Analysis of Current and Future Computer Science Needs via Advertised Faculty Searches for 2021

Craig E. Wills
Computer Science Department
Worcester Polytechnic Institute
WPI-CS-TR-20-08
November 2020

Abstract

This work uses the same methodology as previous work to study where Computer Science departments are choosing to invest faculty positions using data obtained from advertised tenure-track searches for the current hiring season. This work also provides an opportunity to understand the effects of the COVID-19 pandemic on faculty hiring in Computer Science for hires starting in 2021.

We analyzed ads from 235 institutions seeking to fill hundreds of tenure-track faculty positions in Computer Science. This number is a 40% decrease from last year at this time (mid-November) and the lowest number in six years. The number of tenure-track positions sought shows a 47% decrease from last year and is at a similar level as six years ago. PhD institutions show a 50% one-year decrease where public PhD institutions are at a seven-year low and a 62% one-year decrease in the number of positions being sought.

We clustered the specific Computer Science topics mentioned in ads into 16 areas. In terms of specific areas, we found that the clustered area of AI/Data Mining/Machine Learning again accounts for 20% of all sought positions with Security again second at 18%. The area of Data Science is at 11% of positions, but aggregating the Data Science, AI/DM/ML and Databases clusters again resulted in roughly one-third of all hires sought in these data-oriented areas. The area of Theory/Algorithms again increased with 10% of all positions sought due to an increase in demand for the topic of Quantum Computing.

Differences are also seen when analyzing results based on the type of institution. Positions in the clustered area of AI/Data Mining/Machine Learning have the highest percentages for PhD institutions. Positions related to Security have the highest percentages for MS and BS/BA institutions. These two clustered areas are the two most sought areas for all types of institutions except for top-100 PhD institutions in which Theory/Algorithms is the second-most sought area.

We do plan to follow-up these results to understand if the severe reductions in searches observed up until mid-November is the result of non-existent or simply delayed searches for 2021. We plan to augment this report with those results in January 2021.
1 Introduction

The number of faculty searches in Computer Science during this hiring season for tenured and tenure-track positions starting in 2021 again affords the opportunity to study areas of Computer Science where departments are choosing to invest in new faculty hires. This is the seventh such report detailing results from a study of faculty hiring ads in Computer Science. It uses a similar methodology as done in previous years [2, 3, 4, 7, 9].

In addition, the longitudinal aspect of this work provides an early look at the effects of the COVID-19 pandemic on faculty hiring in Computer Science. A survey conducted by the Computing Research Association on the impact of COVID-19, which was conducted this summer and published in August [1], queried chairs of Computer Science academic units on many aspects of the pandemic impact including faculty hiring, but did not address it for the upcoming 2021 hiring season.

The focus of this work has always been to study where departments specifically, and the discipline more generally, are choosing to invest precious tenure-track faculty positions. It is an opportunity to understand where Computer Science departments think they are in terms of current needs as well as where they think they are going, but this year in particular it is an opportunity to see how many institutions are searching and with how many positions.

With this focus, there are a number of caveats to our study:

1. Our study is not exhaustive in that it does not necessarily take into account all searches currently underway for this hiring season. We describe the methodology used to discover ads, but ads may have been only placed in other venues or not have been placed in the timeframe of our study.

2. While our study focuses on preferred areas for faculty applicants, not all ads identify such preferred areas. These searches are accounted for in the data, but are not considered when analyzing particular areas of interest.

3. Similarly not all ads identify the specific number of positions being sought. In analyzing these searches we make an assumption on the number of positions being sought.

4. Our study analyzes searches and not hires. The number and areas of actual faculty hires may not match what is being sought.

2 Methodology

We used four primary sources for obtaining ads for Computer Science faculty positions: the Computer Research Association (CRA) Job postings\(^1\) the Association for Computing Machinery (ACM) list of jobs\(^2\), the Chronicle of Higher Education Vitae site\(^3\) and the HigherEdJobs site\(^4\).

