

Dear nominating committee,

I am registering my intent to run for a position on the Executive Committee.

My motivation to join the EC is driven by my belief that biocurators are the primary guardians of our research heritage. Every bit of knowledge gained is predicated on hypotheses drawn from existing knowledge. And when knowledge is lost, the rate at which we can gain a better understanding of the world we are a part of is slowed.

I have had the pleasure of facilitating biocurators since the profession was born in early nineties (coincident with the early Internet). I would like to contribute to the EC a computational perspective, how to best enable this crucial work through improvement in their work environment. The Web is a wonderful soapbox for publishing data of all kinds. Labs of all sizes are distributing their results and conclusions to the entire world. As laboratory instruments continue to improve, the sheer volume of this information is going to continue to increase super-linearly. Providing access to it is straightforward from a technical standpoint, but if you are a consumer, trying to locate particular data, then trolling the Web is not much better than rolling dice. Zipf's law applies, and the things you are most interested in can only be found in a tiny minority of the items available on the Web. Biocuration vastly improves the odds to put it mildly.

As far as my background, my research focuses on developing methods for improved integration and analysis of biological data assets. I began working in this area when I directed the Berkeley *Drosophila* Genome Project (BDGP) Informatics group, ultimately culminating in the successful publication of the annotated *Drosophila* genome in March 2000. More recently I collaborated with Drs. Lincoln Stein and Gos Micklem on the Data Coordinating Center for modENCODE, collecting both raw data and meta-data (ontologically described) from almost 40 different international laboratories. Over these years, I have brought people from a wide range of different backgrounds together: software developers, biological curators, students (undergraduate and graduate), and post-doctoral researchers to work efficiently and productively. My work has centered upon developing curational tools, covering a wide spectrum of science: from genomic feature annotation using tools such as Apollo and JBrowse (a genomics feature editor now used by numerous small genome projects and educators), to data management (e.g. the Chado database used by hundreds of small genome projects), to the most advanced use of ontologies as a formal semantic framework for organizing data. Most relevant here is the development and use of ontologies as semantic frameworks for data integration. I am one of the original founders of the Gene Ontology Consortium on which I am currently a PI. In addition my group continues to oversee the logical design and development of the other elemental key ontologies such as: Cell Types, Qualities, Sequence, and UBERON (a cross-species anatomical ontology that enables cross-species comparisons in a systematic manner. I am keenly interested in practical, real-world applications to solve biology's data management challenges. As part of the MONARCH initiative to link diseases to animal model organism resources (funded by the NIH office of the director) we are exploring the practicalities of reasoning over representations of phenotypes. In addition one of my roles in the Gene Ontology Consortium is the development of the Phylogenetic Annotation Inferencing Tool (PAINT) tool, which provides us with the ability to gain additional insights based on evolutionary relationships.

Best regards, Suzi

p.s. Matriculated from the University of Michigan, both undergraduate (Microbiology) and graduate (Systems Engineering)