

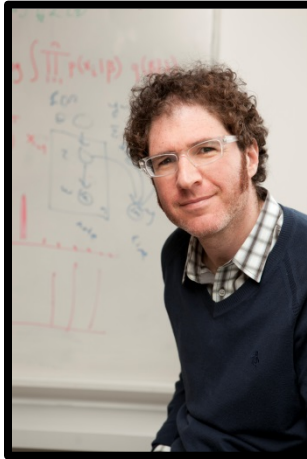


Institute for Data Sciences and Engineering
COLUMBIA UNIVERSITY

SEMINAR EVENT

DATE: THURSDAY,
SEPTEMBER 12
TIME: 6:00 P.M.
LOCATION:
DAVIS AUDITORIUM

*A light reception will take
place after the event in the
lobby*



SPEAKER: David Blei

Associate Professor of Computer Science at Princeton University

DATE: THURSDAY,
SEPTEMBER 12
TIME: 6:00 P.M.
LOCATION:
DAVIS AUDITORIUM

*A light reception will take
place after the event in the
lobby*

Abstract:

Probabilistic topic models provide a suite of tools for analyzing large document collections. Topic modeling algorithms discover the latent themes that underlie the documents and identify how each document exhibits those themes. Topic modeling can be used to help explore, summarize, and form predictions about documents.

In this talk, I will review the basics of topic modeling and describe our recent research on collaborative topic models, which simultaneously analyze texts and corresponding usage data. We studied collaborative topic models with a large collection of 80,000 scientists' libraries and the abstracts of the corresponding articles.

Traditional topic modeling algorithms take a document collection as input and analyze the texts to estimate its latent thematic structure. However, for many collections, there is an additional type of data: how people use the documents.

Bio:

David Blei is an associate professor of Computer Science at Princeton University. He received his PhD in 2004 at U.C. Berkeley and was a postdoctoral fellow at Carnegie Mellon University.

His research focuses on probabilistic topic models, Bayesian nonparametric methods, and approximate posterior inference. He works on a variety of applications, including text, images, music, social networks, and scientific data.

500 WEST 120TH STREET
NEW YORK, NY 10027

PHONE: 212-854-5660
IDSE.COLUMBIA.EDU