We again augmented these sources with positions posted on the SIGCSE mailing list, which often includes ads for more undergraduate-focused institutions. We considered ads posted on these

\(^1\)https://cra.org/ads/
\(^2\)http://jobs.acm.org/jobs/search
\(^3\)https://chroniclevitae.com/job_search/new
\(^4\)https://www.higheredjobs.com/faculty/
venues between August 2020 and mid-November 2020, which is the same timeframe used in our previous studies. We recognize that the announcement of searches may be delayed this year due to the impact of the COVID-19 pandemic. By using the same timeframe for our study, we are not explicitly accounting for delayed searches, but we do address this potential impact in the future work portion of this report.

Only ads for tenured and tenure-track positions by departments containing Computer Science or closely-related programs were considered. We did not consider non-tenure-track positions such as lecturers, instructors or researchers and we only considered institutions awarding at least a BS or a BA degree. Searches for Deans or Department Chair positions were noted, but not considered because they do not reveal information regarding areas. Similarly, searches for other departments and programs with interest in faculty with Computer Science background were noted, but also not considered.

3 Results

3.1 Institutions and Positions

Using this methodology our resulting dataset contains information for faculty searches from 235 institutions (204 are in the U.S.). 182 (77%) of these institutions indicate a specific number of positions being searched for with the remaining searches using non-specific phrases such as “multiple positions,” “several positions” or just “positions” to indicate the number. As comparison, our previous-year study [10] found searches for 394 institutions (356 in the U.S.) with 77% of these institutions indicating a specific number of positions being searched for.

The left-side of Figure 1 shows all seven years of results for the number of institutions searching for tenure-track faculty. It shows a 40% decrease from last year at this time and the lowest number in six years.

In terms of the total number of positions, in the past we experimented with treating such “Multiple Position” searches as meaning two, three or four positions and settled on a value of three. In related work [5, 6, 8] where we surveyed institutions on their hiring outcomes we found those seeking multiple positions responded seeking a median of 3 and mean of roughly 3.5 positions. The right-side of Figure 1 shows the total number of positions searched for using a value of three for “multiple position” searches for a total of 413 positions. This value represents a 47% decrease from last year and is at a similar level as six years ago. We analyze the number of institutions and positions based on the type of institution later in the report.

Finally, in terms of institutions and positions we did encounter additional ads for Computer Scientists that were noted, but not considered in our analysis. We found seven Dean and Chair leadership searches (significantly down from 18 last year) as well as many faculty searches in other departments. These other departments include Electrical & Computer Engineering, Information School/Science/Technology, Bio-related, Health and Business. Ads found for these other departments were not considered in our analysis.
3.2 Results by Topic

In the same way that not all ads list a specific number of positions, it is also the case that not all ads list specific or preferred topics of interest. 177 (75%) of the 235 institutions listed specific topics, similar to the 73% from last year. In studying particular topics of interest, we only considered the ads from these institutions for our analysis.

In the initial step of our study, we determined the number of times that a specific topic was mentioned in an ad. Thus an ad for a single faculty position with preferred interest for the topics of HCI, Security, Machine Learning and Robotics would count one “mention” for each of these four topics. Another institution looking to focus three positions for the topic of Security would be one mention for Security. A total of 810 specific topics are mentioned in ads (versus 1357 last year).

While mentioned topics are one metric, another approach is to consider a faculty search as a “vote” for a topic of current and future need. Using this approach a single position with four topics of interest would be investing 0.25 positions for each topic, while three positions focused in a single topic would invest 3.0 positions in that single topic.

The problem with weighting topics based on the number of positions is that not all ads list a specific number of positions. We again use the fixed value of three for multiple-position searches resulting in a total of 413 “positions” being searched for with 312 (76%) of the positions indicating preferences for specific topics. Figure 2 shows the percentage of mentions and positions for topics with at least one percent for either mentions or positions. They are shown in rank order based on the percentage of positions.

---

5 We use the term “topic” to refer to sub-domains of Computer Science listed in ads and the term “area” to refer to a clustering of topics.
The results show that the topic of Security again accounts for the highest percentage of both mentions and positions, although it accounts for relatively more positions. Security was also the top topic for each metric the past four years. Data Science has the second highest percentage of positions with Machine Learning having the second highest percentage of mentions. The two topics are swapped for the third rank with AI ranked fourth for each metric. Other topics with smaller percentages, but significant increases relative to last year in the percentage of positions include Quantum Computing (again a big jump) and AI Fairness.

