Information Technology: Doctoral Theses

DOCTORAL THESIS TITLE:

"Essays on Collective Intelligence" by Yiftach Nagar (2016)

COMMITTEE:

Thomas Malone (Chair), MIT Sloan; Jeffrey V. Nickerson, Stevens Institute of Technology; Iyad Rahwan, MIT Media Lab

ABSTRACT:

This dissertation consists of three essays that advance our understanding of collective-intelligence: how it works, how it can be used, and how it can be augmented. I combine theoretical and empirical work, ranging from qualitative inquiry, to lab experiments, and design, in order to explore how novel ways of organizing, enabled by advancements in information technology, can help us work better, innovate, and solve complex problems. The first essay offers a collective sensemaking model to explain structurational processes in online communities; the second essay offers a way to improve predictions in semi-structured environments by combining human and machine intelligence; the third essay offers a way to accelerate the review of complex intellectual artifacts in open innovation challenges.

In the first essay, I observe conversations that take place as Wikipedia members negotiate, construct, and interpret its policies. Logs of these conversations offer a rare – perhaps unparalleled – opportunity to track how individuals, as they try to make sense, engage others in social interacts that become a collective processes of sensemaking. I draw upon Weick's model of sensemaking as committed-interpretation, which I ground in a qualitative inquiry into policy discussion pages, in attempt to explain how structuration emerges as interpretations are negotiated, and then committed through conversation, and as they are reified in the policy. I argue that the wiki environment provides conditions that help commitments form, strengthen and diffuse, and that this, in turn, helps explain trends of stabilization observed in previous research. The proposed model may prove useful for understanding structurational processes in other large wiki communities, and potentially in other radically open organizations.

In the second essay, we characterize a class of semi-structured prediction problems, where patterns are difficult to discern, where data are difficult to codify or quantify, and where changes occur unexpectedly. These prediction problems are quite common, yet can be extremely difficult to predict, and are often associated with high stakes. Although prior research has shown that in relatively stable environments, statistical models provide better predictions than human experts, and albeit recent advancements in artificial intelligence, we argue that humans can still play an important role together with computers in improving predictions. We conducted laboratory experiments in which we used prediction markets, human judgment, and averaging to combine predictions from groups of people and artificial intelligence agents. We find that the combined predictions were both more accurate and more robust than predictions made by groups of only people or only machines. We discuss the appropriateness of these methods in different contexts.

The third essay addresses a critical bottleneck in open-innovation systems: the process of reviewing and selecting the best submissions. This bottleneck is especially problematic in settings where expertise is needed to review large volumes of complex intellectual artifacts. To help reduce the review load from experts, we offer a computational approach that relies on analyzing sociolinguistic and other characteristics of submission text, as well as activities of the crowd and the submission authors, which can help filtering low quality submission and prioritizing work. We developed and tested models based on data from contests done in the Climate CoLab – a large citizen science platform that hosts innovation challenges – and find that they are able to accurately predict expert decisions about the submissions, and can potentially lead to substantial reduction of review labor, and accelerate the review process.

DOCTORAL THESIS TITLE:

"Using Web Data and Services: Technology, Theory and Evidence" by Xitong Li (2014)

COMMITTEE:

John Hauser (co-chair), Stuart Madnick (co-chair), Erik Brynjolfsson

ABSTRACT:

Nowadays many firms and individuals have been publishing online data and services. For example, Xignite.com has published a number of Web services that provide financial data. Groupon.com and LivingSocial.com are constantly publishing the sales data on their webpages. TripAdvisor.com and Yelp.com provide a tremendous amount of user-generated reviews about hotels and restaurants. This thesis aims to explore (1) what technology could be developed to facilitate using online data and services, and (2) what economic theory and mechanisms are driving the impact of using online data and services. Specifically, the thesis is composed of three essays that describe the technology, theory and empirical evidence about using online data and services.

