Computational Social Sciences

Elisa Bertino
Sorin Adam Matei *Editors*

Roles, Trust, and Reputation in Social Media Knowledge Markets

Theory and Methods

Springer

Computational Social Sciences

A series of authored and edited monographs that utilize quantitative and computational methods to model, analyze, and interpret large-scale social phenomena. Titles within the series contain methods and practices that test and develop theories of complex social processes through bottom-up modeling of social interactions. Of particular interest is the study of the co-evolution of modern communication technology and social behavior and norms, in connection with emerging issues such as trust, risk, security, and privacy in novel socio-technical environments.

Computational Social Sciences is explicitly transdisciplinary: quantitative methods from fields such as dynamical systems, artificial intelligence, network theory, agent-based modeling, and statistical mechanics are invoked and combined with state-of-the-art mining and analysis of large data sets to help us understand social agents, their interactions on and offline, and the effect of these interactions at the macro level. Topics include, but are not limited to social networks and media, dynamics of opinions, cultures and conflicts, socio-technical co-evolution, and social psychology. Computational Social Sciences will also publish monographs and selected edited contributions from specialized conferences and workshops specifically aimed at communicating new findings to a large transdisciplinary audience. A fundamental goal of the series is to provide a single forum within which commonalities and differences in the workings of this field may be discerned, hence leading to deeper insight and understanding.

For further volumes:
http://www.springer.com/series/11784

Elisa Bertino • Sorin Adam Matei
Editors

Roles, Trust, and Reputation in Social Media Knowledge Markets

Theory and Methods

Springer

Editors
Elisa Bertino
Computer Science
Purdue University
West Lafayette
Indiana
USA

Sorin Adam Matei
Brian Lamb School of Communication
Purdue University
West Lafayette
Indiana
USA

ISBN 978-3-319-05466-7 ISBN 978-3-319-05467-4 (eBook)
DOI 10.1007/978-3-319-05467-4
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014949192

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Specialized knowledge and expertise, especially of the kind that can shape public opinion, have been traditionally conceived to be the domain of individuals holding degrees awarded by higher learning institutions or occupying formal positions in notable organizations. Their expertise is validated by reputations established in an institutionalized marketplace of ideas with a limited number of "available seats" and a stringent process of selection and retention of names, ideas, topics, and facts of interest. However, the social media revolution, which has enabled over 2 billion Internet users not only to consume, but also to produce information and knowledge, has created a secondary and very active informal marketplace of ideas and knowledge. Anchored by platforms like Wikipedia, YouTube, Facebook, and Twitter, this informal marketplace has low barriers to entry and has become a gigantic, and potentially questionable, knowledge resource for the public at large.

The availability of these new knowledge markets poses some important research questions concerning the ways in which knowledge producers and users interact, and how knowledge is created and evolves. Credibility and quality of such knowledge is also a critical issue. Notions such as expertise and reputation need new definitions and metrics. Tools and methodologies to carry out research to answer those questions are needed, perhaps based on extensive data analyses. With the goal of creating and fostering a research community able to answer these questions, and also to identify novel related research directions, the US National Science Foundation founded the KredibleNet project—a multidisciplinary project involving researchers from different disciplines including computer science, social sciences, and statistics.

This book is the result of the first invitational KredibleNet workshop, which was held at Purdue University in April 2013. The workshop was an interactive forum to present the latest theoretical and methodological advances related to social media social roles, structures, and reputation research.

The workshop included sessions of "featured" presentations, followed by discussions, and two "round table" sessions, in which the discussants proposed possible future research agenda items related to social media roles, authority, and trust.

The core workshop topics included:

1. What are the most important existing or emerging Social Media knowledge markets and what distinguishes their working mechanisms?
2. What is the social structure of the emerging social media knowledge markets and what are the main motivator factors that fuel the individuals that are central in these social structures?
3. Which approaches based on social network analysis techniques can be used for defining "expert" reputation in informal marketplaces of ideas?
4. Which social graph topological configurations are associated with specific functional roles and levels of reputation in social media knowledge markets?
5. How do functional roles, reputation and authority emerge on social media knowledge generation projects and how can they be operationalized, measured and explained?
6. How does trust and knowledge credibility connect to specific functional roles and authority structures?
7. How can a new theoretical understanding of credibility, roles, and trust be turned into specific actionable tools and approaches to moderating knowledge market?
8. What are the most promising yet under-researched areas in the field of social media knowledge markets, especially with respect to authorship and reputation?

Workshop presenters were invited to author chapters for the book. As a result, the book represents a comprehensive research coverage concerning questions of trust and reputation in the new knowledge market.

The book is organized is several parts. The first part introduces the book and consists of two chapters. Chapter 1, titled "A Research Agenda for the Study of Entropic Social Structural Evolution, Functional Roles, Adhocratic Leadership Styles, and Credibility in Online Organizations and Knowledge Markets" by Sorin Matei at al., provides an overview of the KredibleNet project and the research agenda that has been formulated as part of the project. Notable research directions outlined as part of this agenda include how to use network analysis techniques for modeling functional roles and reputation, how to assess the stability of leading functional roles, and how to extend data analytics and statistic methods for functional role analysis. The chapter also introduces the novel concept of social entropy as one of the metrics for modeling collaborative spaces. In this context, entropy explains the degree of social organization in knowledge building spaces by measuring contribution inequality. Chapter 2, titled "Building Trusted Social Media Communities: Organizations, Motivation, Reputation" by Ben Shneiderman introduces several basic concepts including reputation, trust, and credibility and discusses how these concepts form the foundation for credible web-based communities.

The second part of the book focuses on methods for researching trust and credibility and consists of four chapters. Chapter 3, titled "*Semantic and Social Spaces: Identifying Keyword Similarity with Relations*" by Yun Huang et al. addresses the problem of identifying the expertise and topics of individuals participating in knowledge networks. Such information is critical in order to assess the quality of information in those networks. The chapter then proposes an approach based on semantic tagging and text analysis. Chapter 4, titled "*Emergent Social Roles in Wikipedia's Breaking New Collaborations*" by Brian Keegan focuses on temporal patterns of activities and collaborations of Wikipedia editors when dealing with breaking news. The chapter, based on an extensive analysis of four case studies, identifies several different patterns followed by editors of such news. Chapter 5, titled "*Words and Networks: How Reliable are Network Data Constructed from Text Data*?" by Jana Diesner, focuses on the key problem of designing methods supporting the reliable construction of network datasets that are then used for research in computational social networks. Chapter 6, titled "*Predicting Low-Quality Wikipedia Articles Using User's Judgments*" by Ning Zhang, Lingun Ruan, and Luo Si, investigates the problem of assessing the quality of Wikipedia contents. The approach make uses of the Wikipedia reader feedback data to build a regression model able to predict the quality of articles.

The third part of the book focuses on tools for increasing trust and transparency and consists of Chap. 7, titled "*From Invisible Algorithms to Interactive Affordances: Data after the Ideology of Machine Learning*" by Bernie Hogan. The chapter addresses the important issue of how to support user navigation in information networks. The chapter introduces two different approaches based on two different "ideologies" and discusses how these approaches can be used for reputation analysis. An interesting point made by the chapter is that the "dominant ideology" used for information presentation is based on sorting and that this "ideology" is not well suited to the study of reputation and credibility.

The fourth part of the book focuses on novel research directions. It consists of three chapters. Chapter 8, titled "*Iron Law of Oligarchy: Computational Institutions, Organization Fidelity, and Distributed Social Control*" by Howard Welser, makes the point that recent developments in on-line communities and social networks can help overcoming the tendency that all organizations have in structuring themselves as oligarchies. The chapter elaborates on challenges and on the fundamental design elements needed to achieve a distributed control in organizations. The author proposes that a solution to the tendency of organizations to suffer mission drift and to allow the top agents of power to exploit it to their own advantage is to share dependence on the success of the organization across all agent roles, from the top ones to the rank and file. More importantly, he proposes a comprehensive system of contribution monitoring, visualization, and conditioning of rewards on inputs introduced in the system. Chapter 9, titled "*Cultural Differences in Social Media: Trust and Authority*" by Mei Kobayashi, makes the point that people with different cultures may exhibit different behavior in cyber space with respect to the perception of trust and reputation and therefore large-scale studies are needed to better assess the impact of cultural differences. The chapter also identifies the applications that may benefit for such an assessment and includes an extensive review of existing work. Chapter 10, titled "Convincing Evidence" by Andrew Gelman and Keith O'Rourke, focuses on statistical tools and makes the important observation that authorship, reputation, credibility and past experience play an important role in decisions about statistical procedures. The chapter also elaborates on issues related to the use of big data in research.

The book is concluded by Chap. 11, titled “The Trajectory of Current and Future Knowledge Market Research: Insights from the First KredibleNet Workshop” by Sorin Adam Matei, Brian Britt, Elisa Bertino, and Jeremy Foote, which reports the results of the discussions at the workshop with the goal of organizing such discussions into a research roadmap. Two broad research areas emerged during the workshop focusing respectively on theoretical frameworks and on methodologies to assess these theories. For each such area, the chapter covers the current state of the art and promising research directions.

As this book is result of a multidisciplinary effort to assess the current state of the art and identify novel research directions, we trust that the reader will find in the book interesting and novel research perspectives.

Purdue University, Lafayette, IN, USA

Elisa Bertino
Sorin Adam Matei

Acknowledgments

The work reported in this book and the KredibleNet workshop have been partially supported by the US National Science Foundation under award number 1244708 and a Purdue Discovery Park Fellowship. We would like also to thank Dr. Richard Buckius, Purdue Vice President for Research, and Dr. Alan Rebar, Director of Purdue Discovery Park, for their constant support to research in the area of computational social science.

Contents

Contributors

Elisa Bertino Purdue University, West Lafayette, IN, USA

Brian Britt Purdue University, West Lafayette, IN, USA

Noshir Contractor The Science of Networks in Communities (SONIC) research group, Northwestern University, Evanston, IL, USA

Jana Diesner GSLIS/UIUC, Champaign, IL, USA

Jeremy Foote Purdue University, West Lafayette, IN, USA

Andrew Gelman Columbia University, New York, NY, USA

Bernie Hogan Oxford Internet Institute, University of Oxford, Oxford, UK

Yun Huang The Science of Networks in Communities (SONIC) research group, Northwestern University, Evanston, IL, USA

Brian C. Keegan Northeastern University, Boston, USA

Mei Kobayashi IBM Research, Koto-ku, Tokyo, Japan

Baozhen Lee School of Economics and Management, Jiangsu University of Science and Technology, Zhenjiang, China

Chuanhai Liu Purdue University, West Lafayette, IN, USA

Sorin Adam Matei Purdue University, West Lafayette, IN, USA

Keith O'Rourke O'Rourke Consulting, Ottawa, ON, Canada

Lingyun Ruan Google, Mountain View, CA, USA

Ben Shneiderman University of Maryland, College Park, MD, USA

Luo Si Purdue University, West Lafayette, IN, USA

Department of Computer Science, Purdue University,West Lafayette, IN, USA

Cindy Weng The Science of Networks in Communities (SONIC) research group, The Science of Networks in Communities (SONIC) research group, Northwestern University, Evanston, IL, USA

Howard T. Welser Department of Sociology and Anthropology, Ohio University, Athens, OH, USA

Ning Zhang Department of Computer Science, Purdue University,West Lafayette, IN, USA

Michael Zhu Purdue University, West Lafayette, IN, USA

Part I
Introduction

Chapter 1
A Research Agenda for the Study of Entropic Social Structural Evolution, Functional Roles, Adhocratic Leadership Styles, and Credibility in Online Organizations and Knowledge Markets

Sorin Adam Matei, Elisa Bertino, Michael Zhu, Chuanhai Liu, Luo Si and Brian Britt

Introduction

The new social media enabled by the Internet and the Web have deeply changed the ways in which individuals interact and how knowledge is created and exchanged, which is opening up interesting new research questions for social science. A key question is how the notions of expertise and reputation will evolve as a consequence of the emergence and broad use of social media. Addressing such questions is crucial for many different domains, from traditional academic settings and processes (e.g., promotion procedures), to research funding (e.g., assessing the impact of research results), homeland security and intelligence (e.g., detecting campaigns aiming at spreading deceiving information), and healthcare (determining the source and credibility of health information on the Net).

