Analysis of Current and Future Computer Science Needs via Advertised Faculty Searches for 2020

Craig E. Wills

Computer Science Department
Worcester Polytechnic Institute
Worcester, MA 01609

WPI-CS-TR-19-06

December 2019

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 faculty 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 394 institutions seeking to fill hundreds of tenure-track faculty positions in Computer Science. This number is a slight drop from the past two years, but still a 48% increase over the past five years of our studies. The number of tenure-track positions sought shows a decrease of 6% and 11% from the past two years, but still a 54% increase over the past five years. The number of positions being sought decreased for all types of institutions except for those offering only a BS/BA.

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 accounts for 20% of all sought positions with Security dropping to second from last year at 16%. The area of Data Science dropped to 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 had a big increase with 8% of all positions sought due to increases for the Theory and Quantum Computing topics. We found that 22-58% of all hires are for areas that are, or may be, interdisciplinary in nature.

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. Over 35% of positions for PhD institutions are in data-oriented areas.
1 Introduction

The wealth of faculty searches in Computer Science during this hiring season for tenured and tenure-track positions starting in 2020 again affords the opportunity to study areas of Computer Science where departments are choosing to invest in new faculty hires. This is the sixth such report detailing results from a study of faculty hiring ads in Computer Science. It uses a similar methodology as done in previous years [1, 2, 3, 6, 8]. The longitudinal aspect also allows insight into the number of positions and the areas being sought over time.

The primary focus of this work is 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.

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

As in past years, we used three 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\(^2\), and the Chronicle of Higher Education Vitae site\(^3\). 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 venues between August 2019 and mid-November 2019, which is the same timeframe used in our previous studies.

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

\(^1\)https://cra.org/ads/
\(^2\)http://jobs.acm.org/jobs/search
\(^3\)https://chroniclevitae.com/job_search/new
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 394 institutions (356 are in the U.S.). 305 (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 [8] found searches for 409 institutions (364 in the U.S.) with 75% of these institutions indicating a specific number of positions being searched for.

The left-side of Figure 1 shows five years of results for the number of institutions searching for tenure-track faculty. It shows a 4% decrease over the previous year’s five-year high, but a 48% increase over the five years shown.

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 [4, 5, 7] 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 780 positions. This value represents an 6% two-year and 11% one-year decrease, but still a 54% increase over the past five years. This is the first year of our study in which the number of institutions searching and the number of positions being searched for has decreased from the previous year. We analyze these decreases 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 18 Dean and Chair searches (down from 19 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. 287 (73%) of the 394 institutions listed specific topics, similar to the 75% 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 1357 specific topics are mentioned in ads (versus 1512 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 780 “positions” being searched for with 587 (75%) 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.

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 three years, although the 14% of positions this year compares to 18% in last year’s results. Machine Learning is the topic with the second highest percentage of mentions and positions. This topic jumped from last year when it was ranked fourth in percentage of positions. Data Science is third in percentage of positions (dropping from second last year) and Artificial Intelligence is third in percentage of mentions (up from fourth last year). Other topics with smaller percentages, but significant increases relative to last year in the percentage of positions include Theory and Quantum Computing.

---

4We 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.
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 [8] with that study describing small changes made to the classification based on comparisons with other topic classifications.

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 first time in our studies, it shows that the AI/DM/ML clustered area has both the highest percentage of mentions and positions (20% each). The Security area now ranks second in percentage of mentions (13%) and positions (16%) with DataSci having the third highest percentage of positions (11%) and Sys/Net having the third highest percentage of mentions (13%).