3.3 Clustering Topics into Areas

Figure 2 does not show topics that appear less frequently in ads nor does it group similar topics, such as Data Science and Data Analytics or Security and Privacy. To address these issues, we clustered topics into 16 areas. These clustered areas and the set of topics constituting the area are shown in Table 1. Topics with a small number of mentions and not clearly fitting into a cluster are included in two other clusters—one with topics in traditional Computer Science (OtherCS) and one with topics more interdisciplinary in nature (OtherInter). These are the same clustered areas as used in last year’s study [10].

3.4 Results Based on Clustered Areas

Given the clustered areas in Table 1, Figure 3 shows the same results as Figure 2 except it uses the 16 areas rather than the topics directly. The areas are again ordered by percentage of positions. For the second year in a row it shows that the AI/DM/ML clustered area has both the highest percentage
<table>
<thead>
<tr>
<th>Area</th>
<th>Constituent Topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI/DM/ML</td>
<td>Artificial Intelligence, AI Fairness, Data Mining, Deep Learning, Machine Learning Natural Language Processing, Optimization, Reinforcement Learning, Text Mining</td>
</tr>
<tr>
<td>Arch</td>
<td>Architecture, Hardware</td>
</tr>
<tr>
<td>Compiler/PL</td>
<td>Compilers, Programming Languages</td>
</tr>
<tr>
<td>CompSci</td>
<td>Biological Computing, Bioinformatics, Biometrics, Computational Biology, Computational Medicine, DNA, Neuroscience, Scientific Computation</td>
</tr>
<tr>
<td>DataSci</td>
<td>Big Data, Data Analytics, Data Science, Data Systems, Visual Computing, Visualization</td>
</tr>
<tr>
<td>DB</td>
<td>Databases, Data Management, Information Retrieval, Information Systems</td>
</tr>
<tr>
<td>HCI/IntMedia</td>
<td>Augmented Reality, Games, Human-Computer Interaction, Human Factors, Multimedia, Virtual Reality</td>
</tr>
<tr>
<td>ImageSci</td>
<td>Graphics, Image Processing, Vision</td>
</tr>
<tr>
<td>Mobile</td>
<td>Human-Centered Computing, Mobile Systems, Pervasive Computing, Ubiquitous Computing</td>
</tr>
<tr>
<td>Security</td>
<td>Block Chain, Cryptography, Forensics, Information Assurance, Malware, Privacy, Security, Trusted Computing</td>
</tr>
<tr>
<td>SoftEngr</td>
<td>Software Development, Software Engineering, Software Systems</td>
</tr>
<tr>
<td>Theory/Alg</td>
<td>Algorithms, Logic, Quantum Computing, Theory, Verification</td>
</tr>
<tr>
<td>OtherInter</td>
<td>Computer Engineering, Economics, Engineering Education, Health, Health Informatics, Race, Social Science, Sustainability</td>
</tr>
</tbody>
</table>
of mentions (19%) and positions (20%). The Security area again ranks second in percentage of mentions (15%) and positions (18%) with DataSci having the third highest percentage of positions (11%) with Sys/Net having the third highest percentage of mentions (13%). The Theory/Alg has the fourth highest percentage of positions at 10%.

![Clustered Area Percentage by Mentions and Positions](image)

Figure 3: Clustered Area Percentage by Mentions and Positions

The right-most clustered area in Figure 3 is the Data Oriented cluster that further aggregates results for the DataSci, AI/DM/ML and DB clusters. This aggregated cluster was introduced because of overlap between the three data-oriented clusters while still retaining the three distinct clusters as defined in Table 1. As shown in the figure, this aggregated cluster accounts for 32% of mentions and 34% of positions.

