

PIMS-UMANITOBA DISTINGUISHED COLLOQUIUM

April 16, 2026 | 10:30 AM CT | 225 St. Paul's College

Recent Advances in Modelling Text as a Dependent Variable

Latent Dirichlet Allocation and Text embeddings have revolutionized the use of text as data leading to vastly improved document clustering and retrieval and recommender systems by finding lower dimensional structure in the text. Furthermore, these computational advances have seen substantial success when modelling text as an independent variable to be associated with numerical or categorical dependent variables. However these tools bring forward substantial challenges when considering inferential questions related to text in that these tools are highly dependent on random seed initialization. This talk focuses on the scenario where lower dimensional structure of the text is of inferential interest, with particular consideration to how to model the evolution of text structure in association with changes in covariates. We consider deterministic transformations of the text into data and model sampling variability directly as part of the regression models. We consider applications to product reviews and economic speeches from Central Banks.



DAVE CAMPBELL
CARLETON UNIVERSITY

Dr. Dave Campbell is a Professor in the School of Mathematics and Statistics and the School of Computer Science at Carleton University, where he leads a collaborative research team focused on inferential algorithms at the intersection of statistics, machine learning, computing, and applied mathematics, often in partnership with industry and government. From 2021 to 2023, he was on leave to lead the inferential Data Science team at the Bank of Canada, working on projects including cybersecurity, inflation analysis, banknote demand forecasting, and data privacy. Prior to joining Carleton in 2019, he was a faculty member at Simon Fraser University, where he helped establish one of Canada's first Bachelor's degrees in Data Science, and he has held leadership roles in the Statistical Society of Canada, including serving as inaugural President of the Data Science and Analytics Section and current President of the Business and Industrial Statistics Section.

LEARN MORE:

