KERI LARSON
Assistant Professor of IS
Williams School of Commerce,
Economics, and Politics
Washington & Lee University

Research Areas

- Text mining/unstructured data analytics
- Neuro-IS
- Healthcare IT/IS
- Organizational responses to innovation
- Identity

Research Agenda

The areas of interest listed above encompass my current streams of research. My agenda for the next three years involves targeting a steady stream of manuscripts to major IS journals (*MIS Quarterly*, *ISR*, etc.) derived from my dissertation work in social media text analytics, my program of Neuro-IS research, and various complementary research projects that fall under the rubrics of Healthcare IS, organizational responses to innovation, and identity and IS.

Text Mining/Unstructured Data Analytics

Conducted under the mentorship of Dr. Rick Watson, my dissertation work addresses the organizational dilemma of how to derive *meaningful* knowledge from social media exchanges, concentrating on social media analytics and the business value organizations might derive from textual data generated by these technologies.

I have carved four manuscripts from this work. The first two papers are conceptual and are supported in part by a wealth of archival data originating from multiple campaigns of an Atlanta social media marketing firm. The third paper, an experiment, tests the efficacy of a theoretically crafted social media analytics system based on natural language processing (NLP) concepts. Stemming from a need for a fourth treatment for a second run of the experiment to strengthen my results, I am currently working out an additional methodology paper for training a machine learning algorithm to sort out large sets of domain-unspecific text in a valid and reliable manner. My coauthors and I are using a large number of crowdsourced contributions to attempt to eliminate idiosyncratic bias from the process of building a training set from human contributions.

As my major area of interest, unstructured text analytics represents both a serious challenge and huge potential source of value for organizations. Data analytics is a core IS concern and promises plenty of problems to be resolved, ensuring forward movement with this research focus. A further description of each paper and its targeted publication follows.

1. "Tying Social Media Strategy to Organizational Decision-Making: A Social Media Analytics Framework" (with Richard Watson) – From the first paper of my dissertation I have an *ICIS* publication ("The Value of Social Media: Toward Measuring Social Media Strategies"), which forms the basis of this journal-length paper intended for submission to *MISQ* this calendar year. This paper presents a model of the social media stakeholder ecosystem suitable for situating a wide range of social media research problems. It

unpacks the processes that occur between the firm and its customer stakeholder group across social media applications and delineates granular steps that must be measured by organizations interested in monitoring the success of their social media strategies via the analysis of unstructured textual data. Based on AE and reviewer responses to the *ICIS* submission, this journal submission is poised to fill an important gap in the area of social media analytics research. Indicative comments include:

> "This is an excellently structured, well thought through and timely paper."
> "…this paper represents one of the most thoughtful discussions of social media research I have read."
> "Thank you for submitting this timely paper on social media metrics."
> "Overall, I find this paper well-conceived and well-written. This is a timely stream of research that will be well received and create interesting conversation at ICIS."
> "It is clear to me that this paper represents the first steps in what is likely to be a major contribution to social media research."

2. "The Analysis of Unstructured Data: *Meaningful* Measurement of Social Media Interactions" (with Richard Watson) – Framing further inquiry into the analyses of voluminous and dynamic corpora of customer-generated social media data, this paper reviews the state of the art of textual analysis, a technique that can provide the deep level of qualitative analysis needed to fully ascertain important tends in firm/customer and customer/customer social media exchange. The paper advances a theoretical argument for the inadequacy of existing text mining applications and techniques for supporting organizational decision-making based on social media-derived intelligence, concluding with the articulation of a set of design principles for a social media analytics system based on natural language processing capabilities.

This paper will be submitted as a design science piece to *JAIS*. There is further opportunity to craft a second manuscript from this paper that proposes a new methodology for reviewing and synthesizing tools as a mechanism for creating a firm foundation for advancing knowledge, similar in spirit to writing a literature review (Webster and Watson, 2002). This review piece, co-authored with Rick Watson, will be targeted to *MIS Quarterly*.