### 3.5 Results Comparison with Previous Years

Figure 4 shows a more complete comparison of clustered area results based on percentage of positions for the past five years of our studies. Clustered areas percentages for all years were determined based on the 2021 clustering of topics shown in Table 1. Clustered areas are ordered based on 2021 percentages.

The results show the percentage of positions are being targeted for AI/DM/ML hires has grown steadily over the five years and this area again has the highest percentage of positions for the 2021 hiring season. The Security area again ranks second with DataSci still ranked third. The Theory/Alg area, which includes the topic of Quantum Computing, took another jump this year and is ranked fourth in the percentage of positions. The Sys/Net area has stayed relative level around 10% across the five years. The aggregated DataOrient (AI/DM/ML, DataSci and DB) cluster is at a five-year high with a third of all positions being sought for these areas.
4 Results by Type of Institution and Highest Degree Offered

As means to better understand the results we augmented the dataset to include additional information about each institution.

4.1 Results by Type of Institution

Table 2 shows a breakdown of results based on whether the institution is public, private or non-U.S.-based. The results show that 106 (45%) of the institutions are public and account for 178 (43%) of the total positions. These percentages are down from 49% and 51% in last year’s study. There are 98 (24%) private institutions accounting for 160 (39%) of positions. The 31 non-U.S.-based institutions account for 8% of the total and 75 (18%) of positions.

Table 2: Summary of Position Searches by Institution Type

<table>
<thead>
<tr>
<th>Institution Type</th>
<th>Number of Institutions</th>
<th>Advertised Number of Positions</th>
<th>Total Positions</th>
<th>% Positions w/ Specific Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public</td>
<td>106</td>
<td>70 (66%) 9 (8%) 4 (4%) 23 (22%)</td>
<td>178</td>
<td>77%</td>
</tr>
<tr>
<td>Private</td>
<td>98</td>
<td>67 (68%) 13 (13%) 3 (3%) 15 (15%)</td>
<td>160</td>
<td>72%</td>
</tr>
<tr>
<td>Non-U.S.</td>
<td>31</td>
<td>11 (35%) 2 (6%) 3 (10%) 15 (48%)</td>
<td>75</td>
<td>79%</td>
</tr>
<tr>
<td>All</td>
<td>235</td>
<td>148 (63%) 24 (10%) 10 (4%) 53 (23%)</td>
<td>413</td>
<td>75%</td>
</tr>
</tbody>
</table>
4.2 Results by Highest Degree Offered

We also characterize each institution based on the highest degree it offers. For example, undergraduate-only programs may not have the same needs as PhD programs. For this portion of the study we augmented our dataset to include the highest degree offered by each program—BS/BA, MS or PhD. Our dataset includes 104 PhD institutions—significantly down from 183 last year. In order to study faculty investments at the most prominent U.S. programs, we further subdivided this group by using the U.S. News Rankings of the 100 Best Graduate schools\(^6\) for the top-100 U.S. and then more PhD institutions including those not in the U.S. The “PhD100” list accounts for 48 (vs. 82 last year) institutions in our dataset. The remaining PhD programs, including 29 non-U.S.-based, are denoted as “PhDMore.” Table 3 shows summary results based on the four highest degree types. The left-side of Figure 5 shows longitudinal results for the number of institutions searching over a seven-year period. The number of institutions for all types dropped this year with PhD100 and MS at seven-year lows.

Table 3: Summary of Position Searches by Highest Degree Offered

<table>
<thead>
<tr>
<th>Institution Type</th>
<th>Number of Institutions</th>
<th>Advertised Number of Positions</th>
<th>Total Positions w/ Specific Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>PhD100</td>
<td>48</td>
<td>12 (25%) 4 (8%) 6 (12%) 26 (54%)</td>
<td>136</td>
</tr>
<tr>
<td>PhDMore</td>
<td>56</td>
<td>30 (54%) 4 (7%) 4 (7%) 18 (32%)</td>
<td>111</td>
</tr>
<tr>
<td>MS</td>
<td>41</td>
<td>29 (71%) 7 (17%) 0 (0%) 5 (12%)</td>
<td>59</td>
</tr>
<tr>
<td>BS/BA</td>
<td>90</td>
<td>77 (86%) 9 (10%) 0 (0%) 4 (4%)</td>
<td>107</td>
</tr>
<tr>
<td>All</td>
<td>235</td>
<td>148 (63%) 24 (10%) 10 (4%) 53 (23%)</td>
<td>413</td>
</tr>
</tbody>
</table>