At the same time, the fact that communication has been migrating to social media makes it possible to collect extensive datasets for use in research. A major problem, however, is that just having huge collection of datasets, which document human interactions in detail, is not sufficient. We also need data management and analytical tools that can support timely, effective, and efficient knowledge extraction processes from such data. In general, currently available tools have not been designed to deal with massive interaction data, and they are particularly unable to deal with specific questions concerning expertise and reputation. Understanding which new tools we need and how to design and build these tools requires input from a broad multidisciplinary community involving experts from different research communities, including the social sciences, computer sciences, and statistics.

KredibleNet was designed as a broad, multidisciplinary community effort focused on researching expertise and reputation in the new social media, and on designing and

Pre-print draft—to appear in 2014 in Matei, S. A. and Bertino, E. *Roles, trust, and reputation in Social Media Knowledge Markets: Theories and Methods.* Contract Signed with Springer Verlag.

S. A. Matei (✉) · E. Bertino · M. Zhu · C. Liu · L. Si · B. Britt
Purdue University, West Lafayette, IN, USA
e-mail: smatei@purdue.edu

E. Bertino, S. A. Matei (eds.), *Roles, Trust, and Reputation in Social Media Knowledge Markets,* Computational Social Sciences, DOI 10.1007/978-3-319-05467-4_1,

building new large-scale data analysis and management infrastructures to support this research. In what follows, we discuss the fundamental questions that KredibleNet hopes to answer, with a review of the previous research and an explanation of our research agenda, including details regarding the datasets we can offer the research community and the possible strategies to explore them. Next, we describe the roadblocks, and explain why we believe that these questions can only be resolved via a broad and interdisciplinary approach. We also discuss our current efforts on dealing with a few of these roadblocks. Finally, we discuss what we see as the future directions for knowledge market research, and our goals for how the KredibleNet community can be involved in that research.

In essence, this chapter presents a summary of the assets and of the visions that we will bring to bear to make them valuable to a broader community of scholars and practitioners as well as a way to document and explain present and future challenges.

Knowledge and Expertise and the New Social Media Knowledge and expertise, especially of the kind that can shape public opinion, have been traditionally perceived to be the domain of individuals who hold degrees awarded by higher-learning institutions, or those who occupy formal positions in notable organizations. Their expertise is validated by reputations established in an institutionalized marketplace of ideas with a limited number of "available seats" and a stringent process of selection and retention of names, ideas, topics, and facts of interest.

With the advent of online communication and social media, however, knowledge creation has become a much more complex process. The communication revolution has enabled over 2 billion Internet users to not only consume, but also to produce information, creating a secondary and very active informal marketplace of ideas and knowledge. Anchored by platforms like Wikipedia, YouTube, Facebook, and Twitter, this informal marketplace has low barriers to entry and has become a gigantic, and for some questionable, knowledge resource for the public at large. Furthermore, the informal mechanisms for knowledge creation and sharing openly challenge traditional notions of authority and reputation sanctified by the institutionalized, expert-based marketplace of ideas (Weinberger 2011).

Reputation and Authority Changes in the nature of reputation and authority represent an important research topic, since such changes fundamentally reshape the knowledge production process. They are of equally great importance for commercial interests, as sites and companies like Klout and Social Mention demonstrate. The search string "online reputation measure" on Google Scholar produces no less than 12,000 scholarly articles within the social sciences alone. Even more impressively, Amazon.com offers no less than 1039 paperback and 355 hard cover books on the subject of "online reputation." Measures for reputation have been proposed and interesting algorithms have been deployed, both by academic and industry researchers (Adler and de Alfaro 2007; Arazy et al. 2010; de Alfaro et al. 2011; Dellarocas 2006; Hasan et al. 2009; Hennis et al. 2011; Jøsang and Golbeck 2009; Kraut and Resnick 2012; Masum and Tovey 2012; Matei et al. 2010a; Welser et al. 2007).

Research Challenges Yet, significant gaps remain to be filled. First, the basic definition of reputation and the metrics proposed for measuring it are not always convertible

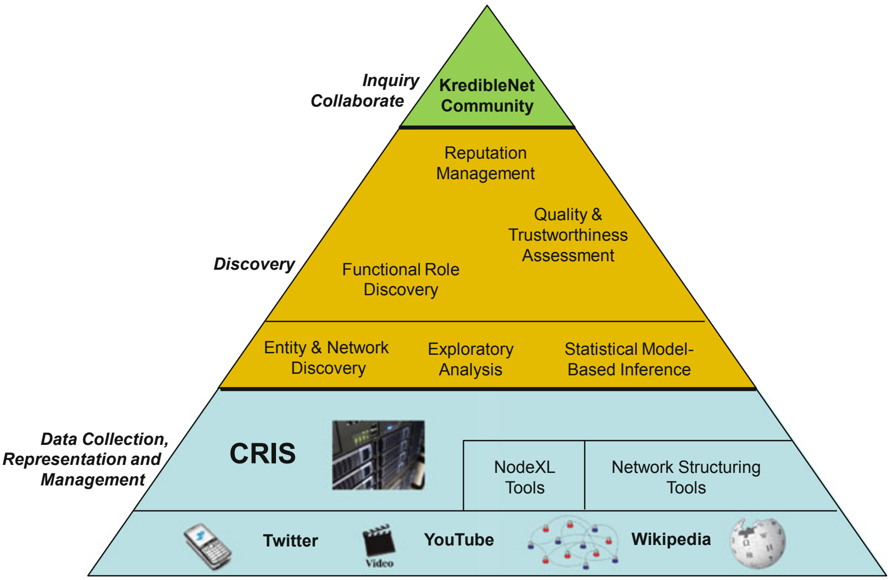


Fig. 1.1 The data-driven discovery and collaboration strategy of KredibleNet

from one domain of reference to another. Second, nonproprietary data, tools, and procedures used to measure online reputations are often not readily available or shareable. Finally, the data-intensive nature of a project that would support research on social media itself presents numerous challenges including data heterogeneity, entity and network discovery, and data size.

Project Goals To meet these challenges, our project aims to create a community of scholars and practitioners interested in defining, measuring, and operationalizing reputation as a new and essential lens for understanding and evaluating the knowledge that is generated and consumed online. Additionally, we aim to develop a new generation of data management and analysis tools and techniques to support reputation research on extremely large datasets, while making these tools and techniques available to spur further research and related activities.

The community will grow through a series of workshops and a collaboratively designed cyber-infrastructure prototype that aims to bring reputation related tools and visualization techniques closer to the public. KredibleNet will serve as an experimental laboratory for the research community (Fig. 1.1), and will offer live social media data to researchers who wish to experiment with data-intensive analytic strategies to better define the nature of and extend the reputation measurement problem.

Techniques and Tools Answering the many research questions about credibility and reputation will require the development of novel techniques, which will be used as building blocks for the broader project, and for the transformation of these techniques into tools. The tools and techniques will first be defined using the Wikipedia revision dataset (Britt 2011), which contains 250 million data points representing all of the edits made to the Wikipedia articles throughout its first 9 years of life (Adler and de Alfaro 2007). These tools and techniques will then be applied to other available social media datasets including Twitter and YouTube. Novel techniques that we will investigate in the project include:

- Techniques for mapping the graph of editorial, communicative, and semantic interactions between contributors
- Techniques for entity discovery (rule- and learning-based) that aim to decompose the graph of editorial collaboration into its basic building blocks
- Techniques to uncover structures of interaction utilizing social entropy theory and related methodologies (Matei et al. 2010a; Matei et al. 2010b)
- Techniques for associating social network building blocks with functional roles and reputations
- Techniques for predictive quality rating of wiki content and for associating quality with contributions from individuals who perform specific functional roles that are associated with measurable reputations
- Techniques for predicting the development of social media projects over time

A central analytic effort is directed at differentiating recurring patterns of interaction within an intact social media interaction graph, namely that created by the Wikipedia editors over a period of 9 years. Such patterns (building blocks) would, in effect, be subnetworks associated with specific functional roles and reputations, as performed by the people who contribute to Wikipedia. These building blocks will be incorporated into the tools able to assess reputation and credibility not only with respect to a single medium, but also across all three media (Wikipedia, Twitter, and YouTube). At the same time, we intend to uncover the degree to which Wikipedia's collaborative structures have emerged by utilizing social entropy and network statistics (such as betweenness centrality) as main investigation tools (Matei et al. 2010a; 2010b). A core concern is to understand the socio-evolutionary dynamics of knowledge production spaces.

Other Applications In addition to the research area described above, the proposed infrastructure and tools can be used to support the research in the area of security for social networks, determining, for example, how wiki and social media spam attacks and campaigns originate and spread. We anticipate that the target data will be organized as graphs, and auto-correlative processes under conditions of uncertainty, and as such, the infrastructure and tools under development can be used for any research area that requires integrating and analyzing graph or auto-correlative data via Bayesian methodologies.

Problem Space Overview As important as online reputation of the contributors to social media projects is, the concept has yet to be rigorously defined and measured. After all, reputation is typically associated with fame and public recognition. Yet, as we will argue in more detail below on social media, reputation might be seen as a function of the amount and frequency of contributions multiplied by the velocity at which the content is disseminated. Recognition is implicit, defined by the "viral impact" of content shared or contributed. Such impact presupposes hard work and constant presence online. To gain and maintain a high reputation online, one must perform tasks and deliver tangible results constantly and credibly.

Further, reputation is the product of fulfilling specific functional roles, including a contributor, disseminator, project coordinator, and so forth. It is not a simple individual attribute but a relational entity, predicated on the intensity and diversity of implicit ties that individuals establish by sharing or contributing content online.

If reputation is defined relationally, we need to reconceptualize how it is measured and assigned to actors. Furthermore, a new understanding of reputation needs to be incorporated into tools that measure and visualize its magnitude for each social media, while the meaning of these tools should be communicated to the public effectively. The research agenda that will emerge through our project will propose guidelines for generating tools and services that measure reputation relationally and will make reputation measurements and visualization an integral and essential part of ordinary individuals' online knowledge production and consumption.

Another issue that our project will tackle, which is a challenge in and by itself, relates to collecting and analyzing the data necessary to understand, measure, and assign reputation to social media actors. Measuring and visualizing reputation from social media data requires access to and the ability to handle massive datasets, typically in a graph format. These datasets are several orders of magnitude greater than what a typical social scientific research project deals with and involves complexities that are not yet fully understood. In some of the sections below, we will detail a possible approach that uses network analysis to find clusters of interactions in the middle of which nodes (users, concepts, ideas) with higher levels of reputation can be found.