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 33% of both
<table>
<thead>
<tr>
<th>Area</th>
<th>Constituent Topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI/DM/ML</td>
<td>Artificial Intelligence, Algorithm Fairness, Computational Linguistics, Data Mining, Deep Learning, Knowledge Representation, Machine Learning/Perception, Multi-Agent Systems, Natural Language Processing, Optimization, Reinforcement Learning, Text Mining</td>
</tr>
<tr>
<td>Arch</td>
<td>Architecture, Computer Organization, Hardware</td>
</tr>
<tr>
<td>Compiler/PL</td>
<td>Compilers, Programming Languages</td>
</tr>
<tr>
<td>CompSci</td>
<td>Biological Computing, Bioinformatics, Biometrics, Biosensors, Computational Biology, Computational Science, Neuroscience, Scientific Computation</td>
</tr>
<tr>
<td>DataSci</td>
<td>Big Data, Data Analytics, Data Engineering, Data Science, Text Analytics, Visualization</td>
</tr>
<tr>
<td>DB</td>
<td>Databases, Data Intensive Systems, Data Management, Information Retrieval</td>
</tr>
<tr>
<td>HCI/IntMedia</td>
<td>Augmented Reality, Animation, Assistive Technologies, Computer-Supported Cooperative Work, Games, Human-Computer Interaction, User Interfaces, 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, Cloud Security, Cryptography, Forensics, Information Assurance, Privacy, Security</td>
</tr>
<tr>
<td>Theory/Alg</td>
<td>Algorithms, Formal Methods, Graph Analysis, Logic, Quantum Computing, Theory, Verification</td>
</tr>
</tbody>
</table>
Figure 3: Clustered Area Percentage by Mentions and 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 2020 clustering of topics shown in Table 1. Clustered areas are ordered based on 2020 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 now has the highest percentage of positions for the 2020 hiring season. The Security area, which has been consistently at the top over the five years is now ranked second, doubled during this time. DataSci is still ranked third with many of these positions going to AI/DM/ML The Sys/Net clustered area has been relatively steady around 10% over the five years. Theory/Alg, which now includes the topic of Quantum Computing, took a big jump this year and is the ranked fifth of the percentage of positions. The areas of SoftEngr, Robotics/CPS and HCI/IntMedia rank next at percentages closer to their five-year low than high. Mobile is again at its lowest percentage (1%) over 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.

3.6 Results for Interdisciplinary Clustered Areas

Another question we again examined is how the interdisciplinary nature of Computer Science is affecting hiring. Specific clusters in Table 1 that are more interdisciplinary include the Data
Science, Robotics/CPS, Computational Science, and Other Interdisciplinary clusters. Combining the results for these clusters from Figure 3 shows that 22% of the positions are for these more interdisciplinary areas. Moreover, other clusters such as AI/DM/ML and Security either support interdisciplinary work or may include work with other disciplines. Including these two clusters, which have some amount of interdisciplinary nature, results in up to 58% of the positions being interdisciplinary in nature. The total percentages are a bit lower than previous year results.

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 192 (49%) of the institutions are public and account for 397 (51%) of the total positions. The included non-U.S.-based institutions tend towards seeking higher numbers of positions.

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 (simply signified as BS in previous years), MS or PhD. Our dataset includes 183 PhD institutions—up from 178 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\(^5\), which were updated in 2019, for the top-100 U.S. and then more PhD institutions including those not in the U.S. The “PhD100” list accounts for 82 (vs. 87 last year) institutions in our dataset. The remaining PhD programs, including 38 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 five-year period. The number of institutions for all types except BS/BA dropped a bit in this most recent year.

Table 3 reveals differences between the different types of institutions. Ads for 81% of the BS/BA institutions are for a single position while 65% of the ads for PhD100 institutions are for multiple positions. The overall percentages are generally comparable as last year.

As shown, the distributions translate into a total number of 267 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 five-year results for the number of positions

---

\(^5\)http://grad-schools.usnews.rankingsandreviews.com/best-graduate-schools/top-science-schools/computer-science-rankings

### 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>% Positions w/ Specific Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Public</td>
<td>192</td>
<td>88 (46%)</td>
<td>42 (22%)</td>
</tr>
<tr>
<td>Private</td>
<td>164</td>
<td>107 (65%)</td>
<td>24 (15%)</td>
</tr>
<tr>
<td>Non-U.S.</td>
<td>38</td>
<td>6 (16%)</td>
<td>3 (8%)</td>
</tr>
<tr>
<td>All</td>
<td>394</td>
<td>201 (51%)</td>
<td>69 (18%)</td>
</tr>
</tbody>
</table>

### 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>% Positions w/ Specific Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>PhD100</td>
<td>82</td>
<td>10 (12%)</td>
<td>9 (11%)</td>
</tr>
<tr>
<td>PhDMore</td>
<td>93</td>
<td>23 (25%)</td>
<td>23 (25%)</td>
</tr>
<tr>
<td>MS</td>
<td>75</td>
<td>51 (68%)</td>
<td>16 (21%)</td>
</tr>
<tr>
<td>BS/BA</td>
<td>144</td>
<td>117 (81%)</td>
<td>21 (15%)</td>
</tr>
<tr>
<td>All</td>
<td>394</td>
<td>201 (51%)</td>
<td>69 (18%)</td>
</tr>
</tbody>
</table>
Figure 5: Five-Year Counts of Institutions Searching and Positions Being Sought by Highest Degree Offered

being searched for by the four types of institutions. It shows that the number of positions being sought for all types of institutions except for BS/BA decreased in 2020.

