrated the distinctions between these groups.[27]

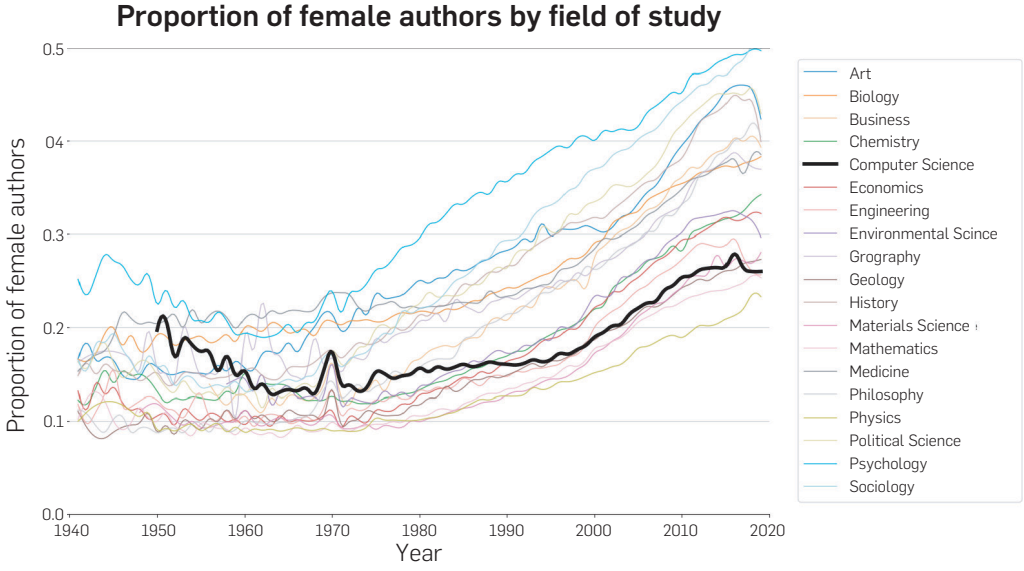
Lastly, our projection of female author proportion uses data from the last 50 years to project more than 100 years into the future. We understand the inaccuracies of making such an extensive forecast with limited data. The goal of our projection is not to provide a definitive answer to the question of when gender parity will be reached among Computer Science authors; rather, the projection signals that even under optimistic growth, the gender gap will likely not close in the near future without some form of community or external intervention. Observed recent trends also suggest that the increase in female representation among Computer Science authors may be slowing in the last five years. The long range forecasts we show may not adequately capture changes on this shorter time scale. Our forecasts also do not reflect changes that would result from newly introduced or as yet unimplemented interventions.

Prior work. Inequality in gender representation is a well-documented and studied issue in academia. Studies have shown that existent and perceived gender biases may affect many aspects of career and academic success, including but not limited to a woman's choice of college major,[21] crediting in scientific publications,[10] access to mentorship,[9,20,23] rate of promotions,[7] opportunities for collaboration,[1] as well as publishing and citation trends.[18,19] All of these factors can lead to imbalanced representation of women in certain fields of study.

With the increasing digitization of scholarly communication and availability of publication-related metadata, scholars have been able to better quantify inequality in authorship. Cohoon et al.[8] analyzed 86,000 ACM conference articles and showed increasing representation of women authors publishing at Computer Science venues, which strongly correlated with increasing numbers of female Computer Science PhDs.[8] West et al.[27] analyzed 1.8 million articles from JSTOR, a large multidisciplinary repository of academic literature, and revealed that although gender gaps are shrinking in academic publications, women were found to be significantly underrepresented as last and single authors. Elsevier, a large publisher of research articles, in an analysis of data from Scopus and ScienceDirect, reported the presence of gender imbalance among authors and inconsistent trends toward equal representation in different fields.[1] A study in 2018 confirmed continuing gender disparities among Nature Index journals, commonly considered some of the most reputable sources of academic literature, and in particular, limited representation of women among last authors, who are often perceived as more senior.[3] Our work demonstrates that the gender gap is persistent and relatively large among Computer Science authors, which is consistent with the results of these studies.

A study of gender bias in authorship conducted by Holman et al.[13] projected the closing of the gender gap in various fields based on recent trends. Through analyzing 9.1 million articles from PubMed, the authors projected that

Figure 4. The proportion of female authors among 19 fields of study. Proportion is plotted if there are more than 1,000 author-article units for which we could obtain gender information in a particular year.



gender parity would be reached in around 20 years in certain biomedical fields such as Molecular Biology, Medicine, or Biochemistry. Holman et al.'s analysis of a small corpus of Computer Science preprints from arXiv showed that gender parity in Computer Science will be reached in more than 100 years from the present.[13]. Also corroborating our estimate is related work from Way et al.,[26] which forecasts that gender parity in CS faculty hiring will be reached around 2075. Due to the long duration of faculty careers, parity in hiring would be expected to precede parity in publication and overall representation. Our results confirm and expand upon the results of this prior work. We use a significantly larger corpus of literature metadata to place the trends observed in Computer Science in the context of other fields of study. Additionally, we provide an assessment of co-authorship trends, which demonstrate a gap in cross-gender collaborations among CS authors.