Table 3 reveals differences between the different types of institutions. Ads for 86% of the BS/BA institutions are for a single position while 54% of the ads for PhD100 institutions are for multiple positions, which is down from 65% last year. As shown, the distributions translate into a total number of 136 (down from 267 last year) positions for PhD100 institutions. We note that this number is particularly sensitive to the number of positions assumed for “multiple position” searches as over half of these searches are not specific in the number of positions being sought.

The right-side of Figure 5 shows seven-year results for the number of positions being searched for by the four types of institutions. It shows that the number of positions being sought by PhD100 and PhDMore institutions dropped by roughly 50% in 2021 with PhD100 and MS institutions at seven-year lows for positions sought.

The last column of Table 3 shows that only 57% of positions from BS/BA institutions identify specific areas of interest while 83% of PhD100 institutions do so with the percentages for PhDMore and MS institutions in between. In order to understand differences on areas of interest between different types of institutions for 2021 searches, we break down the results in Figure 3 based upon the type. Figure 6 shows the results (in the same rank order as Figure 3) grouped by the four types of institutions.

\(^{6}\)http://grad-schools.usnews.rankingsandreviews.com/best-graduate-schools/top-science-schools/computer-science-rankings
Figure 5: Seven-Year Counts of Institutions Searching and Positions Being Sought by Highest Degree Offered

Figure 6: Clustered Area Percentage by Highest Degree Offered
Figure 6 shows a number of interesting results. AI/DM/ML is of most interest for PhD100 and PhDMore institutions. Positions related to Security have the highest percentages for MS and BS/BA institutions. DataSci accounts for a higher percentage of positions for BS/BA institutions. Again, the Theory/Alg area is particularly strong for PhD100 institutions with this area having the second-highest percentage for this set of schools. Finally, the last set of results show that 35% of positions being sought by PhD100 institutions are for the Data Oriented aggregated cluster with the other type of institutions just a bit less.

4.3 Results by Combination of Institution Type and Highest Degree Offered

A third summary of positions searches is shown in Table 4 where institutions are classified based on a combination of type and highest degree offered. For this analysis, PhD100 and U.S. PhDMore institutions are combined as are U.S. MS and BS/BA (designated as “MB” in the table and subsequent graphs). The non-U.S. institutions are dropped in this analysis, but their results are shown in Table 2 as well as included in Table 3 and Figure 5.

Table 4: Summary of Position Searches by Institution Type and Highest Degree Offered

<table>
<thead>
<tr>
<th>Institution Type</th>
<th>Number of Institutions</th>
<th>Advertised Number of Positions</th>
<th>Total Positions w/ Specific Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1    2   3+ Multiple</td>
<td></td>
</tr>
<tr>
<td>Pub/PhD</td>
<td>45</td>
<td>23   1   4 (9%) 17 (38%)</td>
<td>96</td>
</tr>
<tr>
<td>Prv/PhD</td>
<td>30</td>
<td>10   5   3 (10%) 12 (40%)</td>
<td>78</td>
</tr>
<tr>
<td>Pub/MB</td>
<td>61</td>
<td>47   8   0 (0%) 6 (10%)</td>
<td>82</td>
</tr>
<tr>
<td>Prv/MB</td>
<td>68</td>
<td>57   8   3 (4%)</td>
<td>82</td>
</tr>
<tr>
<td>All U.S.</td>
<td>204</td>
<td>137  22  7 (3%) 38 (19%)</td>
<td>338</td>
</tr>
</tbody>
</table>

The results in Table 4 show that public institutions account for the majority of PhD-producing schools while there is a higher number of private MS&BS/BA institutions. 40% of private PhD institutions are searching for multiple positions while over 80% of PhD institutions identify specific topics on interest in their ads. Figure 7 shows seven-year results for the number of institutions searching and the number of positions sought based on this institution classification. The left side of the figure shows a one-year decline for all institution groups with a 51% one-year decline for public PhD institutions. It shows the number of public and private PhD institutions searching to be at seven-year lows. The right side of the figure shows the number of positions sought decreased in 2021 for each grouping of institutions. This drop-off is particularly striking for public PhD institutions where there is a seven-year low and a 62% one-year decrease in the number of positions being sought.

