Are AI Ethics Conferences Different and More Diverse Compared to Traditional Computer Science Conferences?

Daniel Acuna deacuna@syr.edu
Lizhen Liang lliang06@syr.edu
Syracuse University, School of Information Studies

Introduction

A great deal of research in artificial intelligence occurs in the context of computer science conferences. Recent cases of biases in AI systems have shook the community and society, and made these CS conferences self-reflect on the social implications of their work (O'Neil, 2016).

This introspection has motivated the community to create new conferences that address these issues. However, it stands to reason to ask whether these new "AI ethics" (AIE) venues have different representations from traditional CS conferences: real-world biased AI decisions mostly affect communities already underrepresented in CS. Are we, as a field, replicating past issues with CS conferences in AIE conferences? How are the authors, institutions, countries, contents, and citations different? Are researchers affected by biases in AI better, represented in these new venues? We hypothesize that we can use meta-science (Fortunato et al., 2018) to start answering some of these questions.

In this work, we study the characteristics of AIE conferences and contrast them to popular traditional CS conferences. We do so from the science of science point of view (Fortunato et al., 2018), which uses a variety of quantitative methods to study scientific processes and research behaviors using publication data, citation data, and author and affiliation statistics (Fortunato et al., 2018). We now describe the data and methods used in our study.

Data

We use the Microsoft Academic Graph (MAG) (Sinha et al., 2015), which contains exhaustive publication, citation, authorship, and affiliation data. We also use Semantic Scholar for content analysis, which includes the abstract of publications (Fricke, 2018). With MAG and Semantic Scholar, we can locate publications from our AI ethics and computer science conferences of interest.

For AIE conferences, we select FAccT and AIES for our analysis. For CS conferences, we choose the top 10 meetings by their combined impact and productivity. These CS conferences are AAAI, ACL, CVPR, ECCV, EMNLP, HLT-NAACL, ICCV, ICML, NAACL, and NeuIPS. From both data sets, we are able to find 381 publications from AIES and FAccT conferences from 2018 to 2020. To make comparisons relevant, we restricted publications from CS conferences to the same time frame. We found 14,179 publications from popular CS conferences, published between 2018 and 2020. Finally, we get geographical information about affiliations and author features related to citations such as h-index and fields of study from the data set. We retrieve features related to institutions and authors with publications accepted by those conferences from the affiliations.

Methods

Gender and race estimation. In order to investigate the trend of gender and ethnic diversity of authors, we built a BERT-based prediction model that predicts gender and race given the author's name.

Author's impact and productivity (seniority). By analyzing the citation network in MAG, we calculate the h-index for each author who published in AIE or CS conferences in any given year.

Inter-conference impact. With the citation network, we could locate publications in CS conferences citing publications in AIE conferences and vice versa. For CS conferences, we were able to get the percentage of citations to publications from AIE conferences, and for AIE conferences, we were able to get the rate of citations to publications from CS conferences.

Publication diversity. We are also able to find each author's publication in the last ten years and the corresponding field of each of those publications. For each author who has publications accepted by AIE conferences and CS conferences, we estimate an author's field by using the most common field. With the estimated areas of authors, we are able to measure the diversity of a conference by calculating the entropy of that conference given the estimated author fields.

Geographic diversity. We are able to reverse geocode the latitude and longitude of affiliation provided by MAG.
Conference content analysis. To further understand the differences between AIE and CS, we measure the differences between publications from different fields by measuring the use of words and topics.

#	AIE conferences		CS conferences	
1	Google	5%	Google	3%
2	Carnegie Mellon University	4%	Carnegie Mellon University	2%
3	IBM	4%	Chinese Academy of Sciences	2%
4	Stanford University	3%	Tsinghua University	2%
5	University of Oxford	2%	Microsoft	2%
6	Cornell University	2%	MIT	1%
7	Microsoft	2%	Stanford University	1%
8	Duke University	2%	Peking University	1%
9	University of Cambridge	2%	IBM	1%
10	UC, Berkeley	2%	UC, Berkeley	1%
	Others	65%	*Others*	80%

Top 10 institutions publishing from 2018 - 2020

#	AIE conferences		CS conferences	
1	computer science	46%	artificial intelligence	52%
2	artificial intelligence	24%	computer science	36%
3	mathematics	2%	mathematics	2%
4	medicine	1%	mathematical optimization	1%
5	economics	1%	algorithm	<1%
6	social media	1%	biology	<1%
7	psychology	1%	adversarial system	<1%
8	sociology	1%	medicine	<1%
9	biology	1%	discrete mathematics	<1%
10	mathematical optimization	1%	psychology	<1%
	Others	5%	*Others*	17%

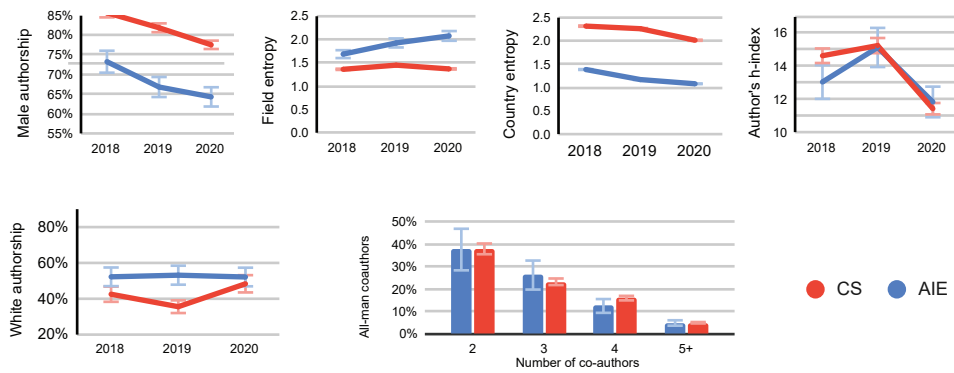
Top 10 fields publishing from 2018 - 2020