In addition, we are interested in uncovering the emergence and temporal evolution of interaction structures in collaborative spaces, such as those that are created by people coediting articles on Wikipedia. To this end, we utilize a modified version of social entropy theory (Matei et al. 2010a; 2010b; see also Shannon 1948) by integrating other approaches, including network analytic methodologies. As entropy is a core concept of our broader theoretical approach, and a starting point for our initial exploratory research in this area, including a doctoral dissertation (Britt 2013), it might be useful to detail in a few words how we operationalize it and how it could be useful for this project or cognate projects that will emerge from our research. As proposed by Matei et al. (2010a):

> [s]ocial entropy could be used to measure how structured or unstructured a group is. More specifically, we reformulate Shannon's theory of information to suggest that:
>
> 1. Information and "structure" go in the opposite direction of entropy;
> 2. Information and structure, especially in the social realm, are intrinsically connected; and,
> 3. Structure (of a language, symbol system, or group organization) can be measured with one synthetic indicator, namely entropy.
>
> We emphasize the connections between social entropy and structure because groups are more than mere aggregations of people who share the same space. A group is the structure of ties

> between its individuals. Individuals that occupy specific roles in this structure communicate, contribute or interact in a specific way. The distribution of outputs in the group will follow the curve of abilities, productivity, task and power allocation specific to each role. Employing Shannon's entropy measure to describe group efforts, communicative patterns and collaborative patterns, we expect that as a group becomes more structured (i.e., roles emerge, tasks are assigned or assumed, power and information starts flowing from specific nodes to other nodes), imbalances in the distribution of communication or work will appear.
>
> In other words, as the group starts to form and its structure to emerge, group units (individuals) start behaving in predictable and non-random way. This predictable pattern entails a specific amount of unevenness. It is important to mention that "specific" has no normative meaning in our research. We have no a priori preference for any given level of unevenness, nor do we think that unevenness is demanded by "natural," individual characteristics. Rather, we propose that unevenness, while ever present, is a dynamic group process. Any group member can theoretically occupy any level of contribution or interaction. For each group and type of structure, some of which can be flatter while other more hierarchical, there is a "specific" level of unevenness and social entropy that needs to be observed and explained, not predicated.

To make these research goals a reality, several challenges need to be addressed: (1) transforming contribution data into analyzable interaction graph data, (2) entity and network discovery within the social interaction data, and (3) calculating entropy values and social network statistics over time for the entire collaborative space within the social medium (e.g., Wikipedia). Other less important, but still relevant, challenges are represented by managing large social media data (in our case, databases containing hundreds of millions of data points), new statistical procedures for analyzing noisy network data, and data integration.

Expected Results To meet these challenges, our project aims to create a community of scholars and practitioners who seek to define, measure, and operationalize reputation as a new and essential component of the knowledge that is generated and consumed online. The community will grow around a series of workshops and a collaboratively designed reputation measurement and dissemination platform prototype, which aims to bring reputation-related tools and visualization techniques closer to the public. The platform will mine and serve social media data collected on a large scale (over 250 million data points) and will become an experimental laboratory for the research community. It will offer live social media data for the development of data intensive analytic strategies to better define the nature and extension of the reputation measurement problem.

The research program that will emerge from our activities will produce the following major results:

1. Heuristic tools and strategies for detecting under what conditions functional roles and reputations emerge in social media knowledge spaces, how their footprints can be captured and measured, what their main characteristics are, and how the networks in which they are embedded influence knowledge production and consumption. Social entropy and network analytic methods and tools for understanding how social media collaborative structures emerge will also be central to our effort.

2. A blueprint and an early prototype for a set of integrated online data analytic and integration services to monitor the flow of social media more broadly. These will provide actionable intelligence on where functional nodes and roles appear and what their potential reputation and impact could be. A platform that explores the use of social entropy for measuring interactions online has been produced and can be experimented with at http://veffort.us.
3. A collaborative cyberinfrastructure platform for researchers, public policy experts, media, or citizens to pose questions, and to interact with each other in the process of mining social media. This platform will further provide insight about the functional roles, reputations, and actors that continuously emerge in social media.

The Workshops and the Publications To successfully build a community that can directly engage and address the described challenges, our research program includes two workshops covering a distinct set of research topics. The first one is dedicated to social-scientific research methods in the context of social media, while the second one focuses on computer scientific and statistical analyses that can help in understanding social media.

Each workshop will be conducted using a time-tested academic process wherein topics will be identified and distributed, and then papers will be solicited for submission and review. Papers will be presented with moderated discussions surrounding the topic captured by assigned members of each group. All content is to be curated for a workshop report which will include a roadmap for action and the guidelines for social media reputation research. The reports and selected papers will be published in an edited volume of the Computational Social Science series at Springer Publishing House.

The KredibleNet Cyberinfrastructure The workshop results will support building a cyberinfrastructure upon the extensible Purdue *The Computational Research Infrastructure for the Sciences* (CRIS) research infrastructure. This cyberinfrastructure will be designed to make reputation-related tools, visualization techniques and results, and social network data freely and openly accessible. In particular, CRIS will integrate the tools and the data made available by the NodeXL project. The NodeXL Graph Gallery Website (http://nodexlgraphgallery.org) makes tools for improving the variety and quality of data available to the social network analysis community. The collections of images and datasets on the site have created a learning community that exchanges best practices and improves on them over time. Related projects like ThreadMill, and tools to extract networks from data sources like Facebook and message boards, significantly expand the scope of this learning community.

Novel Contributions

KredibleNet is one of the first projects to develop a data-intensive approach for cross-disciplinary research in the area of reputation and credibility in social media. As such, it aims to achieve major advances in several disciplines, including the social sciences, communication, computational social science, data analytics, and data management and security. Specific novel research contributions include:

- Techniques for querying raw knowledge construction and coeditorial interactions in social media, including techniques for structuring and defining interactions relationally, attaching metadata to interactions, tracking interactions over time, and measuring system level states and interaction
- Techniques for integrating nontextual data in the collaborative corpus and explaining the knowledge-based interactions that surround such data
- Approaches based on social network analytic techniques for defining "expert" reputation in informal marketplaces of ideas as a functional (achieved) role, not as an ascribed (institutionally sanctioned) role. As functional roles are seen as nodes of interaction in a network graph, metrics will also be defined in terms of graph measures and features including:
  - The definition and measurement of social media reputation as a relational (graph) phenomenon
  - The identification of topological configurations that are associated with specific functional roles and levels of reputation
  - Statistical measures that can be used to define the roles of connections (autocorrelative interactions, thus reputation related) in shaping outcomes (individual productivity, power in the network, etc), such as autocorrelation-based regression models
  - Topological characteristics that make some nodes more generative of new subnetworks with specific impact on knowledge production and gaining reputation
- Entity discovery strategies for defining privileged nodes and subnetworks
- Strategies for translating collaboration graphs across domains, adapting graph measures to each domain and defined translational measures, and calibrating reputation across domains (e.g. defining a common unit of measurement)
- Techniques for calculating social entropy values, as they change over time, for entire collaborative spaces (e.g., Wikipedia)

Research Roadmap

While the project is primarily focused on creating a community of researchers interested in redefining online reputation and better understanding knowledge production via social media, the synergies that the project creates are channeled to address a number of very specific research challenges. In what follows, we expand on the core challenge of defining reputation as a function of functional roles and on using social entropy as a collaborative metric. Our approach, discussed below, is to view this problem as a longitudinal network analysis problem within a dynamic social system framework. In developing our approach, we will also address other subsidiary challenges, especially those related to constructing network analytic tools.

Functional Roles and Reputations as a Network Analysis Problem

Reputation, which is one of the core concepts of our project, is by definition a relational construct. Any social media actor has a publically acknowledged and measurable impact on the communication space with which he or she interacts.

In the context of our project, reputation focuses on the amount of "social capital" (number and strength of coeditorial ties) or the "social weight" an actor has in a computer-mediated network of interactions (Britt 2011; Coleman 1988; Lochner et al. 1999; Matei 2004). With the emergence of social media, reputation becomes more than public acknowledgment of fame; it is, in fact, an attribute of an achieved position in a network of interactions.

People are influential and important by the amount of content they share, contribute to, or manipulate. As all these behaviors are relational, they exist through and in a network of interactions. Furthermore, the number of relationships, their direction, intensity, diversity, and the specific locations of the individuals connected by those relationships within the broader network topology illuminate the specific roles individuals play in the network (Britt and Matei forthcoming). In other words, knowing the topology of a specific network of social media interaction allows us to derive the functional role and reputation of each node or individual.

Under this understanding of the concept, reputation becomes a social structural concept that can be measured using social network analysis. Reconceptualizing reputation in this manner requires a substantial amount of theoretical and methodological work and likely represents the greatest challenge that our project addresses. In addition, we are interested in conceptualizing social structure as a social entropy problem and in detecting temporal trends in the emergence of social structures at the macrosystem level (e.g., the entire Wikipedia corpus). This line of research also broadly lays the groundwork for an investigation into the nature of collaborative structures as well as, more specifically, the presence of functional leadership groups and the likelihood that specific individuals are quasi-permanent members of such groups. In what follows, we present the research questions, preliminary data, approaches, and methodologies that we will employ to meet these challenges.

The Research Questions

a. How can online reputation in informal knowledge markets be defined and measured?
b. Can the concept of "functional role," which is defined by the interactions in an online network of contributions to a social media project, provide sufficient conceptual content and measurable dimensions for redefining the idea of "reputation" in informal knowledge markets?
c. What is the impact of "reputation," as redefined by this project, on network growth, knowledge production, and content quality?
d. How do functional leadership groups emerge and how likely they are to be "sticky" (to include a quasi-permanent group of individuals)?
e. What is the structural evolution of a social media space, as captured by social entropy measures?

The Challenge Addressing our research questions requires constructing a network of interaction from simple contributions to an informal knowledge marketplace, which by itself represents a major challenge. This approach is demanded by the emergence of social media—online communication and knowledge creation environments that rely on user contributions—such as Wikipedia, Facebook, or YouTube, or Twitter, which have created invaluable opportunities for tracking and reconstituting intact and exhaustive networks of interaction.

Such systems distinguish themselves through three essential characteristics: self-organization, a quasi-complete record of the social and intellectual transactions that support self-organization, and services that allow systematic retrieval of these interactions as semi-structured graph (who interacts with whom) data. These characteristics allow us to generate network graphs representing the manner in which communication and collaboration are structured into networks of communication, interaction, or work.

Such graphs may represent a wide variety of collaborative and knowledge dimensions. They may run the gamut from direct, person-to-person communicative exchanges to implicit connections that can be discerned from coeditorial activity. However, interaction on social media is often based on common work on or with a body of content, rather than on direct, person-to-person interaction.

As an example in case, when we retweet a message, add new content to a Wikipedia article, or share a picture, we implicitly "collaborate" with other people. We do not address them by name or often even suggest that they are the explicit targets of our sharing or editing work, yet we ultimately engage their attention or work, and a specific type of mediated interaction is constructed. The core research challenges of our project entail understanding these interactions, the social benefits derived from interacting in such a manner with a network of invisible collaborators, the ways in which the interaction space grows and how it can be measured, and the implicit reputation and influence on the social media space one gains from these same interactions.

Research Datasets and Methodologies Our research challenge will be met by gathering a community of scholars and applying the insights and guidelines generated through their interactions within our project on a prototype of an analytic environment. This cyberinfrastructure will take advantage of a complete database of implicit interactions in a given social media project—namely, the revisions made to Wikipedia articles from 2001–2010. This is an essential component of our project, as it is a unique network graph dataset containing all implicit collaborative interactions of over 21 million editors (roughly 3 million logged-in users and 19 million editors identifiable by their IP addresses) who have contributed to 7 million Wikipedia articles. This dataset will be available to the research community established by KredibleNet as a working laboratory.

According to Alexa.com, Wikipedia is the sixth most visited website in the world, attracting 15 % of the world Internet population every day. Further, more than half of the US students use Wikipedia on a regular basis for coursework (Miller 2010). Therefore, Wikipedia can be considered one of the most important informal knowledge markets in the world. Its ability to shape public opinion and learning is tremendous.