Figure 8 shows the percentage of positions for each of the clustered areas using this classification for institutions. The figure both shows similarities and differences with results shown in Figure 6. The area of AI/DM/ML accounts for 29% of the positions for private PhD institutions, 19% for public PhD institutions and a smaller percentage for MS&BS/BA institutions. In contrast, Security is again the area of most interest for both MS&BS/BA institution type combinations, but it is represented by a much lower percentage (8%) for private PhD institutions. There are fewer
Figure 7: Seven-Year Counts of Institutions Searching and Positions Being Sought by Institution Type and Highest Degree Offered

Figure 8: Clustered Area Percentage by Institution Type and Highest Degree Offered
differences between the four type combinations for DataSci, although it has the smallest representation (8%) for public PhD institutions. In contrast, Theory/Alg (20%) is the area with the highest representation for public PhD institutions and is the second-most represented area (with 14% of positions) for private PhD institutions. This area has low representation for MS&BS/BA institutions. The last set of results in Figure 8 shows that 42% of positions for private PhD institutions are in the Data Oriented cluster with 27% for public PhD, 31% for private MS&BS/BA, and 36% for public MS&BS/BA institutions.

5 Summary and Future Work

This work uses the same methodology as previous work to study where Computer Science departments are choosing to invest faculty positions using data obtained from advertised faculty searches for the current hiring season. This work also provides an opportunity to understand the effects of the COVID-19 pandemic on faculty hiring in Computer Science for hires starting in 2021.

This work uses the same methodology as previous work to study where Computer Science departments are choosing to invest faculty positions using data obtained from advertised tenure-track searches for the current hiring season. While the number of and areas for faculty searches does not necessarily translate into the same for faculty hires, we believe that they provide insight into current and future needs within the discipline.

We analyzed ads from 235 institutions seeking to fill hundreds of tenure-track faculty positions in Computer Science. This number is a 40% decrease from last year at this time and the lowest number in six years. The number of tenure-track positions sought shows a 47% decrease from last year and is at a similar level as six years ago. PhD institutions show a 50% one-year decrease where public PhD institutions are at a seven-year low and a 62% one-year decrease in the number of positions being sought.

We clustered the specific Computer Science topics mentioned in ads into 16 areas. In terms of specific areas, we found that the clustered area of AI/Data Mining/Machine Learning again accounts for 20% of all sought positions with Security again second at 18%. The area of Data Science is at 11% of positions, but aggregating the Data Science, AI/DM/ML and Databases clusters again resulted in roughly one-third of all hires sought in these data-oriented areas. The area of Theory/Algorithms again increased with 10% of all positions sought due to an increase in demand for the topic of Quantum Computing.

Differences are also seen when analyzing results based on the type of institution. Positions in the clustered area of AI/Data Mining/Machine Learning have the highest percentages for PhD institutions. Positions related to Security have the highest percentages for MS and BS/BA institutions. These two clustered areas are the two most sought areas for all types of institutions except for top-100 PhD institutions in which Theory/Algorithms is the second-most sought area.

A continued direction for future work is to examine how these searches translate into actual hires. Such follow-up was done in previous years [5, 6, 8], but not done this past year as the CRA COVID-19 impact survey [1] included faculty hiring outcomes.

More immediately, we do plan to follow-up these results to understand if the severe reductions in searches observed up until mid-November is the result of non-existent or simply delayed searches for 2021. In recent years we have continued to track ads from mid-November to the end of the calendar year, which has typically resulted in another 10-15% of institutions searching and
positions being sought. We plan to continue to collect data until the end of this calendar year to see if more searches were delayed due to the impact of the pandemic with plans to augment this report with those results in January 2021.

References