3. Textual Analysis of Social Media Interactions: A Natural Language Processing Approach" (with Richard Watson and Michael Scholz) – From the third paper of my dissertation I have an *ECIS* publication ("The Impact of Natural Language Processing-Based Textual Analysis of Social Media Interactions on Decision Making"). This manuscript experimentally tests the proposed analytics approach against sentiment analysis and manual approaches to mining knowledge from social media data, and is demonstrated to provide superior support for organizational decision-making through improved problem detection. Of particular consequence is that while problems and opportunities are more accurately detected from social media data using an NLP-based approach as expected, sentiment analysis surprisingly appears *no more useful* than randomly reading segments of social media data manually. These results lead to a recommendation for a more useful system for monitoring firm-level effects of social media. I am currently working on a fourth treatment, which I will add to the experiment prior to re-running it on a larger scale. Upon conclusion of the second experimental run, this paper will be submitted to *MISQ* mid-2015.

4. "Forced Ranking of Social Media-Generated Textual Data" (with Richard Watson and Michael Scholz) – Stemming from the need for an additional experimental treatment based on output from a machine learning (ML) algorithm trained on a large set of sorted textual data, this paper sets forth a methodology for sorting large bodies of text in order of importance. Supervised ML algorithms learn from past data in order to better to distinguish, sort, or mine future points of data. However, building a robust ML model able to process information reliably depends on a high-quality set of training data. Creating a training set of text sorted in order of importance is not a trivial task, and is subject to the problem of limited cognitive processing ability in humans. Large sets of force-ranked training data are unlikely to be replicable over time by a single rater, much less across multiple individuals. While structured numerical data have an inherent order, qualitative data require the imposition of some type of judgment to give them meaning relative to the set. Ranked sets are therefore bound to be idiosyncratic to their judges. As a result, algorithms trained on the same data ordered by two different raters are likely to sort or mine differently from one another, meaning the two will produce different results. Because machine learning holds such great promise for a variety of qualitative data-dependent tasks, it is imperative that we devise a valid and reliable method for creating the ordered training sets ML algorithms require. This paper reviews the literature on tournament and pairwise ranking aggregation, and proposes a crowdsourced approach to force ranking large textual data sets based on the Crowd-BT algorithm. This paper will be submitted for review at *ECIS* this December, and based on feedback will be targeted to an appropriate journal in 2015.

<div align="center">

Neuro-IS

</div>

As a member of the Comprehensive Neuroscience Center at UAB, I have collaborated with a team of business and neuroscience scholars to develop a program of research using psychophysiological measures to assess IS phenomena. The first paper supported by our ongoing data collection will be submitted to the *MISQ* special issue on innovation in September 2014. In addition to this current manuscript, we anticipate developing at least two additional empirical papers investigating additional team-level phenomena via psychophysiological measures. Porting direct measures into IS-relevant contexts could enable us to revisit a wide range of constructs that have proven difficult to reliably assess with indirect measures. Functional neuroimaging techniques promise to shed light on decision-making and other important IS constructs as they operate in the brain's areas and structures, complementing our theoretical understanding.

5. "Leading by Dominance: A Neurological Approach to Understanding the Influence of Lead Users on Open Innovation Systems" (with Allen Johnston, Paul di Gangi, Anthony Hood, and Cali Fidopiastis) – Organizations such as Dell, IBM, Starbucks, and SAP have successfully adopted open innovation systems crowdsourcing sites to generate new ideas for their product/service portfolio. A large part of the success of open innovation systems is the degree to which organizations are able to harness the power of *lead users* (DiGangi and Wasko 2009; DiGangi, Wasko, and Hooker 2010), individuals with characteristics that allow them to forecast need, project value, and identify early market trends, and are socially centric to other innovators (von Hippel 2005). Using the phenomenon of shared cognition, a measureable psychophysiological phenomenon, we reexamine its theoretical conceptualization as both a process and an outcome (Ensley and Pierce 2001) generally concerned with group performance. It has been demonstrated that during goal-driven activities, shared cognition leads to consensus among deliberators, and that this shared cognition is the result of the interplay of cognitive dominance and cognitive cooperation among group