The Wikipedia data, collected by Luca de Alfaro (Adler and de Alfaro2007) and made available in a network format by KredibleNet, will be used to test and validate research questions related to the presence of specific functional roles, defined as position in a subnetwork of interaction. Subnetworks will be uncovered using entity discovery procedures. Two different approaches will be used to identify subnetworks. The first approach will directly utilize the network layer or the coeditorial network data and use model-based clustering methods for uncovering groups of editors that form subnetworks. The particular statistical network models used are stochastic block models and latent space models. The second approach relies on the user log layer or the user contribution data, and it consists of several steps. In Step 1, it segments all wiki editors using their weekly contribution profiles with the aim to identify the elite editors, who are the major contributors as well as the most active editors. In Step 2, the elite editors are clustered into different types. In Step 3, treating those elite users as the center nodes, editor subnetworks will be constructed and validated by the network layer data.

The dataset, which is currently available through Purdue's supercomputing environment, consists of several layers.

1. *Primary layer* The primary layer is the dataset created by Luca de Alfaro (one of the senior personnel in our team) and his colleagues at the University of California-Santa Cruz (Adler and de Alfaro 2007). It contains 250 million records representing all revisions made to Wikipedia articles from January 2001 through June 2010. The metadata for each revision include, among other pieces of information, the username and the unique identification number for the user making the revision; the unique identification number for the article being revised; the timestamp of the revision; and a measurement of the quantity of content changed through the revision, which de Alfaro et al. (2010) termed the edit distance, or "Delta." The text file containing this information comprised over 67 GB of hard drive storage space.
2. *Network layer* The network layer, which is also available as a 120 GB database, includes the implicit editorial interactions between over 21 million unique contributors (approximately 3 million logged-in users and 19 million IP addresses), who have contributed to 7 million Wikipedia articles between 2001 and 2010.
3. *An entropy layer* A weekly normalized value of interaction entropy has been calculated to measure the degree of collaborative evenness and diversity on Wikipedia.
4. *A user log layer* A user-by-week table was generated (21 million users $\times$ 495 weeks) was created to describe the individual weekly contributions of each individually identifiable user.

The primary layer has been used to generate the other three layers. First, social entropy values for collaborative interactions were calculated at a weekly level. The entropy measure for each week indicates the degree to which the contributions (as measured by delta values) are even or uneven. Initial findings suggest that following an initial period of relative instability, entropy increases over time, reaching a plateau by the fifth year of the project (2006).

Second, the user log layer took the form of a summary database for the weekly contributions of 22 million unique Wikipedia editors (3 million utilizing registered accounts and 19 million unique, anonymous IP addresses). These contributions were tracked over time in order to detect the emergence of functional leadership groups and the temporal stability ("stickiness") of such groups.

Third, the network layer represents critical semantic information for structuring the raw data into a social network. In building this layer, we started from the premise that on Wikipedia, collaboration occurs whenever individuals coedit the same article. The weights of coeditorial ties (network edges) represent the *probability* of one editor engaging the content left by a previous editor.

In order to establish this probability, we define edge weights using two factors: the quantity of information contributed by each editor, and the number of intervening revisions between those made by the two editors. After all, the more substantial the first editor's contribution is, the more visible that contribution will be to subsequent editors, and the more likely those later editors are to interact with the prior contribution. Likewise, the more substantial the contribution of the second editor, the more likely that editor is to have reviewed a larger portion of the article, and the more likely he or she therefore is to have seen and evaluated the first editor's contribution. However, as more editors continue to revise the article after the first revision, the contribution made by the first editor may be changed or removed entirely, reducing the likelihood that subsequent editors will have the opportunity to interact with the content of the first revision. As such, the more revisions made between those of any two given editors, the less likely the second editor is to be able to engage the content of the first.

This approach falls in line with Isard's (1954) gravity model of trade, which approximates the likelihood of two potential trade partners forming an exchange relationship based on the quantity of goods held by each party and the distance between them. Similarly, in developing our network layer, we evaluated the likelihood of two editors of the same Wikipedia article forming a collaborative relationship by considering the quantity of content contributed by each editor and the temporal distance between their revisions, as measured by the number of intervening revisions that might obscure the contribution of the first editor. Thus, while Isard's original model approximates the likelihood of connections between parties located in physical space, our modified version, the gravity model of online interaction targets the probability of connections between revisions, and the editors who developed them, over time (Britt 2011).

Further details on the approach used to define network edges are given below, followed by a short overview of our plans for using this data for research in the KredibleNet community.

Network Edge Definition The definition of edges represents the most critical step for constructing the network layer. In accordance with the gravity model of online interaction, a connection (e.g. edge) from a given editor A to editor B was formed when editor A contributed to a Wikipedia article, and editor B revised that article at a later point in time; in other words, editor B was working with the content that editor A developed.

The significance of each revision was quantified using de Alfaro's (Adler et al. 2008) "Delta" score, while the distance between a pair of revisions was assessed by using the timestamps to determine the order of revisions to a given article and taking the difference between revision numbers for the pair in question. For instance, the distance between the second and fifth revisions to a given article is $5 - 2 = 3$.

Using the gravity model of online interaction, then, the weight of the connection formed between editors A and B based on their two revisions is given by the product of the two revisions' Delta scores divided by the squared distance between them. *The gravity model of online interaction formula is:*

$$F_{ij} = \frac{M_i M_j}{D_{ij}},$$

where M_i and M_j represent the Delta scores of two revisions, and D_{ij} is the squared distance between those revisions.

Of course, on Wikipedia, a given editor may revise an article multiple times, creating multiple instances of one's own work over the development of an article. Each of these instances has a different temporal location in the article's growth, and thus there are different temporal distances between that contribution and others in the same article. Connections are formed between revisions, which we can then extrapolate to represent additional connections between the authors of those revisions—or, more appropriately, further growth in the connection between the two authors.

In addition, collaborators may not be limited to partnerships that stem from revising a single article repeatedly. Two individuals might connect by mutually collaborating on a number of articles across the Wikipedia community, which would strengthen the bond between those fellow editors just as recursive interactions on a single article would.

Thus, while each article forms its own miniature community, the full Wikipedia network consists of all individuals from all of these subcommunities. Any redundant edges moving to and from the same two actors are summed in creating this master network, reflecting relationship growth over the course of collaborative activities throughout the community. As such, the more that any two editors built upon their relationship by mutually editing the same set of articles repeatedly, the more that their collaborative connection grew.

From network layers to hypotheses about functional roles and reputations The final network graph constructed using this methodology and divided into weekly summary sequences will be available to the research community gathered by KredibleNet as a "sandbox" for testing statistical algorithms, socio-technical theories or hypotheses. Some of the hypotheses will focus on the emergence of functional roles, division of labor, or reputation derived from functional roles in the Wikipedia knowledge creation marketplace.

The network graph resides in the computing cloud made available by the Rosen Center for Advanced Computing and the cyberinfrastructure environments of the Cyber Center and Discovery Park at Purdue University. It will be accessible to the project participants and to the larger academic and professional communities through a set of Web services and query interfaces facilitated by CRIS and NodeXL.

The services we provide will allow data query and retrieval, statistical analysis for hypothesis testing and visualization of various aspects of the Wikipedia network graph. For example, researchers who would like to focus on the coeditorial networks that have contributed to articles related to climate change could query the dataset for all articles related to a particular concept, such as "global temperature record" or "paleoclimatology," or a whole category like "climate history," and obtain the full list of articles within their query. Further, researchers will be able to generate the networks of interaction for a specific article or for the entire query. Such graphs can be delivered as graphML or UCINET files, which can then be further imported into other network analysis tools, such as the previously mentioned NodeXL framework.

These examples, however, represent mere simple exploratory procedures. Our primary goal is to support more sophisticated analytic strategies, especially those focusing on revealing functional roles and reputations, collaborative structures, or functional leadership dynamics.

Research Task—Exploring Functional Role and Reputation Measurements A significant amount of research that will be conducted on the Wikipedia coeditorial network dataset will consist of detecting network signatures for functional roles and for deriving the implicit "reputation" (social capital) held within those roles.

This type of research demands a sophisticated approach to entity discovery in communication networks. Such entities (subnetworks) will be categorized according to their contribution to the structuration of the larger coeditorial network and their potential attribution to the functional roles that the nodes found in their central locations may perform. A later section (see Sect. 4.2) describes the various strategies that can be used for entity discovery. In the rest of this section, we focus on what we will gain by identifying specific entities in a social media network, in this case, Wikipedia.

Identifying subnetworks in a broader, undifferentiated network of collaboration will be equivalent to identifying the constitutive building blocks of a building (Matei et al. 2010a). A large network of collaboration is made of many overlapping local networks in which people interact with the same content and connect with one another through that content. These small groups are further divided into the functional roles that their members assume: information aggregator, coordinator, proof-reader and so on, each discernible through those members' patterns of interaction with each other and with the content (Welser et al. 2007).

Analytically, when the larger network is decomposed into simple subnetworks, we indirectly reveal how local collaborative structures exist. Such microcollaborative clusters are seen as footprints of routinized interactions, which we will categorize into a variety of functional roles. Revealing the nature of the functional roles can be done by comparing the subnetworks they anchor and dividing them into a taxonomy. This would require a semi-inductive process. Practically speaking, the taxonomy will take into account the nature of the connectivity of the individuals with highest level of betweenness centrality in each of the subnetworks. Node multiplexity, connectivity intensity, breadth and diversity, as well as overall contribution to group contributions will be used for defining recognizable functional roles of editorial, administrative, or "political" consequence and for classifying each subnetwork. Associations between roles and interaction structures will be identified by comparing individual and network-level attributes using appropriate statistical modeling.

Furthermore, roles isolated into subnetworks of interaction could be further contextualized in the coeditorial network of a specific article or topic by the effect they have on group productivity, coordination, and norm enforcement. This will become, in effect, the root procedure for assigning "reputation" to any individual node. This measure of reputation will be refined by observing affinities between the nodes and the manner in which the local regions in which nodes interact pass on key information to other regions, allowing them to self organize. Reputation will therefore incorporate role characteristics and the impact of these roles on the local and global networks of interactions.

Once the local networks have been "sequenced" in this manner and divided into building blocks, these subnetworks (and the roles they define) may themselves be treated as superordinate "nodes" within the complete network. We approach networks and the functional roles associated with them within the framework of a lattice structure, in which simpler structures contribute to creating increasingly complex structures. Similarly, analyzing combinatorial affinity with entity discovery statistical methodologies allows us to assess the likelihood of simple graphs associating with one another in a specific order and following a statistically meaningful pattern to form more complex structures. This operation may be repeated until all the nodes, both local and superordinate, are joined in a meaningful structure.

To summarize, the process of detecting network graph building blocks, and ultimately the functional roles and reputations most responsible for the structuration of a social media space, involves breaking down the undifferentiated networks of collaboration social media into their most basic units (local graphs/communities associated with functional roles). The next step is to define the possible range of such subnetworks and associating the most central nodes in each one with the "functional role" that it performs. Then, we measure the impact of these central nodes on the connectivity of rest of the network and on the productivity of the nearby nodes. These two types of impacts effectively represent the "reputation" component of the functional roles, or their "social capital" potential. We thus aim to derive a definition of reputation whose measure is restricted to a function of a node's specific role and impact in the network.

Research Task—Quality Impact of Functional Roles Our efforts to identify specific functional roles and to measure their impact on knowledge production processes would not be complete without an assessment of the connection between functional roles and the quality of the articles produced by editors with specific reputations and impacts on the network. In order to address this key issue, we make use of the Article Feedback Tool, which Wikipedia launched in 2010. This tool allows Wikipedia users reading a particular article to rate it on several dimensions:

- *Trustworthy*: "Do you feel this page has sufficient citations and that those citations come from trustworthy sources?"
- *Objective*: "Do you feel that this page shows a fair representation of all perspectives on the issue?"
- *Complete*: "Do you feel that this page covers all the essential topic areas that it should?"
- *Well-written*: "Do you feel that this page is well-organized and well-written?"

This tool has been very popular. In 2011, there were about 30,000 ratings every day. About 97 % of users submitting ratings were not logged-in editors, indicating significant uptake by registered and anonymous readers alike.

