1. Contact Info

Name: Kai Zhu

School: Boston University

Email: kaizhu@bu.edu

Short bio: I am a 4th year PhD student in the Information Systems Department at Boston University. I am broadly interested in computational social science with a focus on bias, inequality, and polarization in various digital information systems, including Wikipedia, broadcast TV, Twitter and other social media.

2. Coursework

I have taken course mainly in the field of Economics, Management, and Computer Science. A selected sample of my coursework in graduate school is as following:

1) Management

Economics of digitization (MIT, Erik Brynjolfsson)

Empirical Research of Innovation and Digitization (Harvard, Shane Greenstein and Karim Lakhani)

Information Economics (BU, Marshall Van Alstyne)

Causal Inference for Management Research (BU, Tim Simcoe)

Social Network: Theory and Application (MIT, Sinan Aral)

Digital Experiments: Design and Analysis (MIT, Dean Eckles)

2) Economics

Microeconomics 1, 2

Macroeconomics 1, 2

Econometrics 1, 2, 3

3) Computer Science

Machine Learning

Computational Tools for Data Mining

Deep Learning and Neural Network

3. Research

1. Content Growth and Attention Contagion in Information Networks: A Natural Experiment on Wikipedia (Under Review), with Dylan Walker and Lev Muchnik

Abstract: Online collaborative systems 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 the information network. We leverage a large-scale natural experiment to study how exogenous content contributions to a Wikipedia article affect the attention it attracts 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. Finally, 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 online collaborative systems and information networks.

2. Political Polarization and Information Diversity in Televised News (Working-in-Progress), with Dylan Walker

Abstract: The political polarization of mass media has a dramatic impact on political and societal outcomes and strongly shapes the national dialogue surrounding important issues. Polarization has recently received increased attention from academic researchers. However, we still know very little about the geographical distribution of political polarization and the factors that determine it. In this study, we investigate the underlying drivers of political polarization in televised news arising from the supply and demand sides of the market. Specifically, we consider the supply side as comprised of parent networks of local TV channels such as FOX or NBC and owners of local channels such as the Sinclair Broadcast Group. The demand side is comprised of local audience populations with differing political preferences. We theorize that together these drivers determine the (equilibria) polarization of content in local markets. To understand how this phenomenon manifests, we leverage computational techniques in natural language processing to investigate and quantitatively estimate supply- and demand-side drivers on 6 years of complete transcript data of televised local news across more than 1000 local TV channels in the US from 2012-2018.