The last column of Table 3 shows that only 60% of positions from BS/BA institutions identify specific areas of interest while 89% of PhDMore institutions do so with the percentages for PhD100 and MS institutions in between. In order to understand differences on areas of interest between different types of institutions for 2020 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.

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 PhDMore and MS institutions. The Sys/Net area has the highest percentage for MS and BS/BA institutions. 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 roughly 36% of positions being sought by PhD-granting institutions (a bit higher for PhD100) are for the Data Oriented aggregated cluster, while roughly 29% of positions for MS and 27% for BS/BA institutions are in the Data Oriented cluster.
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</th>
<th>% Positions w/ Specific Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pub/PhD</td>
<td>92</td>
<td>21 (23%) 19 (21%) 12 (13%) 40 (43%)</td>
<td>253</td>
<td>84%</td>
</tr>
<tr>
<td>Prv/PhD</td>
<td>45</td>
<td>6 (13%) 10 (22%) 4 (9%) 25 (56%)</td>
<td>129</td>
<td>74%</td>
</tr>
<tr>
<td>Pub/MB</td>
<td>100</td>
<td>67 (67%) 23 (23%) 5 (5%) 5 (5%)</td>
<td>144</td>
<td>73%</td>
</tr>
<tr>
<td>Prv/MB</td>
<td>119</td>
<td>101 (85%) 14 (12%) 2 (2%) 2 (2%)</td>
<td>141</td>
<td>57%</td>
</tr>
<tr>
<td>All U.S.</td>
<td>356</td>
<td>195 (55%) 66 (19%) 23 (6%) 72 (20%)</td>
<td>667</td>
<td>74%</td>
</tr>
</tbody>
</table>

The results in Table 4 show that public institutions account for the clear majority of PhD-producing schools while there is a higher number of private MS&BS/BA institutions. 56% of private PhD institutions are searching for multiple positions while over 80% of public PhD identify specific topics on interest in their ads. Figure 7 shows five-year results for the number of
institutions searching and the number of positions sought based on this institution classification. In particular, the number of positions sought (right side of figure) decreased in 2020 for each grouping of institutions. Again, a caveat is that the high percentage of “multiple position” ads for private PhD institutions makes the count highly dependent on the number of positions assumed for such ads.

Figure 7: Five-Year Counts of Institutions Searching and Positions Being Sought by Institution Type and Highest Degree Offered

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 28% of the positions for private PhD institutions, 23% for public PhD institutions and a much lower 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 (12%) for private PhD institutions. There are not strong differences between the four type combinations for DataSci. The end result is that 41% of positions for private PhD institutions are in the Data Oriented cluster with 35% for public PhD, 30% for private MS&BS/BA, and 26% for public MS&BS/BA institutions. Other notable results include Systems/Networking as a relatively low percentage area for PhD institutions and Theory/Alg with a relative high percentage area for PhD institutions.
4.4 Interdisciplinary Results by Institution Type and Highest Degree Offered

Figure 9 shows the five-year percentage of positions devoted to areas that are more and some amount of interdisciplinary in nature based upon the highest degree offered of each institution. The results show that the impact of interdisciplinary areas is even more pronounced for PhD institutions with roughly 25-60% of all positions devoted to these areas by public and private PhD institutions in 2020. In contrast, public MS&BS/BA institutions have a range of 15-48% and private MS&BS/BA institutions have a range of 18-54% of positions devoted to interdisciplinary areas. These percentages tend to be lower than previous years.

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. 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 394 institutions seeking to fill hundreds of tenure-track faculty positions in Computer Science. This number is a slight drop from the past two years, but still a 48% increase over the past five years of our studies. The number of tenure-track positions sought shows a decrease of 6% and 11% from the past two years, but still a 54% increase over the past five years. The number of positions being sought decreased for all types of institutions except for those offering
Figure 9: Investment in Interdisciplinary Clustered Areas by Different Sets of Institutions
only a BS/BA.

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 accounts for 20% of all sought positions with Security dropping to second from last year at 16%. The area of Data Science dropped to 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 had a big increase with 8% of all positions sought due to increases for the Theory and Quantum Computing topics. We found that 22-58% of all hires are for areas that are, or may be, interdisciplinary in nature.

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. Over 35% of positions for PhD institutions are in data-oriented areas.

A continued direction for future work is to examine how these searches translate into actual hires. Such follow-up was done for the past three years [4, 5, 7] and should continue to look both at numbers of actual hires as well as the areas in which these hires occur.
References