We and some of our collaborators have conducted several preliminary data analyses to study the correlation between these ratings and attributes such as page length, the number of revisions, and the number of registered and anonymous users who had contributed to the article at the time of a given rating. These analyses have already produced some interesting results; for instance, we have found that the number of users involved in editing an article is an important positive indicator for the "Complete" dimension, but it appears to be unimportant in whether an article is deemed "Trustworthy."

In the context of the KredibleNet project, we plan to utilize the Wikipedia coeditorial network data to predict quality and trustworthiness based on the functional roles and reputations of each article's most influential editors. In the process, we intend to define a specific quotient of "contribution impact" for each contributor, with the measure sensitive to the moment in time and the place in the sequence of edits of a specific contribution. This work will build on and extend work done by one of our external collaborators, Luo Si (Cetintas et al. 2011; Zhang et al. 2012).

Research Task—Assessing the Structure of Collaboration and the Stability of Functional Leadership Groups The uneven distribution of rewards and benefits within working groups is a constant of human affairs and a topic of much debate and political conflict. Pareto's "80/20 rule," which says that 80 % of the output or benefit is produced or enjoyed by 20 % of the members of any given group, has increasingly come under debate, and the economic, social, organizational, and especially ethical justification for systems of contributions and benefits that follow uneven distribution curves is questioned every day.

The current heated debates that surround income distributions in the US and Western Europe about the proper amount of taxation to mitigate social and economic inequalities invoke the rule often and in a contradictory manner, even when the numbers are not as dramatic as those postulated by Pareto's rule. The top 20 % American earners, for instance, each of whom makes over $ 101,582 per year, reap 51 % of the combined national personal income—a majority of the American wealth, to be sure, but not nearly as extreme a fiscal division as the 80/20 rule would suggest (US Census Bureau 2012).

Team efforts are often equally skewed in corporate contexts, as are material rewards for those efforts. CEO, for instance, are often accused of reaping a disproportionate share of compensation, with the average CEO earning 380 times the annual salary as the average American worker (Liberto 2012). At the microlevel, the distribution of effort in teams generally follows the same rule, with a small proportion of team members responsible for most team productivity.

Bales (1953) observed that across numerous experiments of small groups, the distribution of effort was consistently unequal across group members. Kumar et al. (2010), in turn, showed the prevalence of this inequality apart from work communities, even in online social networks, while Ortega et al. (2008) demonstrated its applicability to the online community of Wikipedia editors. Bruno (2010) extended this notion, arguing that the observed inequities illustrate not merely the natural variation of individual effort in the statistical sense, but much more significantly, the adoption of different functional roles by community members.

Such observations and the public debates that they spur are very important because they speak about the nature of human collaboration and about the incentive systems that can be used to stimulate it. Furthermore, the ethical dimension of such distributions is just as important to consider, particularly given their relationship with our core concern for equality, the keystone of democratic societies.

With the advent of the Internet, and especially the widespread adoption of online communication and interaction, many scholars and practitioners invested a great deal of hope in the capacity of such new media technologies to level the societal playing field, so to speak. These individuals hoped for the Internet to create more opportunities for individuals to contribute more evenly and to be rewarded commensurately in a variety of contexts, from business to media to educational endeavors. One famous example comes from Raymond (1999), who argued that freely allowing more programmers to enter the Linux development community greatly enhanced the workflow and the final product. After all, he argued, "Given enough eyeballs, all bugs are shallow."

Yet, claims that the Internet will usher in an era of egalitarianism have been dampened by empirical observations from Ortega et al. (2008), and Bruno (2010), among others (for instance, Correia et al. 2006; Huberman 2001; Madey et al. 2002; Mockus et al. 2002), that seem to suggest an alternative scenario of inequality (Matei and Bruno 2012). Shirky (2008) offered an especially prominent assessment of this imbalance, observing that the time users spend taking part in Internet communities tends to follow the power law distribution previously observed in such domains as income and organizational effort.

This imbalance, however, results in both dramatic benefits and serious consequences. For instance, Kuk (2006) found that many open source software developers strategically form small clusters with others who are especially resourceful, and that taking advantage of this opportunity, which is afforded by the openness of the communities in question, ultimately enhances the collaborative effort, that is, as long as this concentration is not taken to an extreme. Moore and Clayton (2008), in contrast, examined a phishing (malicious spam attack) patterns and showed that the power law distribution of contributions among its users made it slow to pinpoint problems, susceptible to manipulation, more prone to inaccuracies, and generally less complete than proprietary sites serving the same purpose.

In short, Internet production projects, especially voluntary collaborative efforts which are open to newcomers, display uneven distributions of contributions and rewards, which are on par with those of other domains, especially offline interactions, some of them explored as early as the 1950s (Bales 1953). The same kinds of disparities among users persist on Wikipedia, which is perhaps the most famous example of a voluntary collaborative community. According to our own calculations, supported by previous independent work (Ortega et al. 2008), if Wikipedia was a nation and words were its wealth, the top 1 % contributors would own more than 90 % of the "riches." In other words, a small minority of editors controls the vast majority of information on perhaps the world's single most widely-used knowledge resource. And, as we will show, even if the method for calculating member contributions varies, the observed disparity remains quite consistent.

The prevalence of uneven distributions of contributions across online projects is widely known, and past work has proposed social entropy as a method for measuring it (Matei et al. 2010b), but the ultimate meaning of such distributions has not been sufficiently explored. A core question that remains unanswered is whether such skewed distributions are merely a result of random processes, or whether there is something special about top contributors such that the digital world grows to be dominated by an increasingly stable group of elites.

In the context of Wikipedia, which is the empirical ground of the present study, if we were to examine the top 1 % contributors over time, what should we expect? That the composition of this group changes all the time, as some users join and others leave the project, and as some find time to contribute and others do not? Or, alternatively, would we observe contributors who become one of the top 1 % tend to remain top contributors for a long time, joining an elite group of similar long-term leaders? In other words, does the elite group tend to become durable over time, preserving its membership even as wave after wave of new contributors join the project every day?

This question has not only descriptive, but also inferential consequences. Let us assume, for the moment, that the contribution and collaborative processes on Wikipedia, or those of any other similar project, are dominated by a small group of stable, long-term elites at the top of the project, and that this group is consequently responsible for a vast share of the content contributed. In this case, we may ask such questions as:

- Will the project exhibit a system-level structuration process through which routine activities and functional roles stabilize and become accepted standards over time?
- What are the distinct phases through which this structuration progresses?
- How does the emergence of a stable elite influence other aspects of the structuration process, including the possible solidification of other functional roles in the community, and how does this promote or detract from project productivity?
- Moreover, is the durability of the contributing elite group an indicator that social media projects such as Wikipedia gradually develop "functional roles," which can be defined as a measurable pattern of social collaboration?

- Finally, do the top users subjectively view themselves as being part of a stable leadership team within the project as a whole, and do they act accordingly by forging alliances with other top members, eventually taking on more formal functions in the project and attaining increasingly consequential leadership positions?

In short, one of our core research tasks is to ascertain how voluntary knowledge production social media projects such as Wikipedia evolve over time, especially including how they become structured. We aim to measure these structures using social entropy, to assess the degree to which membership becomes organized into functional roles and elites, to observe how the elites function as a group, and to explore the impact of these processes in order to better grasp how social media change the way human organizations work.

In the subsidiary, we would like to find a number of synthetic indicators that could be used to develop a continuum for social media projects. At one end of this spectrum, we would have a situation where the leadership is forever changing, with each individual doing his or her part only to abandon the leadership post after a very short time, following the "wisdom of the crowds" paradigm in which no particular individuals dominate the collaborative effort. At the other end, we would observe collaborative groups in which the top contributors are almost all long-term members who tend to know and promote one other and who possess a nuanced, subjective understanding of their leadership positions gleaned from their extensive time working within the project. Such indicators would help us to enhance our understanding of leadership roles in the social media era and their potential impact on human organizational behavior in general.

The strategy for determining the top collaborative groups (in terms of output) and their likelihood to be sticky, over time, will rely on a variety of methodological approaches. The general approach starts by differentiating between edits generated by users that have created an account and those created by anonymous users, which can only be identified by IP addresses. The characteristics of the two groups of editors are compared, including output distributions, percentiles, and frequencies. The comparison results show that, although the groups are different from each other, the top IP editors cannot be ignored. According to preliminary investigations, 65 % of the editorial effort is generated by a mere 80,000 editors of the total 22 million. Of these 80,000 top editors, 50 % (40,000) are anonymous and can only be identified by IP addresses. Therefore, both anonymous and logged in editors will be included in the analysis.

All editors are rank-ordered based on their overall contributions, and two thresholds for the ranks are identified, which can be used as the cut-off points for defining or identifying the elite editors. The first threshold is T_1, including the top 80,000 editors. The second threshold is T_2. There are 1 million editors with contribution above T_2, and collectively they account for 90 % of the total contribution. In addition to the total contribution, the number of active weeks of an editor further reflects the editor's duration, editing frequency, and reputation in wiki. For the top 1 million editors, their average weekly contribution and the total number of active weeks are calculated, and a two-dimensional density surface is fitted to the resulting data.

The fitted density surface is then used to identify editors with either large average weekly contributions or high editing frequencies. The identified editors are considered the elite editors.

Hidden Markov models are used to model the editing patterns of the elite editors, which are clustered into groups. The fitted models and identified clusters can be used to study the stickiness of the elite editor groups and further be used to construct the editorial subnetworks required by the entity discovery procedure described above.

Data Analytics

In order to address these research questions, we will make use of machine learning techniques (including statistical machine learning) and sophisticated data analysis tools to support new analytic and statistical approaches for analyzing the KredibleNet data. We first propose an initial set of data analytic approaches, and we expect to develop a fuller, broader set of requirements and a preliminary design for the analytic techniques and tools during the workshops. Below, we discuss the three critical analytic challenges.

Entity and Network Discovery Entity recognition and network discovery (e.g., association or relationship discovery) are two important research problems to be addressed before we can measure the credibility or reputation of entities.

One key task of entity recognition is to identify different types of entities such as people, their respective structures of interaction in a given implicit network of collaboration, and even entire organizations and locations. In our case, we will focus on decomposing the broad Wikipedia interaction dataset into smaller subnetworks, each of which will then be associated with particular functional roles and reputations.

There are two main approaches to such an entity recognition task: the rule-based approach (such as the one discussed above) and the learning-based approach. Researchers enacting the rule-based approach design different types of rules to identify entities of appropriate types. In contrast, the learning-based approach begins with the development of different types of models based on some training data; these models may then be used to identify entities on test data.

Both approaches have been successful in past research endeavors. One intuition is to combine both types of methods (rule-based and learning-based) together to improve the accuracy of entity recognition. Prior research by Si, Kanungo, and Huang (2005) proposed a meta entity recognition method to combine the results from multiple entity recognition system in order to obtain a superior final product. The authors proposed three meta recognition algorithms, and their initial empirical confirmation demonstrated that these methods substantially improved the entity recognition accuracy of a biomedical application over the use of individual entity recognition systems. The best results were obtained with a conditional random field method that took advantage of structural information for recognition.

One of the most important tasks within this process is to identify the association between different types of entities. In this case, we first aim to address networks based on coeditorial activity, and then to explore clusters of subnetworks associated with functional roles. This is similar to previous research (Balog et al. 2012; Fang et al. 2010, 2011) on faculty home pages, a conceptual entity which is clearly associated with a particular individual (the faculty member in question) even when the page's author is not explicitly stated.

To explore these relationships, Fang et al. (2010) proposed a joint prediction model, which started by defining a dependence graph for web pages. Next, a conditional undirected graphical model was employed to make joint predictions about page references to individuals/owners by capturing the dependence of the decisions on all the candidate pages.

