

BEHAVIORAL INSIGHTS FROM TEXT CONFERENCE

January 17, 2020

Room 240 Huntsman Hall, 3730 Walnut Street, Philadelphia, PA 19104

8:00 – 8:45 AM – Continental Breakfast

8:45 – 9:00 AM – Introductions

9:00 – 10:15 AM – Session 1

- **Lauren Cohen (Harvard Business School):**

We find that managers strategically shift targets in their communications with investors and markets. Investors should pay close attention to the metrics upon which firms choose to focus, and the subtle changes to those metrics that firms make over time, as moving targets contain important information for future firm value and realizations.

- **Khrystyna Bochkay (University of Miami):**

We develop a dictionary of linguistic extremity in earnings conference calls, a setting where managers have considerable latitude in the language they use, to study the role of extreme language in corporate reporting. Our evidence suggests that market participants are influenced not just by what managers say, but also how they say it, with extreme language playing an important role in management communications with investors.

- **Kent Harber (Rutgers University at Newark):**

The Positive Feedback Bias is the tendency among White instructors to provide selectively more praise and less criticism to students of color. However, language to a Black writer, especially among subjects who felt anxious giving feedback, employed shorter and simpler words conveyed using more rudimentary syntax. In sum, Black writers received conflicted and confusing feedback messages, which explicitly lauded their work while implicitly denigrated their potential.

- **H. Andrew Schwartz (Stony Brook University (SUNY)):**

Facebook language has predicted diagnoses of depression slightly more accurately than screeners and used to assess personality on par with self-report. Here, we will describe ongoing work of the World Trade Center Language and Health Study, applying computational linguistics techniques to oral interviews from a cohort of hundreds of first responders to the 9/11 tragedy.

10:15 – 10:35 AM – Break

10:35 AM – 12:05 PM – Session 2: Research Blitz

- **Sandeep Soni (Georgia Institute of Technology):**

Diachronic word embeddings -- vector representations of words over time -- offer remarkable insights into the evolution of language and provide a tool for quantifying socio-cultural change. Prior work has used such embeddings to identify semantic changes i.e., words whose meaning has shifted. In this paper, we link word embeddings to the documents in which they appear, by situating those documents as leaders or laggards with respect to ongoing semantic changes.

- **Joe Hoover (Northwestern University):**

Acts of hate have been used to silence, terrorize, and erase marginalized social groups throughout history. However, in recent years there has been a marked increase in their frequency, which underscores the importance of developing a better understanding of when, why, and where they occur. We find consistent evidence that acts of hate are strongly tied to moral values associated with group preservation, the so-called Binding values.

- Russell Richie (University of Pennsylvania):

Information retrieved from memory, often in the form of sentence-like thoughts, influences the formation of preferences and beliefs in most naturalistic decision tasks. We use vector representations in a novel modeling framework that implements established theories of memory search and decision making within a single integrated cognitive system. We showcase the power of our framework by applying it to study natural language thoughts in domains including risk perception, consumer behavior, financial decision making, ethical decision making, legal decision making, food choice, well-being, and about society and culture.

- Reihane Boghrati (University of Pennsylvania):

Many have argued that music is misogynistic, but is that actually true? And have any such biases changed over time? Natural language processing of a quarter of a million songs over 50 years tries to address these questions. While both genders are equally likely to be objects of aggression, subtler machine learning approaches paint a more complex picture. Ancillary analyses suggest that lyrics have become less gendered more broadly (though remain gendered) and that temporal changes may be driven by male artists (as female artists were less biased initially).

- Alain Lemaire (Columbia Business School):

In this paper, we examine how the linguistic similarity between the language used by reviewers of a product and prospective customers' own writing style can be leveraged to assess the match between customers and products. Applying tools from machine learning, Bayesian statistics, and computational linguistics to a large-scale dataset from Yelp, we find that the closer the writing style of a restaurant's past reviews are to a prospective customer's writing style, the more likely that customer is to write a review for that restaurant. Our results suggest that recommendation agents should incorporate linguistic matching as part of the recommendation algorithm.

- Alexander Ruch (Cornell University):

Integrating graph and language embedding models' complementary relational and communicative data may be especially helpful if predicting rare events or classifying members of hidden populations. For example, mental health support groups often form in amorphous online groups. Predicting suicidality among individuals in these settings using standard network analyses is prohibitive due to resource limits (e.g., memory), and adding auxiliary data like text to these models further exacerbates complexity- and sparsity-related issues. Merging graph and language embedding models (metapath2vec and doc2vec) avoids such limits.

- Emma Jesch (University of Pennsylvania):

Does media coverage of tobacco and e-cigarettes affect youth and young adult beliefs and behavior? Previously we reported about valence frequency and evidence that anti-tobacco media coverage predicted lower intentions to smoke among non-smokers and lower intentions to continue smoking among smokers (and showed that these intentions were associated with smoking behavior). In this presentation, we extend these results to focus on beliefs associated with intentions to smoke or vape.