Major strides have been made to reduce gender disparities. The presence of an overall structure of sexism in academia continues to be debated,[5,16,17] but many academic institutions recognize the issue and have sought to equalize admissions and hiring procedures. Evidence of movement toward more equitable representation in hiring and publication has been observed in some controlled settings.[6,12,28] How these observations translate into systemic change remain to be seen. Our results suggest, however, that the current pace of change in Computer Science will not result in a rapid closing of the gender gap.

Conclusion

We performed a large-scale analysis of the Computer Science literature (11.8M articles) to evaluate gender trends among authors. Based on trends over the last 50 years, the proportion of female authors in Computer Science is forecast to reach parity beyond the end of this century, and under different assumptions, it may take far longer. In this regard, Computer Science trails other fields of study, where we may want to look for inspiration. We also observed lower than expected numbers of cross-gender collaborations, with a gap of approximately 50000 cross-gender collaborations per year in the last several years.

Unless a major shift occurs that changes the gender makeup of the Computer Science community, the authorship gender gap will likely persist for a long time. Given the pervasiveness of computing technologies in our daily lives, it is of utmost importance that the researchers, designers, and builders of these technologies reflect the diversity of their users. Gender is one type of diversity among many that can be more easily assessed using the types of automated methods we employ. We hope that these findings will motivate members of the community to reflect upon the causes of these disparities, and provide evidence to back up policy decisions to change the status quo.

Acknowledgments

Thanks to Jonathan Borchardt, Matt Gardner, and Candace Ross for the initial analysis that motivated this work. Thanks to Kyle Lo for methodological discussions and Ashish Sabharwal, Maarten Sap, Noah Smith, and Mark Yatskar for helpful comments.