members. We posit that lead users impose their will over others through cognitive dominance, while non-lead users will cognitively cooperate, which we can directly measure. Neurological assessment allows us to capture the interplay of dominance and cooperation during the consensus building or decision-making process. To understand the influence (cognitive dominance) of lead users within user innovation communities (UICs), this controlled laboratory experiment manipulates the lead user of a group (high versus low on lead user characteristics) compared to his cohort (high versus low lead user status compared to the identified lead user) and mode of interaction (face-to-face versus virtual teams), assessing how lead users influence shared cognition within UICs and executive decision making teams through their cognitive dominance over other non-lead users, and whether this varies in virtual communities from face-to-face settings.

This research is the first to predict and test the influence of lead users within open innovation systems, particularly (1) their influence on the generation and processing of innovation ideas within a UIC and (2) their influence on executive decision teams seeking to development implementation strategies from the innovation ideas emerging from a UIC. This research is also the first to apply a neurophysiological assessment of the presence of shared cognition within decision-making teams. We are submitting this manuscript to the *MIS Quarterly* special issue on innovation in September 2014.

Organizational Responses to Innovation

I have a handful of active studies outside of my two major streams of research, most of which developed as interesting side projects from other studies. These include a literature review of identity work in IS research, two practitioner pieces regarding user innovation communities (UICs), and a longitudinal qualitative study of how parallel nursing units co-opt technological innovations differently depending on differences in institutional logics across units. The shared theme amongst these studies is how organizations and communities of practice respond to disruptive technological innovations.

6. "Second-order Responses to Innovation in Healthcare" (with Nicholas Berente and Marie-Claude Boudreau) – The theoretical front-end of this work was published at *ECIS* 2011. We are currently culminating a five-year longitudinal qualitative study of nurses in units that piloted a new EHR system. Our data consist of pre-, mid-, and post-implementation interviews and focus on the differences in implementation across units as a function of institutional logics and evolutionary psychological drives. An empirical paper will be submitted for feedback to *ECIS* this December, and a full-length manuscript targeted to an A journal next year (2015).

7. "Review of Identity Work in Information Systems Research" (with Marie-Claude Boudrea and Christina Serrano) – This paper extends the review work done for a recent *Information and Organization* publication ("IT-Driven Identity Work: Creating a Group Identity in a Digital Environment") into a standalone literature review. We will submit this work-in-progress to *MISQ* in late 2014/early 2015.

8. "If You Build It, Will They Come? Examining the Design and Implementation Challenges of User Innovation Communities" (with Allen Johnston, Paul di Gangi, and Anthony Hood) – The widespread adoption of crowdsourcing and open innovation techniques within the business environment has led many

organizations to discover the value of incorporating users into the innovation processes. While there have been many successful examples of such practices (e.g., SAP, Proctor & Gamble, Dell, and Starbucks), initiating, designing, and implementing user-driven innovation strategies can be a complex challenge to an organization well accustomed to managing its innovation practices internally and outside pubic scrutiny. This article identifies the challenges organizations are likely to encounter during the initiation, design, and implementation stages of establishing a UIC. Through three case examples, we illustrate these challenges and present mitigation techniques to ensure successful implementation. This piece is targeted to *MISQ Executive*.

9. Business Models of User Innovation Communities: A Business Model Perspective (with Allen Johnston, Paul di Gangi, and Anthony Hood) – This practitioner piece reviews of the literature on open innovation by focusing exclusively on the business models underlying a particular type of open innovation initiative—online user innovation communities (UICs), then draws on emerging theory and practice on business model innovation to organize our literature search according the framework of the business model canvas (Osterwalder & Pigneur, 2010), a tool gaining widespread acceptance for promoting both value capture and creation from innovation efforts. We will submit this paper to *AoM* and, based on feedback, will target a management practitioner journal.