In the first essay, we present a comprehensive classification of the data misinterpretation problems that may occur when using online services and develop an approach to automatic detection and reconciliation of data interpretation conflicts in Web services composition. The approach uses a lightweight ontology augmented with modifiers, contexts, and atomic conversions between the contexts. The WSDL descriptions of Web services are annotated to establish correspondences to the ontology. Given the naive Business Process Execution Language (BPEL) specification of the desired Web services composition with data interpretation conflicts, the approach can automatically detect the conflicts and produce the corresponding mediated BPEL. We develop a prototype to validate and evaluate the approach.

In the second essay, we explore how herding and social media word-of-mouth (WOM) drive product sales when consumers are influenced by certain data and information published on the online shopping sites. Using a panel data set consisting of about 500 deals from Groupon.com, we find both herding and Facebook-mediated WOM lead to additional product sales, whereas Twitter-mediated WOM has no significant impact on sales. More importantly, we theorize the interaction effect between herding and social media WOM and show herding and Facebook-mediated WOM are complements in driving product sales. The complementarity supports the current practice of daily-deal sites where both mechanisms are often implemented together. We explore the underlying mechanisms of those effects and find supportive empirical evidence.

In the third essay, I study if online reputation (restaurants' displayed Yelp ratings), which helps update consumers' perception of product value, is a causal factor that affects consumers' decisions of endorsing via Facebook and purchasing products (the restaurants' vouchers). I build a stylized Bayesian learning model and derive the hypotheses: (1) a higher online reputation leads to increased social media endorsements and voucher sales, but only when it is built upon a sufficient amount of review ratings; (2) these effects are greater for restaurants with more reviews; and (3) these effects are greater for restaurants with a larger variance in the review ratings. Interestingly, the third hypothesis contrasts to the predictions by some established theories (e.g., cue diagnosticity theory). I test the hypotheses using data of Groupon and LivingSocial deals and find supportive empirical evidence. In particular, I find the effect of displayed Yelp ratings on Facebook Likes is greater when the variance of ratings is larger, suggesting that perhaps consumers are risk averse in the decision-making of endorsing a product to their peers.

DOCTORAL THESIS TITLE:

"Essays on Social Networks and Information Worker Productivity", Lynn Wu (2011)

COMMITTEE:

Erik Brynjolfsson (Chair), Roberto Fernandez, Ray Reagans, Sinan Aral

ABSTRACT:

This thesis is about social networks and their performance implications. It contains three essays.

In the first essay, entitled Social Network Effects on Performance and Layoffs: Evidence from the Adoption of a Social Networking Tool, I study the performance implications from social networks constructed from electronic communication networks. By studying the changes in employees' networks and performance before and after the introduction of a social networking tool, I find that a structurally diverse network (low in cohesion and rich in structural holes) has a positive effect on work performance. The size of the effect is smaller than traditional estimates, suggesting that omitted individual characteristics may bias the estimated network effect. I

consider two intermediate mechanisms by which a structurally diverse network is theorized to improve work performance: information diversity (instrumental) and social communication (expressive) and quantify their effects on two types of work outcomes: billable revenue and layoffs. Analysis shows that the information diversity derived from a structurally diverse network is more correlated with generating billable revenue than is social communication. However, the opposite is true for layoffs. Friendship, as approximated by social communications, is more correlated with reduced layoff risks than is information diversity. Field interviews suggest that friends can serve as advocates in critical situations, ensuring that favorable information is distributed to decision makers. This, in turn, suggests that having a structurally diverse network can drive both work performance and job security, but that there is a tradeoff between either mobilizing friendship or gathering diverse information. Furthermore, it is important to examine the mechanisms by which social communications reduce the risks of being laid off. If social communications promote team effectiveness, delegating decisions rights to managers is optimal. However, if managers choose to optimize their own power at the expense of the firm, the positive impact of social communications on layoffs is evidence that delegating layoff decisions to managers can incur important costs.

