THE STORIED LIFE OF A PROFESSOR OR ACADEMIC IS OFTEN PORTRAYED AS ONE WHERE THIS INDIVIDUAL HAS COPIOUS FREE TIME ON HIS OR HER HANDS, NOT ALTOGETHER REMOVED FROM THE VISION OF A CAREFREE AND GLAMOROUS HOLLYWOOD LIFESTYLE. IT MIGHT JUST BE ME, BUT TODAY’S academics seem to be working more than ever to maintain what’s still a great career and work–life balance. It’s not a question of whether we work the hours, because we do; it’s just a matter of when.

The past several months have been a bit of a whirlwind for me. For the first time ever, I organized a conference from conception to realization—the Chicago Colloquium on Digital Humanities and Computer Science. Suffice it to say, I have a new appreciation for the incredible work others put into conferences, and I’m still recovering. Being the organizer in chief, I truly can’t do justice to describe in detail what actually happened at this conference, given the array of topics that were covered. The technical presentations reminded me a great deal of some of my earlier work in high-performance computing and supercomputing, where seemingly every person present was thinking about how to apply computer science and computational methods to understand just about every problem in the humanities. I know that the digital humanities field might not be familiar to many readers, but this area has really taken off in the past few years and seems to be experiencing something similar to what happened in the 1990s with computational science, where just about every corpus (body of work) is being analyzed in one way or another using algorithmic and data-driven methods—the same methods we’re applying to computational science and engineering.

Many people might wonder what the field of digital humanities actually is. In a nutshell, it’s the application of computational methods to the humanities. To understand why anyone would want to do this, consider the following question: What do you do with a million books? This, of course, refers to the major human undertaking by Google Books to digitize seemingly all of the books on the planet. Your first reaction, if you enjoy reading—as I once did when I had free time—might be to say, “Read them.” If only life were so simple. (And even if you read 500 or more words per minute, you couldn’t read all of them anyway.) We no longer “just read” things; in our technology-driven world there are now so many ways to present, read, perceive, and analyze text. In particular, the use of text analytics and visualization can greatly guide how a person reads a text, especially if the text isn’t well understood or actively studied.

To the end of using emerging methods from digital humanities, I decided to do a wacky experiment. This experiment, in the end, had nothing to do with digital humanities per se, yet would make use of one of the tools from this community, so to speak. To describe my experiment concisely, I wanted to determine whether a given funding opportunity was relevant to my research by examining the text of various US National Science Foundation solicitations. So I paid a visit to Wordle.net, which is a toy for generating word clouds from text that you provide. The site is extremely easy to use. You simply click on the create button and the site gives you a form to enter the text that you’d like to visualize, which you can then copy and paste or use the URL for your records. Because I only wanted to analyze the text of the solicitation related to the actual research being targeted, I opted to copy and paste. I tried the text for a number of research solicitations that were currently “open,” and ultimately found one that generated the word cloud shown in Figure 1.

Suffice it to say, we had a match for my research interests, which will be evident to readers who have read CiSE’s Scientific Programming department.

Wordle is a rather neat tool that basically uses the results of one of the first computer science programs—word count—in a rather novel way. It’s a trivial algorithm per se, but the visualization is less trivial. Careful thought has been given to the layout and presentation, so that your eyes see (at a glance) which words are truly emphasized in the solicitation. Given that my work tends to be focused on the
systems area with a great emphasis on software architecture and design, it became clear to me that this solicitation is one to which I can be responsive, simply by looking at the prominence of certain words in the word cloud. More importantly, it also shed some insight on additional words that I might want to include in my proposal to illustrate my responsiveness to the solicitation. I probably could have figured this out by reading the solicitation word for word, but the word cloud’s analysis tells me something that I can’t get simply from reading.

In the end, I can’t promise you that your proposal will be accepted, but I do think the increased usage of computational and data-driven methods in the humanities is something that should be of interest to all of us—and it should inform our work and methods. More importantly, such methods might actually be useful for understanding the many texts we need to read and analyze, especially when we have such little free time on our hands.

Sometimes when people ask me why I’m involved with CiSE as a computer scientist (as opposed to a true computational scientist), I tell them, “Computing is everywhere. And computer science needs to be a part of what other disciplines do—and vice versa.” When I see projects like Wordle, this serves as yet another reminder of the growing importance of computer science to problem solving in all disciplines and providing greater understanding.

Selected articles and columns from IEEE Computer Society publications are also available for free at http://ComputingNow.computer.org.

Figure 1. Word cloud created in Wordle.net that I used to visually determine whether a given funding opportunity was relevant to my research by examining the text of various US National Science Foundation solicitations.

The American Institute of Physics (AIP) is a not-for-profit membership corporation chartered in New York State in 1931 for the purpose of promoting the advancement and diffusion of the knowledge of physics and its application to human welfare. Leading societies in the fields of physics, astronomy, and related sciences are its members.

In order to achieve its purpose, AIP serves physics and related fields of science and technology by serving its member societies, individual scientists, educators, students, R&D leaders, and the general public with programs, services, and publications—information that matters.

The Institute publishes its own scientific journals as well as those of its member societies; provides abstracting and indexing services; provides online database services; disseminates reliable information on physics to the public; collects and analyzes statistics on the profession and on physics education; encourages and assists in the documentation and study of the history and philosophy of physics; cooperates with other organizations on educational projects at all levels; and collects and analyzes information on federal programs and budgets.

The scientists represented by the Institute through its member societies number more than 134,000. In addition, approximately 8000 students in more than 700 colleges and universities are members of the Institute’s Society of Physics Students, which includes the honor society Sigma Pi Sigma. Industry is represented through the membership of 37 Corporate Associates.


Management Committee: H. Frederick Dylla, Executive Director and CEO; Richard Baccante, Treasurer and CFO; Theresa C. Braun, Vice President, Human Resources; John S. Haynes, Vice President, Publishing; Catherine D’Riordan, Vice President, Physics Resources; Benjamin B. Snavely, Secretary.