Dissecting Scholarly Patterns in Biology and Computer Science

Dinesh Majeti, M.S., Computational Physiology Lab, University of Houston, Houston, Texas, USA
Kyeongan Kwon, M.S., Computational Physiology Lab, University of Houston, Houston, Texas, USA
Panagiotis Tsiamyrtzis, Ph.D., Dept. of Statistics, Athens Univ. of Economics and Business, Athens, Greece
Ioannis Pavlidis, Ph.D., Computational Physiology Lab, University of Houston, Houston, Texas, USA

Corresponding Author: ipavlidis@uh.edu, (713) 743-0101

Keywords: Team science; Scholarly patterns; Faculty evaluation; Biology; Computer Science

Abstract: Biology and Computer Science are at the forefront of scientific and technological investigation, respectively, defining human civilization in the 21st century and absorbing large sums of public funds for research. Therefore, analyzing scholarly patterns in these two disciplines is important. At the same time such a comparative analysis is challenging due to disciplinary heterogeneity and the contextual dependence of available metrics. Biological research requires extensive experimentation that takes considerable amount of time and effort. Biology also follows a team-science model with impetus from the Human Genome Project. In contrast, Computer Science research is based on simulations and is characterized by short cycles. In this investigation we use standard academic measures of performance taking into account this differing context, in order to explore the scholarly cultures of Biology and Computer Science.

Method: We targeted the top 30 Biology and Computer Science departments according to the U.S. News Rankings 2014. For each department we collected publication data for the most cited faculty members (top ~37%) who have a Google Scholar page. The collection took place in September 2014. The data set includes $n = 569$ professors in Computer Science and $n = 500$ professors in Biology, representing an elite sample.

Results: We ran tests on distributions formed out of the departmental means of the corresponding variables.

1. Biologists publish less than computer scientists and have more authors per paper. The mean number of publications per year in Biology (5.21) is significantly lower (t-test, $p < 0.01$) than the respective mean in Computer Science (8.58). In contrast, the mean number of coauthors in Biology publications (5.57) is significantly higher (t-test, $p < 0.01$) than the respective mean in Computer Science (4.71).
2. Biologists publish mainly in journals, while Computer Scientists less so. The percentage of journal publications is significantly higher (t-test, $p < 0.01$) in Biology (73.36%) with respect to Computer Science (19.83%).
3. Biology's elite dominates publications in their top journals a lot more than the Computer Science elite does in theirs. Within the ranked set of all Biology journals, our Biology faculty sample publishes in the top 11%. In contrast, within the ranked set of all Computer Science journals, our Computer Science faculty sample publishes in the top 24%.
4. There is significant correlation between the citations obtained per year versus the number of publications produced per year in both disciplines, with Biology ($p < 0.01$, $\hat{\beta} = 81.15$, $R^2 = 0.54$) having a steeper slope than Computer Science ($p = 0.01$, $\hat{\beta} = 40.81$, $R^2 = 0.19$). The latter suggests that Biology has a tendency for higher mean citation-impact per article than Computer Science.

Conclusion: Differences in scholarship profiles between these two leading disciplines reflect to a significant degree the different nature of the respective enterprises (longer science cycles vs. shorter technology cycles). Monopolization of top disciplinary journals by Biology's elite, however, may be an issue with long term moral and performance implications that deserves further scrutiny.