Three cases of dependencies among units of analysis (in the cited work, Web pages) are typically considered for constructing the graphical model. The statistical model utilizes a discriminative approach so that any informative features can be used conveniently. Learning and inference processes for the joint prediction model are relatively efficient because the dependence graphs resulting from the three cases of dependencies are not densely connected. An extensive set of experiments were conducted on two test-beds to show the effectiveness of the proposed discriminative graphical model, and the cited work built separate models of association detection (i.e., faculty homepage detection) for different domains (e.g., different universities).

The KredibleNet team will use this approach for building models of association detection that use the coeditorial connections on Wikipedia as a graphical model. The overall goal of this effort is to identify dependencies between nodes (editors) defined by mutual coeditorial activity, and to use the observed dependencies to predict the roles of those editors.

Social Entropy Measurement The structuration of a collaborative space is a function of the number of participants (m) and the shares of participants (S_i). More participants and more equal shares imply a higher level of uncertainty about activity within the community, which can be translated conceptually as a higher level of "entropy."

How can we translate this into a synthetic indicator? We can do it, as Shannon (1948) suggested, by measuring the relative degree of disorganization found in any system. We can envision disorganization as the random mixing of various elements, whose relative presence should thus be equal. In this situation, we can also say that the diversity of the system is at a maximum, since all elements are equally (randomly) present.

Shannon's entropy index takes a value of zero when there is an absolute order in the system (one element is prevalent at the expense of all others) and a maximum value (which varies from system to system) when there is perfect disorder and diversity (all elements are equally present). Entropy is therefore a synthetic measure that tells us, at a glance, the extent to which the different components of a social or communicative space are well-represented.

Mathematically, the entropy of a random variable X (which in this case represents the contribution level) is based on the proportion of all contributions made by each individual contributor. For all contributors the collaborative effort, the proportion generated by each individual person, in other words, the percentage of the workload for which each individual was responsible, has the probability mass function $p(x)$. Using this, we can define $H(X)$, the social entropy of the entire system, as follows:

$$H(X) = -\sum_{i=1}^{m} p(x_i)\log_2 p(x_i).$$

Entropy therefore varies from zero to $\log_2 m$.

How do we apply this measure to online collaboration environments? Consider an online communication space in which there is a uniform distribution of contributions by four members, $\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$. The entropy of this communication space is

$$H(X) = -\sum_{i=1}^{4} S_i \log_2 S_i = -\sum_{i=1}^{4} \frac{1}{4}\log_2\frac{1}{4} = \log_2 4 = 2.$$

Now, consider another communication space with four members. Assume that the shares of contribution by these members are unequally distributed, $\left(\frac{1}{2}, \frac{3}{10}, \frac{1}{10}, \frac{1}{10}\right)$. The entropy of this communication space is

$$H(X) = -\sum_{i=1}^{4} S_i \log_2 S_i = -\frac{1}{2}\log_2\frac{1}{2} - \frac{3}{10}\log_2\frac{3}{10} - 2^*\frac{1}{10}\log_2\frac{1}{10} = 1.69.$$

The entropy of the former communication space with a uniform distribution of contributions is higher than the latter one with unequally distributed contributions.

Although entropy is an elegant modality to measure diversity in a system, there are some potential limitations. Entropy reflects not just one, but two system dimensions: richness and evenness. When we collapse them into one index score, we necessarily lose some information (Balch 2000).

Moreover, the two dimensions can contribute in different ways to entropy scores that are very similar, which can lead to all sorts of confusion. For example, two very different online communication groups in terms of composition and contributions can have entropy characteristics that seem to be very similar (Balch 2000).

Consider a communication space (C_1) with four contributions made by two participants. The shares of the two participants are equal, $\left(\frac{1}{2}, \frac{1}{2}\right)$. The second communication space (C_2) has 64 contributions unequally distributed among seven participants, $\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{64}, \frac{1}{64}, \frac{1}{64}\right)$. However, the calculation of entropy provides a counterintuitive result: while the entropy of the first communication space (C_1) is 1, the entropy of the second communication space (C_2) is 2, despite the obvious fact that in (C_2) the contributions are less evenly distributed than in (C_1). This is because (C_2) has more participants. The fact that their contributions are unequally distributed is hidden.

This problem can be solved by normalizing the entropy values. This enables us to compare the evenness of two communication spaces, including over time, by controlling for the number of elements that compose each of them. Normalization can be obtained by dividing the raw entropy score by its maximum $\log_2 m$, which limits its range from 0 to 1:

$$H_o = \frac{H}{H_{\max}}, 0 \leq H_o \leq 1, \text{where } H_{\max} = \log_2 m.$$

Normalized entropy is particularly useful for handling the "lurker" problem in studying diversity in online environments. Lurkers are users who do not make any tangible contributions to an environment; he or she is just an observer. Such lurkers can make an environment potentially richer, but they can also impact diversity. How can we capture both of these aspects of the lurker behavior?

Suppose the following two communication spaces. In both, the contributing members make equal contributions:

$$C_1 = \{\Delta, \Omega\}, P_1 = \{Tom, Jane\}, \text{ and the share distribution is } \left(\frac{1}{2}, \frac{1}{2}\right).$$

$$C_2 = \{\Delta, \Omega\}, P_2 = \{Tom, Jane, Sara\}, \text{ and the share distribution is } \left(\frac{1}{2}, \frac{1}{2}, 0\right).$$

In the second communication space, however, is lurker Sara, who did not contribute to the interaction. Despite this important difference, the nonnormalized entropy of the two communication environments is the same, 1.

$$H\ (X_{c1}) = -\sum_{i=1}^{2} S_i\ log_2 S_i = -\frac{1}{2} log_2 \frac{1}{2} - \frac{1}{2} log_2 \frac{1}{2} = 1$$

$$H(X_{c2}) = -\sum_{i=1}^{3} S_i \log_2 S_i = -\frac{1}{2}\ \log_2\ \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} - 0\ \log_2\ 0 = 1.$$

Although the two raw entropy scores are identical, normalizing these raw entropy values highlights the presence of the lurker in one of the spaces. For example, the max log (m) entropy value for C_1 is

$$H_{\max}(X_{c_1}) = \log_2\ 2 = 1,$$

which yields the normalized entropy value

$$H_o(X_{c_1}) = \frac{H}{H_{\max}} = \frac{1}{1} = 1.$$

For C_2, the maximum entropy value will be

$$H_{\max}(X_{c_2}) = \log_2\ 3 = 1.58$$

and its normalized social entropy is therefore $H_o(X_{c_2}) = \frac{H}{H_{\max}} = \frac{1}{1.58} \cong 0.63$.

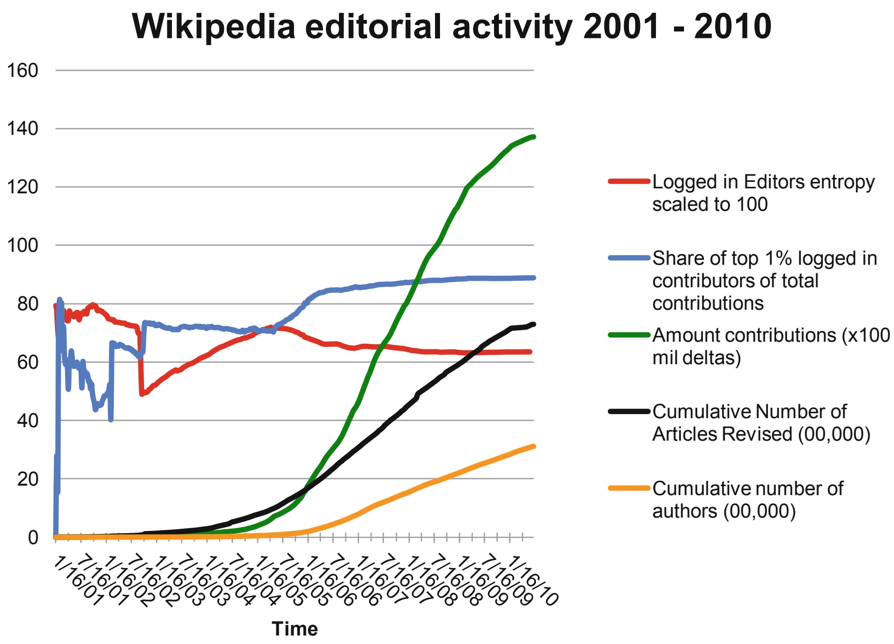


Fig. 1.2 The editorial activity of logged in users mapped on several dimensions: entropy, share of top editors of total editorial activity, total amount of editorial activity, number of authors, and number of articles.

Comparing the normalized entropy of two communication spaces shows us that the second communication space is more diverse than the first one, because the second formula takes into account the presence of the lurker.

We have employed the normalized entropy measure to plot the evolution of contributions on Wikipedia over time. A project-wide weekly entropy value was calculated for 495 weeks (2001–2010), which indicates the degree to which the project has become more or less structured (entropic) on a weekly basis. The red line in Fig. 1.2 presents weekly changes in normalized entropy level on a 0–100 scale over a period of 9 years, against a number of other metrics: share of contributions accounted by the top 1 % contributors, total amount of contributions, number of articles, and number of logged in authors. The Y-axis should be read according to the metric of each measure. For entropy, which is a normalized variable, values span between 0 (no entropy) to 100 (maximum entropy). Share of contributions spans 0–100 %. Amount of contributions, number of articles and authors are unbounded on the upper end and their values on the Y-axis need to be multiplied by the numbers in parenthesis. Entropy and share of top 1 % contributors reflect data for the 3 million users that logged in to make changes, and whose effort account for about 2/3rd of the total editorial work done on Wikipedia over 495 weeks.

The most important preliminary finding is that after a period of wild fluctuations, partly induced by changes in MediaWiki, the underlying platform that supports the online collaboration on Wikipedia, entropy on Wikipedia has reached a stable level around the middle of 2006. Holding steady at values a little over 60, it is a rather low level, as well. To better understand it, it needs to be compared with the share of contributions accounted by the top 1 % logged in contributors (blue line). In July 2006, these editors accounted for no less than 84 % of all editorial effort. By the end of the study period, 4 years later, in July 2010, they accounted for 89 % of editorial effort. As the level of entropy is directly correlated with variation in distribution of effort, the steady entropy level of over 60 (on a 0–100 scale) indicates that Wikipedia had become very well structured and remained in a "steady state" over a prolonged period of time. This is all the more remarkable since during the same period of time (July 2006–July 2010), the amount of contributions has increased almost seven fold, the number of articles four fold, and the number of logged in authors over 13 fold! The fact that entropy decreases in the last 4 years, indicating a larger amount of work being performed by a smaller number of people, is even more remarkable given the explosive increase in number of authors.

Statistical Modeling and Analysis Starting from this descriptive exploration of the data, our further investigation of the Wikipedia coeditorial network will embrace two main statistical strategies: exploratory analysis and statistical model-based inference. We intend to use exploratory analyses to summarize and report features of the network using summary statistics, whereas statistical model-based inferences will allow us to characterize the network using proper statistical models.

Among the various types of models proposed for social networks, two stand out as the most promising. The first type consists of exponential random graph models (Snijders et al. 2006), in which ties among nodes are assumed to be random variables and further dependencies among these random variables are imposed. The second type consists of latent social space models (Handcock et al. 2007). These models postulate the existence of a latent space and further assume that ties, as random variables, are determined by the relative positions of actors within the latent social space. Both models can be used to statistically infer fundamental rules and patterns within a social network based on observed data.

In this project, concepts, models, and methods developed in social network research will be employed and further extended to study online networks for the purpose of developing proper online reputation measurement. Because online reputation is a relatively new concept that has not been well-studied in the literature, new conceptual models and methods need to be developed. In particular, online social networks are usually much larger, more complex, and less structured than face-to-face or "real-life" networks. Data collected from online networks are massive and to use a term from information theory, "noisy."

