

‘Big data’ revolutionizes research on evolution

Software program mines, analyzes digitized documents

BY GENE STOWE, FOR NDWORKS

Grant Ramsey’s laboratory in the Department of Philosophy has produced a Big Data search-and-analysis tool to explore questions of definition, revolution and trending in science. The application, *evoText*, already has access to a half-million articles, and pending agreements will add millions more.

Ramsey, an assistant professor in the Department of Philosophy and the Program in History and Philosophy of Science and a fellow of the Reilly Center for Science, Technology, and Values, came to Notre Dame in 2007 after he earned a Ph.D. at Duke University in the philosophy of biology.

“In graduate school, I went into philosophy because I was interested in the high-level questions of biological science and particularly evolutionary biology,” says Ramsey, whose earlier work was in science and

whose first publication was in the “Canadian Journal of Botany.” He continues to work at the intersection of biology and philosophy, but now chiefly publishes in philosophy of science journals.

One focus of his research is the foundational concepts in evolutionary theory, such as fitness, selection and drift. He was struck by the differences in how different scientists understand these concepts, and wondered how a general theory of evolution was possible without consensus about their definitions. Ramsey has published a series of papers on these concepts, and last year was awarded the prestigious Popper Prize for one of these articles. Another interest of his is the application of concepts from human studies, such as “culture” and “innovation,” to the study of animals. (See theramseylab.org for more of his research projects.)

A few years ago, Ramsey and graduate student Charles Pence, now an assistant professor at Louisiana State University, decided to investigate the use of such concepts

in scientific journals, aiming to identify different definitions and gauge their importance.

“We were interested in the academic journal literature,” Ramsey says. “There really was no tool for doing the kind of algorithmic analysis of the literature that we hoped to do. We decided to try to create a tool for doing text analysis of the journal literature, in particular biology.”

To make *evoText* possible, Pence created *RLetters* (rletters.net) a software program to mine and analyze large numbers of academic journals. This open source software can be used by anyone to mine journal articles of their choosing. But to fulfill the goal of mining the evolutionary biology journal literature, the software was implemented in *evoText* (evotext.org) a website where visitors can perform text analyses on the biology literature. The National Evolutionary Synthesis Center helped support the project.

“We have lots of different questions we’re interested in using it for,” Ramsey says, adding that

the database for scientific journals reaches to the mid-1800s. “One question concerns the origins of novel ideas in science. If we can associate a term or set of terms or phrases with particular ideas, then we can ask questions like, ‘When did this idea arrive? In what kinds of journals, specialized or generalized?’”

Other questions include: To what extent is scientific knowledge revolutionary? Do older scientists change their minds, or do new ideas arise when younger scientists replace them? Do funding agencies like the National Science Foundation tend to fund projects that are risky and cutting-edge, or do they usually

MATT CASHORE



Ramsey

fund more established research programs? And have these foundations become more or less risk averse over time? How much do trendiness and public “splashiness” drive research? “Prior to the digital revolution, one could only speculate about answers to such

questions,” Ramsey says. “But now that we have the digitized journals and text analysis tools, a new horizon of research has opened up. We are excited to see what new avenues of enquiry provided by *evoText* will be traveled by historians and philosophers of science.”