

## ABOUT ME

My research broadly seeks to understand the impact of technology on our economy, culture, and society. Having training in both economics and computer science, I combine machine learning, natural language processing, and causal inference to learn about human behavior and system dynamics in the real-world. I aim to advance our understanding of these dynamic processes using large-scale empirical and data-oriented research, so that we can better leverage our knowledge to facilitate greater business and societal success.

One topic that I am particularly interested in is bias, inequality, and polarization in many digital information systems. Modern digital technology presents the opportunity for the "democratization of information" - ideas, opinions, cultural diversity, knowledge, and education will be accessible anywhere, anytime, and to everybody. As the penetration of digital technology in our every day life continues to advance, we cannot stop ourselves from asking the questions that has it hold up this grand promise? Whether there exist information skewness that we need to be cautious of? I would like to dedicate my research to answer some of these fascinating questions.

## EDUCATION

- 2015 – Ph.D. Candidate in Information Systems, Boston University
- 2015 M.A. in Economics, Indiana University (*Pass all PhD qualify exams with high honor*)
- 2013 B.A. in Economics, Peking University
- 2011 B.S. in Computer Science, Beijing Language and Culture University

## EXPERIENCE

- 2013 – 2015 Data Scientist, Center on Philanthropy at Indiana University
- 2011 – 2013 Research Associate, Artificial Intelligence Lab at Beijing Language and Culture University

## SKILLS

### Methodology:

Econometrics, Causal Inference, Machine Learning and Deep Learning, Text Mining

### Programming:

Primay: Python, R, and SQL

- Frequently used Python packages: NumPy, SciPy, Pandas, Scikit-learn, Scrapy, TensorFlow, Keras, PyTorch, StatsModels, Matplotlib, NLTK, Genism

Secondary: Stata, Java, Matlab, C/C++

## Other Tools:

Cluster Computing, Hadoop, MySQL, BigQuery

---

## RESEARCH

### Content Growth and Attention Contagion in Information Networks: A Natural Experiment on Wikipedia

Kai Zhu, Dylan Walker, Lev Muchnik

- Accepted at **Information Systems Research**
- Presented at *WWW 2017, CODE 2017, WISE 2017, SCECR 2018, HBS 2018, WEBEIS 2019*

**Abstract:** Open collaboration platforms have fundamentally changed the way knowledge is produced, disseminated and consumed. In these systems, contributions arise organically with little to no central governance. While such decentralization provides many benefits, a lack of broad oversight and coordination can leave questions of information poverty and skewness to the mercy of the system's natural dynamics. Unfortunately, we still lack a basic understanding of the dynamics at play in these systems, and specifically, how contribution and attention interact and propagate through information networks. We leverage a large-scale natural experiment to study how exogenous content contributions to Wikipedia articles affect the attention they attract and how that attention spills over to other articles in the network. Results reveal that exogenously added content leads to significant, substantial and long-term increases in both content consumption and subsequent contributions. Furthermore, we find significant attention spillover to downstream hyperlinked articles. Through both analytical estimation and empirically-informed simulation, we evaluate policies to harness this attention contagion to address the problem of information poverty and skewness. We find that harnessing attention contagion can lead to as much as a twofold increase in the total attention flow to clusters of disadvantaged articles. Our findings have important policy implications for open collaboration platforms and information networks