In the literature on statistical analysis of social networks, both the maximum likelihood estimation (MLE) method and the Bayesian method have been suggested as reasonable approaches to fitting statistical models, such as the two types of models previously discussed, to observed data. The Bayesian method using Markov Chain Monte Carlo (MCMC) simulation is generally considered to be more favorable than the MLE method, so we will use MCMC simulation as the primary analytic tool for the current project.

For Bayesian analysis of our online network data, the PIs plan to develop (1) efficient MCMC sampling algorithms for the statistical models built during the research and (2) MCMC methods in parallel and distributed computing environments.

Over the past two decades, much work has been made to create efficient MCMC sampling methods (see, for example, Liang et al. 2010) for Bayesian inference as well as closely related efficient EM-type algorithms for maximum likelihood-based inference (see, for instance, He and Liu 2011; Lewandowski et al. 2010). We believe that, like many proposed efficient iterative algorithms, new efficient algorithms can be developed by making use of statistical thinking that takes the advantage of the specific features in the statistical models for online networks. The development of efficient MCMC iterative algorithms is expected to be an effective approach to statistically analyze online network data using statistical models that deal with complex structures.

The second objective in developing computational tools for statistical inference focuses on the massive data problem. The project will develop MCMC methods in parallel and distributed computing environments. Such methods can dramatically reduce the time and complexity of the approaches needed to sieve through, interpret, and test hypotheses on massive datasets. Chuanhai Liu of Purdue University recently developed an experimental R package called DISC which may be used for experimenting in a Distributed Iterative Statistical Computing environment. This package was written for a graduate course on Large Scale Data Analysis which Liu offered in the fall of 2011. The DISC package can run on any number of network-connected computers, including computer clusters.

DISC is an R package, which abstracts the idea of iterative computation in distributed computing environments, such that the popular strategies for handling distributed data, such as MapReduce (Dean and Ghemawat 2008), become a special case of the DISC framework. The system provides a suite of default configurations and a user-friendly implementation that makes it simple for the user to implement iterative algorithms in a distributed computing environment.

In essence, DISC consists of three types of R sessions, which are structured using the "master–slave" interaction model:

1. User sessions, which provides an interface between the user and DISC system
2. System sessions, such as the MasterProxy session running on the master machine as a daemon process, which helps establishing new connections between the master and "slave," end-R processes, and FileServers on machines with DISC "slaves" that play the role of transferring R objects and files, and
3. Parallel R sessions invoked by the "slaves" to perform computing tasks specified by the user.

DISC provides a simple way to manage distribution of data and parallel processing across multiple sessions of R, either running on multiple cores of the same machine, or in a cluster of machines. DISC was designed to be extensible to make it easy to use experimental algorithms, and allows user to develop and add modules, called DISC applications. Two useful built-in applications included in the current version are DS and MR. The DS application creates and distributes data subsets. Together with the built-in FileServers, the DS application offers data manipulation in a distributed computing environment, especially if one wants to replicate the Hadoop paradigm. The MR application implements the MapReduce algorithm (Dean and Ghemawat 2008).

Glen DePalma and Sanvesh Srivastava, two students in Liu's class implemented EM (Expectation Maximiziation) algorithms on DISC. EM is typically used to compute maximum likelihood estimates using incomplete data (Dempster et al. 1977). Their work and upcoming presentation at the American Statistical Association's 2012 Joint Statistical Meeting (see DePalma et al. 2012 for details) demonstrated that iterative simulation in DISC is indeed promising.

The KredibleNet Cyberinfrastructure

To realize the goals of KredibleNet, the team will deploy a transformative cyberinfrastructure—*CRIS*—to unify a vibrant network of data and scholars. The goal of the CRIS component is not only to test some hypotheses of its own, or to deploy the tools and approaches described in the preceding sections, but to share the findings, the documentation of our approaches and analyses, and the datasets themselves to other researchers, in such a way as to provide springboards for further research. CRIS's primary tenets are to provide an easy to use, trustworthy, cost-effective, and scalable cyberinfrastructure for scientists lacking expertise in computational tools and system administration. CRIS will deploy as a broadly accessible cloud-based community platform, enabling the easy sharing of data with transparent attribution and embedded computational and analytical tools, and facilitating secure community collaboration and research extensions.

The CRIS philosophy is to not reinvent data networks, but to allow existing distributed data and computational tools to be "wrapped" into the system, thus providing broader and more uniform access. In this manner, CRIS brings together the pieces that exist today into an infrastructure to allow scientists to focus their efforts on understanding sustainability efforts in new and innovative ways. This is accomplished by providing: (1) a suite of tools to automatically capture, transform, and analyze data; (2) embedded attribution for all levels of research activity (data, workflows, revisions, etc.); (3) integrated vocabularies for data definition; (4) automatic data quality monitoring; (5) interactive research workflows; (6) easy integration of existing data and computational tools; (7) easy, yet trusted access to workspace information; and (8) long-term storage and access to organized and managed data; all leading to verifiable collaborative research.

Through the use of the CRIS system in KredibleNet, the community will be able to improve the quality of what they can already produce, accomplish more through improved efficiency, enhance their scientific rigor, avoid duplicative efforts, and advance understanding through more complete access to all research components. In particular and as a starting point, CRIS will integrate tools and data made available by the NodeXL project. For example, the NodeXL Graph Gallery Website (http://nodexlgraphgallery.org) makes tools for improving the variety and quality of data available to the social network analysis community.

Conclusions

KredibleNet is an ambitious project, whose aim is to map the relatively uncharted territory of social media roles, reputation, and trust. This goal is to be fulfilled by a collective effort of an academic and practitioner community, which we hope will grow over time. The multifaceted activities of the project—workshops, research tool development, content-specific research, publishing—are meant to converge in a new research agenda, which proposes that social media should be seen as an important "knowledge market," shaped by individual efforts that are generally voluntary, yet relying on a hierarchy of functional, achieved roles that accrue reputations through their linkages to social networks of collaboration. Describing these networks through new statistical methods, decomposing them into simpler subnetworks that can be seen as "social signatures" for specific roles, measuring the degree to which these networks become more complex (less entropic), and the role played by the nodes (roles) that dominate them are at the heart of this enterprise. Furthermore, determining the level of social credibility and trust of such roles and the role of social networks in buttressing credibility and trust is also central to our project. New statistical approaches and theoretical perspectives will be proposed in the process and an online tool suite, which takes advantage of a large, intact social media corpus (Wikipedia) will bring these insights to life. The tool will allow data mining activities to better understand how 21 million individual contributors created 7 million articles, through 250 million edits, over a period of 9 years. Questions regarding the social networks of collaboration, the contours of collaborative interaction over time, the emergence or roles or simpler, descriptive statistics will be possible to be generated by any researcher interested in how knowledge is generated in an informal collaborative information market like Wikipedia. Future directions for taking advantage of the tools and ideas proposed by our project are described in a separate chapter of this volume (see TITLE and pages). We hope, however, that the ideas presented in this chapter will grow over time and will take more tangible shape as the corpus of the literature inspired by our project will emerge and grow, changing the way we think about social interaction in knowledge markets.

Acknowledgements The work reported in this paper has been partially funded by the National Science Foundation (NSF) under award no. 1244708. The work of Sorin Adam Matei has also been supported by a Purdue Discovery Park Fellowship.

References

Adler, B. T., de Alfaro, L. (2007). A content-driven reputation system for the Wikipedia. In *Proceedings of the 16th international conference on World Wide Web—WWW '07* (pp. 261–270). New York: ACM.

Adler, B. T., de Alfaro, L., Raman, P. (2008, May). Measuring author contribution to the Wikipedia. Paper presented at WikiSym 2008, Porto, Portugal.

Arazy, O., Stroulia, E., Ruecker, S., Arias, C., Fiorentino, C., Ganev, V., Yau, T. (2010). Recognizing contributions in wikis: Authorship categories, algorithms, and visualizations. *Journal of the American Society for Information Science and Technology, 61*(6), 1166–1179.

Balch, T. (2000). Hierarchic social entropy: An information theoretic measure of robot group diversity. *Autonomous Robots, 8*, 209–237.

Bales, R. F. (1953). The equilibrium problem in small groups. In T. Parsons, R. F. Bales, E. A. Shils (Eds.), *Working papers in the theory of action* (pp. 111–161). Glencoe: Free Press.

Balog, K., Fang, Y., de Rijke, M., Serdyukov, P., Si, L. (2012). Expertise retrieval. *Foundations and Trends in Information Retrieval, 6,* 127–256.

Britt, B. C. (2011). System-level motivating factors for collaboration on Wikipedia: A longitudinal network analysis (Master's thesis). Retrieved from ProQuest Dissertations and Theses Database. (AAI 1501175). Purdue University, West Lafayette, In.

Britt, B. C. (2013). Evolution and revolution of organizational configurations on Wikipedia: A longitudinal network analysis (Doctoral dissertation).

Britt, B. C., Matei, S. A. (forthcoming). Leading from behind in online knowledge construction communities: A longitudinal network analysis of wiki collaboration. Unpublished manuscript.

Bruno, R. J. (2010). A democracy of unequals: Social differentiation, participation inequality and the collaborative ideal online. PhD Dissertation. Purdue University, West Lafayette, In. http://matei.org/ithink/wp-content/uploads/2010/07/ robertbrunodemocracyofunequals.pdf. Accessed 3 Jan 2014.

Cetintas, S., Rogati, M., Si, L., Fang, Y. (2011). Identifying similar people in professional social networks with discriminative probabilistic models. Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval (pp. 1209–1210). http://dl.acm.org/citation.cfm?id=2010123. Accessed 3 Jan 2014.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology, 94*(Supplement), S95–S120.

Correia, L., Castro Caldas, J., Teles, N. (2006, July). Freedom in knowledge creation. Paper presented at Workshop sobre Ciê ncias da Complexidade: Mathematics and Innovation, ciclo de Encontros da Arrábida, da Fundação Oriente, Arrábida, Portugal.

de Alfaro, L., Adler, B. T., Pye, I. (2010). Computing wiki statistics (WikiTrust). http://wikitrust.soe.ucsc.edu/computing-wiki-statistics. Accessed 10 Sept 2012.

de Alfaro, L., Kulshreshtha, A., Pye, I., Adler, B. T. (2011). Reputation systems for open collaboration. *Communications of the ACM, 54*(8), 81–87.

Dean, J., Ghemawat, S. (2008). MapReduce: simplified data processing on large clusters. *Communications of the ACM, 51*(1), 107–113.

Dellarocas, C. (2006). Reputation mechanisms. In T. Hendershott (Ed.), *Handbook of economics and information systems* (pp. 629–660). Amsterdam: Elsevier.

Dempster, A. P., Laird, N. M., Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological), 39*(1), 1–38. doi:10.2307/2984875.

DePalma, G., Srivastava, S., Liu, C. (2012). The parallel expectation maximization algorithm: Novel ideas in distributed iterative statistical computing. Technical Report, Purdue University. http://www.stat.purdue.edu/~gdepalma/Prim20Research/D_EM_JSM2012.pdf. Accessed 3 Jan 2014.

Fang, Y., Si, L., Mathur, A. (2010). Discriminative graphical models for faculty homepage discovery. *Journal of Information Retrieval, 13*(6), 618–635.

Fang, Y., Si, L., Mathur, A. (2011). Discriminative probabilistic models for expert search in heterogeneous information sources. *Journal of Information Retrieval, 14*(2), 158–177.

Handcock, M. S., Raftery, A. E., Tantrum, J. M. (2007). Model-based clustering for social networks. *Journal of the Royal Statistical Society: Series A, 170*, 301–354.

