
ACM and Infosys Foundation Honor Leader in Machine Learning

David Blei Developed Powerful Modeling Techniques to Speed Information Retrieval from Large Data Collections

New York and Bangalore – April 1, 2014: ACM (the Association for Computing Machinery) and the [Infosys Foundation](#) announced today that David Blei is the recipient of the 2013 ACM-Infosys Foundation Award in the Computing Sciences. He initiated an approach to analyzing large collections of data using innovative statistical methods, known as ‘topic modeling’, that make it possible to organize and summarize digital archives at a scale that would be impossible by human annotation. His work is scalable to collections of billions of documents. It has inspired new research programs across multiple disciplines, with applications for email archives, natural language processing, information retrieval, computational biology, social networks, and robotics as well as computational social sciences and digital humanities.

The [ACM-Infosys Foundation Award](#) recognizes the finest recent innovations by young scientists and system developers in the computing field. An endowment from the Infosys Foundation provides financial support for the \$175,000 annual award. ACM will present the ACM-Infosys Foundation Award at its annual awards banquet on June 21 in San Francisco.

ACM President Vint Cerf said that Blei’s contributions provided a basic framework for an entire generation of researchers to develop statistical modeling approaches. “His topic modeling algorithms go beyond the search and links approach to information retrieval. In an era of explosive data on the Internet, he saw the advantage of discovering the latent themes that underlie documents, and identifying how each document exhibits these themes. In fact, he changed the way machine learning researchers think about modeling text and other objects in the digital realm.”

S.D. Shibulal – CEO and Managing Director, Infosys, said, “The innovative topic modeling method that David Blei has used to analyze data goes to show the capability in executing what was an unthinkable and mammoth task till a few years back. With ever-growing data generation, there is a simultaneous need to archive and interpret this data. Blei’s groundbreaking method has not only made this a simple task but will also help increase productivity significantly.”

Blei led the research that resulted in the simplest topic model, known as LDA (Latent Dirichlet Allocation). This statistical model provides a powerful tool for discovering and exploiting the hidden ‘topics’ or semantic themes in the data. It is scalable to collections of billions of documents with thousands of themes, and applies equally well to images and biological sequences. Blei’s approach is based on a Bayesian framework (a mathematical method based on probability) that exploits hidden variables to draw out the latent thematic structure in the data. LDA has been characterized as the single most important method for analyzing large collections of data. He continues to expand the scope of topic modeling with powerful methods for simultaneously analyzing documents and user behavior.

In a landmark 2003 paper detailing the development of LDA, Blei and his co-authors Michael Jordan and Andrew Ng laid out their method for discovering patterns of word use and connection documents that exhibit similar patterns. The paper, [Latent Dirichlet Allocation](#), has been frequently cited by the growing population of researchers in the topic modeling area.

Background

An associate professor at Princeton University, David Blei has written extensively on topic modeling and his pursuit of new statistical tools for discovering and exploiting the hidden

patterns that pervade modern, real-world data sets. He will join Columbia University in the fall of 2014 as a Professor of Statistics and Computer Science. He will also be a member of Columbia's Institute for Data Sciences and Engineering.

Blei is a recipient of an NSF CAREER Award, an Alfred P. Sloan Fellowship, and the NSF Presidential Early Career Award for Scientists and Engineers. His recognitions also include the Office of Naval Research Young Investigator Award and the New York Academy of Sciences Blavatnik Award for Young Scientists. Blei earned a B.Sc. degree in Computer Science and Mathematics from Brown University, and a Ph.D. degree in Computer Science from the University of California, Berkeley.

About ACM

ACM, the Association for Computing Machinery www.acm.org, is the world's largest educational and scientific computing society, uniting computing educators, researchers and professionals to inspire dialogue, share resources and address the field's challenges. ACM strengthens the computing profession's collective voice through strong leadership, promotion of the highest standards, and recognition of technical excellence. ACM supports the professional growth of its members by providing opportunities for life-long learning, career development, and professional networking.

About The Infosys Foundation

Established in 1996, the Infosys Foundation is the philanthropic arm of Infosys and has the sole objective of fulfilling the social responsibility of the company by creating opportunities and working toward a more equitable society. The Infosys Foundation has made effective strides in the areas of healthcare, education, social rehabilitation, and the arts. The company contributes up to one percent of its profit to the foundation each year.

About Infosys

Infosys is a global leader in consulting, technology and outsourcing solutions. We enable clients, in more than 30 countries, to stay a step ahead of emerging business trends and outperform the competition. We help them transform and thrive in a changing world by co-creating breakthrough solutions that combine strategic insights and execution excellence. Visit www.infosys.com to see how Infosys (NYSE: INFY), with \$7.4B in annual revenues and 158,000+ employees, is Building Tomorrow's Enterprise® today.

For further information, please contact:

<p>Virginia Gold ACM +1 212-626-0505 vgold@acm.org</p>	<p>Sukanya Ghosh Infosys, Ltd. +91 7829915577 Sukanya_Ghosh02@infosys.com</p>	<p>Tara Kozak-Lindsay GolinHarris – United States +1 212-373-6020 TKozakLindsay@golinharris.com</p>
--	---	---