### How Media Ownership Impacts Information Skews: A Study of Televised News Using Massive-Scale Text Transcripts

Kai Zhu, Dylan Walker

- Presented at *PaCSS 2018, SCECR 2019*; To be presented at *WISE 2019*

**Abstract:** Broadcast TV in the United States is an information system comprised of hundreds of local television stations that both produce their own information and syndicate information from other sources (such as major networks). In this system, a few media conglomerates are behind the vast amount of information produced and disseminated. There has been an increasing trend of media consolidation in recent years and this raises the very real concern that conglomerate owners have both the means and motive to skew information. When large owners act coherently, they can skew information to emphasize views, perspectives, framing, coverage and attention. This is important because broadcast media has a dramatic impact on political and social outcomes and undeniably shapes the national dialogue surrounding important issues. Unfortunately, we know little about how ownership affects information skew and lack a systematic empirical evaluation of content and ownership at sufficient scale and detail. In this study, we quantify and investigate the consequences of information system ownership, specifically in terms of diversity of information and political polarization, in one of our most important mass information systems - broadcast televised news. We aim to understand how does media ownership impact information diversity, political polarization, and topic coverage in TV news. To examine the impact of ownership, we focus on the natural experiments of hundreds of station acquisitions by new owners over a six year time span. The change of ownership allow us to disentangle the impact of owner of a TV station from other characteristics that may also correlated with the information produced by a station.

## Media Coverage of Gun Violence in the United States

Kai Zhu, Dylan Walker

**Abstract:** Gun violence in the United States has reached epidemic proportions, where guns have been implicated as the cause of more than 30,000 fatalities and 80,000 injuries annually, according to the Centers for Disease Control and Prevention and the American Public Health Association. The dozens of Mass shooting incidents over the past several years have provoked a national conversation about gun violence, regulation, public policy, and mental health with particular attention to how the media covers and frames gun violence incidents. Despite this attention, our empirical understanding of how the media covers gun violence is remarkably limited, in part due to a lack of comprehensive data on the detailed incidence of gun violence events and of coverage of gun violence events in news media throughout the United States. Our collaboration with a private firm that tracks local and national televised news content has placed us in a position to overcome this hurdle. Using the complete transcript data of all news from local television stations in the United States over a six year timespan and comprehensive records of gun violence incidents throughout the country compiled by Gun Violence Archive, we investigate how media cover gun violence incidents.

## Work in Progress

Text-based Measures of Information Diversity: A Deep Learning Approach

Kai Zhu, Dylan Walker

Measuring the Real-World Impact of Fact Checking in Combating Online Misinformation

Kai Zhu, Dylan Walker

Combating Health-Related Misinformation on Social Media: A Randomized Experiment  
Kai Zhu, Dylan Walker

## Publications in Computer Science

Attribute Reduction Approaches for General Relation Decision Systems (2015), Guilong Liu, Ling Li, Jitao Yang, Yanbin Feng, Kai Zhu, *Pattern Recognition Letters* 65, 81-87

The Relationship among Three Types of Rough Approximation Pairs (2014), Guilong Liu, Kai Zhu, *Knowledge-Based Systems* 60, 28-34

## SELECTIVE COURSEWORK

### Economics

Microeconomics 1 and 2  
Advanced Information Economics  
Contract Theory and Theory of Incentives  
Game Theory  
Industrial Organization  
Econometrics 1 and 2  
System and Panel Econometric models  
Advanced Microeconometrics  
Time Series Econometrics  
Causal Inference in Management Research

### Computer Science

Advanced Machine Learning  
Algorithms Design and Analysis  
Convex Optimization  
Neural Networks for Natural Language Processing  
Deep Learning  
Tools, Techniques, and Theories for Data Mining  
Computational tools for Data Science  
Machine Learning Methods for Social Science Research

### Seminars and Courses around Boston Area

Economics of Information and Technology (Erik Brynjolfsson, MIT)  
Applied Network Theory and Analysis (Sinan Aral, MIT)  
Empirical Studies of Innovation and Digitization (Shane Greenstein, Harvard)  
Advanced Quantitative Research Methodology (Gary King, Harvard)  
Machine Learning in Econometrics (Sendhil Mullainathan, Harvard)  
Design and Analysis of Experiments (Dean Eckles, MIT)

Network Science (Albert-Laszlo Barabasi, Northeastern)

Statistical Methods for Evaluating Causal Effects (Donald Rubin, Harvard)