In the second essay, entitled Water Cooler Networks: Performance Implications of Informal Face-to-Face Interaction Structures in Information Intensive Work, examines the performance implications of face-to-face social networks. We find that their structural properties create unique implications for effective knowledge transfer and productivity. We argue that network theory should incorporate implications of media choice, and particularly differences between face-to-face and electronic communication, when assessing how networks affect individual performance. We introduce a new methodology, using Sociometric badges, to record precise data on face-to-face interaction networks for a group of workers in a large IT manufacturing firm over a one-month period. Linking these data to detailed performance metrics, we find that 1) network cohesion is associated with higher worker productivity, in contrast to previous findings in email data; 2) cohesion in face-to-face networks is associated with even higher performance during complex tasks, suggesting that cohesion complements information-rich media for transferring the complex knowledge needed to complete complex tasks; 3) while information seeking from many colleagues creates disruptions, more interactions with a few key strong tie informants speeds up work. Face-to-face networks have more explanatory power than physical-proximity networks, suggesting that information flows in actual conversations (rather than individuals' correlated exposure to common environmental factors through physical proximity) are driving our results. These results augment our understanding of how media choice and network structure interact, shedding light on the organizational implications of face-to-face interaction. The methods and techniques we introduce are replicable, creating opportunities for new lines of research into the implications of face-to-face interactions in organizations.

In the third essay, entitled Identification of Influence: An Experimental Platform for Understanding the Relationship between Social Networks and Performance, I create an experimental platform for identifying the relationship between social networks and performance. While a large body of literature has examined the correlations between certain network topology and performance, little research has shown a definitive causal link between social network and productivity. I address this problem through conducting three sets of randomized field experiments through an on-line experimental platform at a large information technology firm. The platform enables randomly selected employees to achieve certain network characteristics. By examining work performance before and after the experiment, I hope to tease out the causal linkage between networks and productivity. Furthermore, I plan to distinguish the type of employees (e.g. peripheral actors) that could benefit the most from a change in network structure.

DOCTORAL THESIS TITLE:

"Essays on Information Technology and Intangible Capital", Adam Saunders (2011)

COMMITTEE:

Erik Brynjolfsson (Chair), Tom Malone, Lorin M. Hitt (Wharton)

ABSTRACT:

This thesis consists of three essays related to information technology and intangible capital. The first essay, "Valuing IT-Related Intangible Capital," examines the value of intangible assets in the firm. Using a panel of 130 firms from 2003-2006, we find that intangible assets are correlated with significantly higher market values beyond their cost-based measures. Moreover, we estimate that there is a 30-55% premium in market value for the firms with the highest organizational IT capabilities as compared to those with the lowest

organizational IT capabilities. The second essay, "Has Information Technology Leveled the Competitive Playing Field?" analyzes the relationship between IT and ordinary (non-IT) capital and the competitive dynamics within U.S. industries. Using a panel of industry data from 1998-2005, when an industry becomes more IT intensive, there is more entry and expansion of firms (including entry of new small firms and expansion of large firms from the same and other industries). Yet there is also more turnover of small firms in the industry as well as concentration of the industry into large firms. In contrast, as an industry becomes more ordinary capital-intensive, there is less entry of small firms and fewer establishment openings by large firms; a lower rate of turnover by small firms; and fragmentation of the industry into small firms. In the third essay, "The Value and Durability of Patents in High-Tech Firms" (co-authored with Erik Brynjolfsson and Lorin Hitt), we use data on publicly traded high-tech companies from 1984-2002 to examine the relationship between the firms' market value and their patent-based intangible assets. We find that high-tech firms with patents that are cited by a wide variety of other patents in different patent classes are worth significantly more than firms with patents that are cited by a narrow range of patents. Patent generality is especially valuable in periods of change, when firms are no longer at the leading edge of innovation in a particular year. In these periods, we find that the value of diverse patents across technology categories is positive but not significant and that generality is comparatively more valuable than diversity.