- Ivan Gordeliy (Georgetown University):

In this work, we combine empirical and theoretical approaches to understand linguistic features that differentiate word-of-mouth based on authentic experiences versus word-of-mouth based on experiences that did not happen. We focus on the context of consumer reviews, but our theoretical foundations apply to other contexts, such as news, user-generated content, or police reports and evidence.

12:05 – 1:00 PM – Lunch Break

1:00 – 2:15 PM – Session 3

- **Andrey Siminov (Columbia Business School):**
News consumers in many authoritarian countries read government-controlled sources even when independent sources are available. What drives this demand for government-controlled news? We separate out two potential sources of such demand, preferences for pro-government coverage of sensitive events and persistent preferences for news outlets, in a case study of the online news market in Russia in 2013-2015.
- **Svitlana Volkova (Pacific Northwest National Laboratory (PNNL):**
Deceptive news — misleading, falsified and fabricated content — is routinely originated and spread on social platforms with the intent to create confusion and widen political and social divides. In this talk, I will review computational approaches to identify deception online, explain model predictions and evaluate model robustness, measure the immediate spread of deceptive content and quantify user reactions to it, and discuss user behavior while engaging with deceptive news.
- **Gerard Tellis (USC Marshall School of Business):**
Analysts exploit the sentiment and topics expressed in social conversations to ascertain what people are saying about brands and politicians. However, one key limitation of the current methods is that they fail to consider the location and the geographic influence of social media posts i.e., the how the spatial diffusion of social media posts affects brand preferences at a given location. In this paper, we show how to incorporate users' locations and geographic social influence (geo-influence) on users to monitor brand preference at the micro-geographic level.
- **Olivier Toubia (Columbia Business School):**
Everyday language often describes stories as having shapes (e.g., as having arcs, as pushing the envelope, as going in circles, etc.). In this paper, we propose, apply and test several metrics that describe the shape of stories in word embedding spaces. Our framework is general enough to apply to different types of stories, although empirically we find that different characteristics are relevant in different types of stories.

2:15 – 2:35 PM – Break

2:35 – 3:50 PM – Session 4

- **Fei Long (University of North Carolina at Chapel Hill):**
Using dynamic topic modeling, we analyze the textual content and numerical ratings for product reviews posted between the years 2000 and 2014 on Amazon.com for headphones and cellphones. We find that consumer reviews that appear early in a product's life cycle are systematically different from those that appear later in the life cycle. Furthermore, review ratings decline over time. Next, we use a regression-based approach to relate the numerical ratings of reviews to review content and uncover interesting dynamics. For instance, we find that review text content is more indicative of review ratings earlier in the product cycle (when review content is more feature oriented) than later (when review content is more experience oriented). Notably, we find that the negative effect of time on ratings stays even after controlling for review content, review order and calendar-time.
- **Ashlee Humphreys (Northwestern University):**
As consumers move through the decision journey, they adopt different goals (i.e., information vs. transactional). In this research, we propose that consumer goals can be detected through textual analysis of online search queries and that both marketers and consumers can benefit when paid search results and advertisements match consumer search-related goals. Results from our empirical studies illustrate that matching language with consumer mindsets can increase engagement with search results and search advertising.

- Dokyun Lee (Carnegie Mellon University):

We introduce the Focused Concept Miner (FCM), an interpretable deep learning text mining algorithm to (1) automatically extract interpretable high-level concepts from text data, (2) focus the mined concepts to explain user-specified business outcomes, such as conversion (linked to read-reviews) or crowdfunding success (linked to project descriptions), and (3) quantify the correlational relative importance of each concept for business outcomes against one another and to other explanatory variables. Compared to 4 interpretable and 4 prediction-focused baselines that partially achieve FCM's goals, FCM attains higher interpretability, as measured by a variety of metrics (e.g., automated, human-judged), while achieving competitive predictive performance even when compared to prediction-focused blackbox algorithms.

- Francisco Villarroel Ordenes (University of Massachusetts Amherst):

Despite universal agreement about the importance of frontline employees (FLEs) in the service process, they report higher job dissatisfaction on average compared to other less visible jobs in organizations. This is a major concern in service organizations, where management struggles to identify practices to retain talented FLEs. Deloitte's human capital trends report reveals that in practice, "nearly 80 percent of executives rated employee experience very important (42 percent) or important (38 percent), but only 22 percent reported that their companies were excellent at building a differentiated employee experience." Moreover, there is scarce longitudinal evidence regarding the implications of FLE experience for employee and customer outcomes. Bridging job-demand and resources (JD-R) framework with the broaden-and-build-theory, this research conceptualizes FLE experience and its customer impact.

3:50 – 4:10 PM – Break

4:10 – 5:00 PM – Keynote: Noah Smith (University of Washington)

5:00 – 5:15 PM – Concluding Remarks: Jonah Berger and Bob Meyer

5:15 – 7:00 PM – Cocktails and Light Food (8th floor, Huntsman Hall)