## TEACHING

**Instructor** Introduction to Information Systems

- *Instructor rating: Mean - 4.5 out of 5, Median - 5 out of 5*
- *Description:* The course is designed to provides students with an understanding of the important role that information and information technology play in supporting the effective operation and management of business. We look at how information systems solve problems, create opportunities, and can be a disruptor of traditional business models. To achieve this, the course utilizes case materials, outside readings, e-Learning tools, and hands on MS Access database exercises. The goal is to engage students in understanding how the effective management of information adds value to a business and how to apply information technologies to that process. The four main modules of this course are: Foundations of Management Information Systems, Database and Analytics Skills, and Information Technology as a Disruptor of Traditional Business Models, Business System Skills.

**Course Design** Deep Learning for Business Analytics

- *Description:* I help design and develop a new course about Python and Neural Network for business analytics as part of the recently launched Master of Science in Business Analytics program at Questrom School of Business, Boston University. The course teaches student using Python for data science and deep learning as a tool for business analytics. It covers topics such as theory and fundamentals of Neural Network, Convolutional Neural Network, Recurrent Neural Network, and a Capstone project of application of deep learning methods for business data. The course is comprised of 20 sessions and each session is a 3.5 hour long intensive class including both lecture and hands-on exercise and projects.

**TA** Human-centered Digital Design

- Year: 2017, 2018, 2019 , MBA

○ Instructor: Dylan Walker

Platform Strategy

- Year: 2019, Undergraduate and MBA

○ Instructor: Marshall Van Alostyne

Digital Project Construction

- Year: 2018, MBA

○ Instructor: Benjamin Lubin

Managing Networked Systems

- Year: 2016, MBA

○ Instructor: Benjamin Lubin

Computer Architecture and Systems Software

- Year: 2016, MBA

○ Instructor: Dylan Walker

*Description:* My responsibility as teaching assistant for the above courses including teaching lab sessions, tutoring Python programming, consulting digital projects development, hosting Q&A hours, and grading homework and final exams.

## CONFERENCE PRESENTATION

- 2019 (scheduled) "How Media Ownership Impacts Information Skews: A Study of Televised News Using Massive-Scale Text Transcripts", Workshop on Information Systems Economics (WISE 2019), Munich, Germany
- 2019 "How Media Ownership Impacts Political Bias and Information Diversity: A Large-Scale Study of Broadcast Media", Statistical Challenges in Electronic Commerce Research (SCECR 2019), Hong Kong, China
- 2019 "Content Growth and Attention Contagion in Information Networks: A Natural Experiment on Wikipedia", Workshop on Experimental and Behavioral Economics in Information Systems (WEBEIS 2019), Minneapolis, Minnesota, USA
- 2018 "Content Growth and Attention Contagion in Information Networks: A Natural Experiment on Wikipedia", Harvard Business School, Doctoral Digitization Workshop, Boston, Massachusetts, USA
- 2018 "Political Slant in Local Televised News", Politics and Computational Social Science (PaCSS 2018), Boston, Massachusetts, USA
- 2018 "Content Growth and Attention Contagion in Information Networks: A Natural Experiment on Wikipedia", Statistical Challenges in Electronic Commerce Research (SCECR 2018), Rotterdam, Netherlands
- 2017 "Content Growth and Attention Contagion in Information Networks: A Natural Experiment on Wikipedia", Workshop on Information Systems Economics (WISE 2017), Seoul, South Korea
- 2017 "Content Growth in Network: A Natural Experiment on Wikipedia", The Conference on Digital Experimentation (CODE 2017), Boston, Massachusetts, USA
- 2017 "Content Growth in Network: A Natural Experiment on Wikipedia", Wiki Workshop at World Wide Web (WWW 2017), Perth, Australia

## SERVICE

### Reviewer

Journal:

- Management Science

Conference:

- Conference of Information Systems and Technology (CIST: 2016, 2017, 2018, 2019)
- International Conference of Information System (ICIS: 2018, 2019)
- Workshop of Social Influence (SI: 2018)