#	AIE conferences		CS conferences	
1	USA	74%	USA	42%
2	United Kingdom	9%	China	19%
3	Canada	2%	United Kingdom	4%
4	Netherlands	2%	Germany	3%
5	Germany	2%	Australia	2%
6	Brazil	1%	Canada	2%
7	Switzerland	1%	South Korea	2%
8	Norway	1%	Hong Kong	2%
9	New Zealand	<1%	Japan	2%
10	Spain	<1%	Switzerland	1%
	Others	6.21%	*Others*	18.19%

Top 10 countries (regions) publishing from 2018 - 2020

Science of Science & Computational Discovery Lab

Results



The top countries publishing in these venues are relatively similar. While China makes up 19% of the authorships in CS conferences and is in the top 2 positions after the USA, it does not appear in the top 10 countries in AIE conferences. There are only eight authorships affiliated with China from 2018 to 2020. We found that Google and CMU are at the top of both types of venues Companies tend to be more involved in AIE conferences, appearing at the top 1 (Google), top 3 (IBM), and top 7 (Microsoft). Facebook has a negligible presence in AIE conferences with only four publications. Chinese companies are also involved but only in CS conferences (Top 11: Tencent (not shown), Top 12: Facebook (not shown)). The top fields of publication. Expectedly, computer science and artificial intelligence are at the top. However, in AIE conferences, medicine, economics, psychology, and sociology made it higher in the ranking, suggesting broader applicability of the ideas presented. These rankings show that there are differences in representability between these types of conferences.

We measure field diversity using this entropy across AIE and CS conferences from 2018 and forward . Field diversity in AIE conferences grew significantly from 2018 to 2019 (two-sided bootstrap test [TSBT], $p < 0.01$) and also grew from 2019 to 2020 (TSBT, $p = 0.07$). The field diversity in CS conferences grew significantly from 2018 to 2019 (TSBT, $p < 0.0001$) but then it dropped significantly from 2019 to 2020 (TBST, $p < 0.0001$). Across years, AIE conferences have significantly higher field diversity than CS conferences (AIE: M=1.88, SE=0.5; CS: M=1.38, SE=0.10, $p < 0.0001$). The changes in field diversity in AIE might represent changes across CS rather than something specific about AIE. To test this hypothesis, we performed a difference-in-difference analysis (Cunningham, 2021), where we compare the changes in field diversity from 2018 to 2020 in AIE and CS. Our analysis showed a significant difference between these changes (z-score: 2.44, $p < 0.02$). This result suggests that AIE conferences have higher field diversity than CS conferences; they have been increasing this diversity differently from how CS has been changing its field diversity.

Similarly to how we measure field diversity, we measure the country diversity. We measure this diversity for the AIE and CS conferences from 2018 forward. We found that both AIE and CS conferences had a significant decrease in country diversity (AIE, $p < 0.001$; CS, $p < 0.001$). Moreover, we found that AIE conferences had significantly lower country diversity compared to CS conferences (z-score: -15.25, $p < 0.001$). We also found that AIE conferences had a decrease in diversity that was not significantly different from that of CS conferences ($p = 0.97$). These results suggest that AIE and CS conferences have decreased country diversity alike.

A percentage of 67.6 of the authors in AIE conferences were estimated to be male (SE=2.63%), while 78.8% of authors in CS conferences were estimated to be male (SE=1.44%, Figure 1c). Both types of conferences saw a large decrease in male authorship from 2018 to 2020 (AIE=-12%; CS:-9%), and this decrease was not significantly different among the conference types ($p = 0.568$), suggesting an overall increase in the number of non-male authorship in these venues. We found that CS conferences have more authors of other races in recent years (white: M=37.2%, SE: 1.1%. Asian: M=48.6%, SE = 1.6, z-score = -2.23, $p < 0.05$, 2018-2020). We found also that, on average, 52% of the authors in AIE conferences are white (SE=5.5%) and there is a significant increase in white authorship from 2018 to 2020 (TSBT, $p < 0.05$, Figure 1d). We also found that there are no significant differences in black authorship between CS conferences and AIE conferences (AIE black authorship: M=3.4%, SE=0.0053%, CS black authorship: M=2.7%, SE=0.063%). The number of black authors are not statistically different conferences (z-score: 0.46, p=0.646). Taken together, these results suggest the surprising result that AIE conferences have more white authorship than CS conferences.

To understand whether AIE conferences are having an impact on CS conferences, we performed a regression analysis to estimate the effect of year on the number of citations from AIE to CS papers. We found a positive effect, ($t(18)=2.09$, $p = 0.051$), suggesting an increase of AIE impact on CS. However, we found that AIE conference papers have reduced their citation to CS conference papers. Using a regression model analysis, we found a non-significant negative association between year and citation from AIE to CS ($t(3) = -0.684$, $p = 0.543$). These results suggest that CS conferences are increasingly citing AIE conferences while AIE conferences rely less and less on CS conferences. We then tested the hypothesis that people who publish CS conferences use AIE conferences as venues to publish something different than simply another CS article. We tested this hypothesis by evaluating how many people who publish in CS conferences also publish in AIE conferences. We found that only 0.92% of authors do this. However, we found that people who publish in AIE also publish in CS 25% of the time. This suggests that AIE conferences provide a venue for CS researchers to publish different work. Still, the fraction of scientists who publish at both conferences is very small (0.92%). This suggests that AIE is not taking people "away" from CS conferences, but rather, these conferences complement each other.

Syracuse University School of Information Studies