NBER Digitization Tutorial Application

Matthew Yeaton

November 16, 2020

1 Contact Information

- Email address: mry2105@gsb.columbia.edu
- Physical address: 415 West 118th Street Apt. 31 New York, NY 10027

2 Course of Study

- PhD in Strategy at Columbia Business School (5th year)
- Advisor: Bruce Kogut
- Relevant coursework:
  - Economics: First year microeconomics sequence, Industrial Organization theory seminar, Information Economics theory seminar, Economics of Incentives seminar
  - Statistics and Applied Math: Causal Inference, Bayesian Statistics, Spatial Statistics, Dynamic Programming and Reinforcement Learning, Natural Language Processing, Machine Learning, Experimental Research
  - Strategy: Strategy seminar, Advanced Strategy seminar, Organization Theory seminar, Innovation Diffusion seminar

3 Research Interests

I am very interested in digitization and have several projects around the topic. Digitization is important for researchers both as a phenomenon in itself and as an ever-expanding source of accessible data to test theory. In particular, I am interested in social learning, language

(as the pipes of communication), and coordination. Digital communities are an ideal setting to study these topics as the proliferation of online communities and the relative permanence of the internet gives us access to large databases of repeated communication.

The first project examines the extent to which we can be successful in breaking apart echo chambers and filter bubbles of priorities and opinions. This problem is complicated because platforms and organizations are generally unable to measure and thus incentivize the desired behavior, and so the insights of agency theory are moot. However, this chapter suggests that knowledge of the dynamics of language and organizational structure enables a less invasive solution: small changes to the social network. Using a natural experiment on the website reddit.com, we can test this directly. The website banned an altright subreddit, or subforum, but did not ban any users, which resulted in a shock of the user interaction social network. Across a number of ways of measuring language use, I find that this shock is associated with changes in language away from the jargons of the altright community and towards the language of the broader community. Additionally, this approach suggests that this set of tools may be used more broadly to lend insight to an organization's culture.

The second project seeks to show how language, as a central aspect of organizational culture, enables organizational memory and creates differences between organizations in which types of knowledge can be efficiently communicated. Language in this formulation has connection to routines, in that they are held at the organizational level, and can be a means to transfer knowledge efficiently within an organization, consistent with the knowledge-based view of the firm. Language also takes on an emergent property of storage. This is because we cannot use the language to communicate without also communicating its priorities via its structure. Like an organizational Ship of Theseus, I ask how much of the language of the original members remains when they have all been replaced? I find that by iteratively replacing every member of the organization, a substantial percentage of the new language reflects the founding members' priorities. Moreover, I find that the problem is worse for larger organizations. Even after two full replacements of the organization, that is everyone has been replaced, and their replacements have been replaced, about 15% of the new language reflects the original members' priorities. Common language is a type of limited codifiability that makes certain (possibly tacit) routines and situations more efficient to describe within an organization. Thus, the upshot of common language is that it provides a pathway for organizational memory. On the other hand, this also suggests that it is not enough to change the language and culture of organization simply by changing its members, and thus may be one part of the puzzle of organizational rigidity. I plan to build on the formal model with an online experiment.

The third project develops a formal model of fake news as bias of the relative precision

of signals rather than bias of the signal means. Under this type of bias, Bayesian synthesis of signals can be biased (compared to the true model) even though none of the individual signal means is biased. This model can help explain the "white noise" Twitter strategy of Mexican presidential candidate Enrique Peña Nieto and his PRI party during the 2010 presidential election. The PRI hired a small army of Twitter trolls to manage the impression of the candidate on social media. But while they were not actively hostile to the candidacy, neither did they promote the candidacy directly. Instead, they used a "white noise" strategy that magnified the noise during times of particular bad news about Peña Nieto (e.g. claiming Justin Bieber had died). This type of fake news is realistically possible only in the wild west of digital information diffusion.