References

1. *Gender in the Global Research Landscape.* Technical Report. Elsevier, 2017. https://www.elsevier.com/research-intelligence/campaigns/gender-17.
2. Ammar, W., et al. Construction of the literature graph in semantic scholar. In *NAACL-HLT* (Orleans, Louisiana, June 1–6, 2018).
3. Bendels, M.H.K., Mueller, R., Brueggmann, D., Groneberg, D.A. Gender disparities in high-quality research revealed by Nature Index journals. *PLOS One 13*, 1 (2018), e0189136. https://doi.org/10.1371/journal.pone.0189136.
4. Box, G.E.P., Jenkins, G.M., Reinsel, G.C. *Time Series Analysis: Forecasting and Control*, 3rd edn. Prentice Hall, Englewood Cliffs, N.J., 1994.
5. Boynton, J.R., Georgiou, K., Reid, M., Govus, A. Gender bias in publishing. *The Lancet 392*, 10157 (2018), 1514–1515. https://doi.org/10.1016/S0140-6736(18)32000-2.
6. Ceci, S.J., Williams, W.M. Understanding current causes of women's underrepresentation in science. *Proc. National Acad. Sci. 108*, 8 (2011), 3157–3162. https://doi.org/10.1073/pnas.1014871108.
7. Clifton, S.M., Hill, K., Karamchandani, A.J., Autry, E.A., McMahon, P.J., Sun, G. Mathematical model of gender bias and homophily in professional hierarchies. *Chaos 29*, 2 (2019), 023135. https://doi.org/10.1063/1.5066450.
8. Cohoon, J.M., Nigai, S., Kaye, J. Gender and computing conference articles. *Commun. ACM 54*, (2011), 72–80. https://doi.org/10.1145/1978542.1978561.
9. Decastro, R., Griffith, K.A., Ubel, P.A., Stewart, A.J., Jagsi, R. Mentoring and the career satisfaction of male and female academic medical faculty. *Acad. Med.: J. Assoc. Am. Med. Colleges 89*, 2 (2014), 301–311. https://doi.org/10.1097/ACM.0000000000000109.
10. Feldon, D.F., Peugh, J.L., Maher, M.A., Roksa, M.A., Tofel-Grehl, C. Time-to-credit gender inequities of first-year PhD students in the biological sciences. *CBE Life Sci. Educ. 16*, 1 (2017), ar4. https://doi.org/10.1187/cbe.16-08-0237.
11. Fernandes, J.M., Monteiro, M.P. Evolution in the number of authors of computer science publications. *Scientometrics 110*, (2016), 529–539.
12. Hengel, E. 2017. Publishing while female. Are women held to higher standards? Evidence from peer review. In *Cambridge Working Article Economics*, 2017, 1753. Faculty of Economics, University of Cambridge. https://doi.org/10.17863/CAM.17548.
13. Holman, L., Stuart-Fox, D., Hauser, C.E. The gender gap in science: How long until women are equally represented? *PLOS Biol. 16*, 4 (2018), e2004956. https://doi.org/10.1371/journal.pbio.2004956.
14. Hyndman, R.J., Khandakar, Y. Automatic time series forecasting: The forecast package for R. *J. Statistical Software 26*, 3 (2008), 1–22. https://doi.org/10.18637/jss.v027.i03.
15. Karimi, F., Wagner, C., Lemmerich, F., Jadidi, M., Strohmaier, M. Inferring gender from names on the web: A comparative evaluation of gender detection methods. In *WWW* (Montreal, Canada, April 11 to 15, 2016). https://doi.org/10.1145/2872518.2889385.
16. Lundine, J., Bourgeault, I.L., Clark, J., Heidari, S., Balabanova, D. The gendered system of academic publishing. *The Lancet 391*, 10132 (2018), 1754–1756. https://doi.org/10.1016/S0140-6736(18)30950-4.
17. Lundine, J., Bourgeault, I.L., Clark, J., Heidari, S., Balabanova, D. Gender bias in academia. *The Lancet 393*, 10173 (2019), 741–743. https://doi.org/10.1016/S0140-6736(19)30281-8.
18. Mattauch, S., Lohmann, K., Hannig, F., Lohmann, D., Teich, J. A bibliometric approach for detecting the gender gap in computer science. *Commun. ACM 63*, (2020), 74–80.
19. Mohammad, S.M. Gender gap in natural language processing research: Disparities in authorship and citations. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (2020). Association for Computational Linguistics, Online, 7860–7870. https://doi.org/10.18653/v1/2020.acl-main.702.
20. Moss-Racusin, C.A., Dovidio, J.F., Brescoll, V.L., Graham, M.J., Handelsman, J. Science faculty's subtle gender biases favor male students. *Proc. National Acad. Sci. U.S. 109*, 41 (2012), 16474–16479. https://doi.org/10.1073/pnas.1211286109.
21. Robnett, R.D. Gender bias in STEM fields: variation in prevalence and links to STEM self-concept. *Psychol. Women Q. 40*, 1(2016), 65–79. https://doi.org/10.1177/0361684315596162
22. Santamaría, L.P., Mihaljević, H. Comparison and benchmark of name-to-gender inference services. *Peer J. Comput. Sci. 4*, (2018), e156. https://doi.org/10.7717/peerj-cs.156.
23. Schluter, N. The glass ceiling in NLP. In *EMNLP* (Brussels, Belgium from October 31st to November 4th, 2018). https://doi.org/10.18653/v1/D18-1301.
24. Shapiro, S.S., Wilk, M.B. An analysis of variance test for normality (complete samples). *Biometrika 52*, 3-4 (1965), 591–611. https://doi.org/10.1093/biomet/52.3-4.591.
25. Shen, Z., Ma, H., Wang, K. A Web-scale system for scientific knowledge exploration. In *Proceedings of ACL 2018, System Demonstrations* (2018), Association for Computational Linguistics, Melbourne, Australia, 87–92. https://doi.org/10.18653/v1/P18-4015.
26. Way, S.F., Larremore, D.B., Clauset, A. Gender, productivity, and prestige in computer science faculty hiring networks. In *WWW* (Montreal, Canada, April 11–15, 2016). https://doi.org/10.1145/2872427.2883073.
27. West, J.D., Jacquet, J., King, M.M., Correll, S.J., Bergstrom, C.T. The role of gender in scholarly authorship. *PLOS One 8*, 7 (2013), e66212. https://doi.org/10.1371/journal.pone.0066212.
28. Williams, W.M., Ceci, S.J. National hiring experiments reveal 2:1 faculty preference for women on STEM tenure track. *Proceedings of the National Academy of Sciences of the United States of America 112*, 17 (2015), 5360–5365. https://doi.org/10.1073/pnas.1418878112.

Lucy Lu Wang (lucyw@allenai.org) is a Postdoctoral Young Investigator at the Allen Institute for Artificial Intelligence, Seattle, WA, USA.

Gabriel Stanovsky (gabriel.stanovsky@mail.huji.ac.il) is a Senior Lecturer in the School of Computer Science and Engineering, The Hebrew University of Jerusalem, Israel. This work was done while at the Allen Institute for Artificial Intelligence and the University of Washington.

Luca Weihs (lucaw@allenai.org) is a Research Scientist at the Allen Institute for Artificial Intelligence, Seattle, WA, USA.

Oren Etzioni (orene@allenai.org) is Chief Executive Officer at the Allen Institute for Artificial Intelligence, Seattle, WA, USA.