Hasan, O., Brunie, L., Pierson, J., Bertino, E. (2009). In E. Ferrari, N. Li, E. Bertino, Y. Karabulut (Eds.), *Trust management III* (pp. 65–80). Berlin: Springer.

He, Y., Liu, C. (2011). The dynamic ECME algorithm. *Journal of the Royal Statistical Society: Series B, 74*, 313–336.

Hennis, T., Lukosch, S., Veen, W. (2011). Reputation in peer-based learning environments. In O. C. Santos J. G. Boticario (Eds.), *Educational recommender systems and technologies* (pp. 95–128). Hershey: IGI.

Huberman, B. (2001). *The laws of the web*. Boston: MIT Press.

Isard, W. (1954). Location theory and trade theory: Short-run analysis. *Quarterly Journal of Economics, 68,* 305–322.

Jøsang, A., Golbeck, J. (2009, Sept). Challenges for robust trust and reputations systems. Paper presented at the 5th International Workshop on Security and Trust Management (STM 2009), Saint Malo, France.

Kraut, R. E., Resnick, P. (2012). *Building successful online communities: Evidence-based social design*. Cambridge: MIT Press.

Kuk, G. (2006). Strategic interaction and knowledge sharing in the KDE developer mailing list. *Management Science, 52*(7), 1031–1042.

Kumar, R., Novak, J., Tomkins, A. (2010). *Structure and evolution of online social networks. In link mining: Models, algorithms, and applications* (pp. 337–357). New York: Springer.

Lewandowski, A., Liu, C., Vander Wiel, S. (2010). Parameter expansion and efficient inference. *Statistical Science, 25,* 533–544.

Liang, F., Liu, C., Carroll, R. J. (2010). *Advanced Markov Chain Monte Carlo methods: Learning from past samples.* West Sussex: Wiley.

Liberto, J. (2012, Apr 19). CEO pay is 380 times average worker's— AFL-CIO. CNN. http://money.cnn.com/2012/04/19/news/economy/ceo-pay/index.htm. Accessed 3 Jan 2014.

Lochner, K., Kawachi, I., Kennedy, B. P. (1999). Social capital: An introduction to its measurement. *Health and Place, 5*(1), 259–270.

Madey, G., Freeh, V., Tynan, R. (2002, Aug). The open source software development phenomenon: An analysis based on social network theory. Paper presented at the America's Conference on Information Systems, Dallas, TX.

Masum, H., Tovey, M. (2012). *The reputation society: How online opinions are reshaping the offline world*. Cambridge: MIT Press.

Matei, S. A. (2004). The impact of state-level social capital on emergence of virtual communities. *Journal of Broadcasting and Electronic Media, 48*(1), 23–40.

Matei, S. A., Bruno, R. J. (2012). Individualist motivators and community functional constraints in social media: The case of wikis and Wikipedia. In F. Comunello (Ed.), *Networked sociability and individualism: Technology for personal and professional relationships* (pp. 1–23). Hershey: IGI.

Matei, S. A., Bruno, R. J., Faiola, A., Morris, P. L. (2010a, Sept). Visible effort: A social entropy methodology for managing computer-mediated collaborative learning. Paper presented at the Fourth Global Communication Forum, Shanghai, China.

Matei, S. A., Oh, K., Bruno, R. J. (2010b). Collaboration and communication in online environments: A social entropy approach. In M. Oancea (Ed.), *Comunicare şi comportament organizaţional* (pp. 83–99). Bucharest: Printech.

Miller, M. H. (2010, March 16). Students use Wikipedia early and often. *The Chronicle of Higher Education.* Blog: Wired Campus. http://chronicle.com/blogs/wiredcampus/students-use-wikipedia-earlyoften/21850. Accessed 3 Jan 2014.

Mockus, A., Fielding, R., Herbsleb, J. (2002). Two case studies of open source software: Apache and Mozilla. *ACM Transactions on Software Engineering and Methodology, 11*(3), 309–346.

Moore, T., Clayton, R. (2008). Evaluating the wisdom of the crowds in assessing phishing websites. *Lecture Notes in Computer Science, 5143,* 16–30.

Ortega, F., González-Barahona, J. M., Robles, G. (2008). On the inequality of contributions to Wikipedia. In Proceedings of the 41st Hawaii international conference on system sciences (HICSS '08) (p. 304). Washington, DC: IEEE Computer Society.

Raymond, E. (1999). *The cathedral and the bazaar: Musings on Linux and open source by an accidental revolutionary.* Cambridge: O'Reilly.

Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal, 27,* 379–423, 623–656.

Shirky, C. (2008). *Here comes everybody*. New York: Penguin.

Si, L., Kanungo, T., Huang, X. (2005). Boosting performance of bio-entity recognition by combining results from multiple systems. Paper presented at the 5th ACM SIGKDD International Workshop on Data Mining in Bioinformatics (BIOKDD), Chicago, IL.

Snijders, T. A., Pattison, P. E., Robins, G. L., Handcock, M. S. (2006). New specifications for exponential random graph models. *Sociological Methodology, 36,* 99–153.

U. S. Census Bureau. (2012). Historical income tables: Households. http://www.census.gov/hhes/www/income/data/historical/household/. Accessed 3 Jan 2014.

Weinberger, D. (2011). *Too big to know: Rethinking knowledge now that the facts aren't the facts, experts are everywhere, and the smartest person in the room is the room.* New York: Basic Books.

Welser, H. T., Gleave, E., Fisher, D., Smith, M. (2007). Visualizing the signatures of social roles in online discussion groups. *Journal of Social Structure, 8*(2), 1–32.

Zhang, D., Si, L., Rego, V. J. (2012). Sentiment detection with auxiliary data. *Information Retrieval, 15*(3–4), 373–390.

Chapter 2
Building Trusted Social Media Communities: A Research Roadmap for Promoting Credible Content

Ben Shneiderman

Introduction

Trusted contributors who provide *credible content* are vital nutrients for successful social media communities. When community members can rely on responses to questions, restaurant reviews, or health-care recommendations, they may benefit personally and be more likely to help others. Social capital as well as tangible economic benefits grows when good deeds are rewarded and malicious actions are suppressed. In addition, *reliable resources* of software, hardware, servers, and networks provide the technical foundation, while *responsible organizations* ensure a robust socio-technical foundation.

Techniques for assessing credibility and design principles that encourage trustworthy behavior are still emerging as the web, mobile, and social technologies mature. Early studies of website credibility focused on surface features such as spelling errors, willingness to provide contact information, professional appearance, rapid response, recognizable domain name, recency of content, and volume of information [4–8, 20]. Later work began to emphasize external markers such as verifiable seals of approval (e.g., eTrust, BBB, Microsoft MVP), public reputations based on long-term performance (e.g., eBay, Amazon), references from other users (e.g., likes, confirmations, badges, karma points), and visible histories of activities (e.g., Wikipedia edits, Amazon reviews). These more complex systems are still maturing as community site managers refine designs to promote more credible content that is less subject to deceptive practices [11, 14, 15, 17, 28].

The distinctive open nature of social media communities means that millions of people may post content such as reviews, answers to questions, videos, or comments on blogs. This significant design choice opens up participation broadly, but presents new challenges to researchers and community leaders. Off-topic postings, links to commercial or pornographic sites, and libelous attacks can easily disrupt and undermine a thriving community. The volume of posting means that centralized review is difficult, so automated and social approaches to ensuring credibility are necessary.

B. Shneiderman (✉)
University of Maryland, College Park, MD 20742, USA
e-mail: ben@cs.umd.edu

E. Bertino, S. A. Matei (eds.), *Roles, Trust, and Reputation in Social Media Knowledge Markets,* Computational Social Sciences, DOI 10.1007/978-3-319-05467-4_2,

Dangers exist from those who build reputations artificially or legitimately with the goal of ultimately providing misleading advice [15]. These botnet-facilitated deceptions and cleverly designed moles require more sophisticated filters to detect. Simpler, but effective, threats come from individuals self-promoting their work, companies surreptitiously promoting their products, or political actors undermining opponents. Criminals and terrorists may be few in number but they have more troubling agendas and they are often well organized and knowledgeable. The disturbing reality is that trust is fragile, so that a small fraction of misleading or malicious postings can undermine an otherwise trustworthy community.

While small social media communities are rarely attacked, as they succeed they become more attractive targets, requiring increasingly diligent monitoring to preserve credible content. Wikipedia has developed an especially rich set of protections, since small slips become newsworthy stories that can dramatically undermine a long history of positive reputation. As more people depend on social media communities for travel, health, financial, and legal information, increased research and greater diligence on the part of community leaders is necessary.

A research agenda that addresses all these threats will produce a broad range of recommendations. However, traditional controlled laboratory experiments have little relevance in the large bustling world of social media communities. Reductionist models are less relevant, and the number of uncontrollable variables is large. At the same time, interventions in functioning systems can be difficult to arrange and have their own risks. Therefore, partnerships between industry system managers and academic researchers could prove to be beneficial. By combining applied and basic research, which is informed by practical and theoretical frameworks, high-impact outcomes seem possible. Repeated case studies using design interventions produce data that can support theories, principles, and guidelines. Such systematic interventions in working systems may prove to be the most valuable approach. Of course, automated logging when combined with ethnographic observations, in-depth interviews, and validated surveys have the potential to produce actionable research results.

Research on scalable organizational structures and processes are a further opportunity. Just as large organizations must have a hierarchy, or other structure, online communities will need to have multiple levels of management and leadership. The Reader-to-Leader Framework suggests how multiple levels of participation can be designed into systems [21]. Successful communities have a large number of readers of the content, but often the number of content contributors may be in the neighborhood of one percent of the readers. Those who become active collaborators, engaging in discussions with other contributors are a still smaller circle. Those who rise to leadership positions to guide design processes, cope with problems, and mentor novices is a still narrower circle, but an essential component to a thriving community. In large communities, such as Wikipedia, there are many formal policies and evolving norms, so there is often a great deal for newcomers or aspiring leaders to learn. Creating motivations for readers to become contributors, then collaborators, and eventually leaders is crucial. Then, providing recognition for those who contribute actively or collaborate productively is a further challenge. Research opportunities abound for those seeking to study how visible recognition of positive contributions (downloads, likes, retweets, etc.) and rewards for substantial efforts (leaderboard of most prolific contributors, selection as a Wikipedia featured article, Most Valuable Professional awards).

The leaders help set inspirational agendas, promote behavioral norms by their examples, take the community into new directions, and deal with a wide variety of threats. Successful communities must develop leaders who create resilient social structures to deal with serious threats from hackers who maliciously violate privacy, attack servers, vandalize content, or provide misleading content

Even large communities can go astray, failing to attract, motivate, and recognize contributors adequately. These communities can also face challenges from malicious participants who wish to subvert the community for their own purposes. Worse still, internal dissent, corrupt leaders, or failure to serve stakeholders can rapidly undermine trust, which may be difficult to recover. This was the scenario for Digg's failure (http://www.computerworld.com/s/article/9214796/Elgan_Why_Digg_failed). Therefore, independent oversight by external bodies with high reputation offers a proven approach for corporations, government agencies, or universities that could be valuable in social media communities.

Previous work on web credibility guidelines provides a foundation for social media community credibility, but the shift from a centralized web construction model to an open participatory community environment introduces many new concerns.

The Stanford Web Credibility Project (http://en.wikipedia.org/wiki/Stanford_Web_Credibility_Project) compiled 10 reasonable guidelines [5–7]:

1. Make it easy to verify the accuracy of the information on your site.
2. Show that there is a real organization behind your site.
3. Highlight the expertise in your organization and in the content and services you provide.
4. Show that honest and trustworthy people stand behind your site.
5. Make it easy to contact you.
6. Design your site so it looks professional (or is appropriate for your purpose).
7. Make your site easy to use—and useful.
8. Update your site's content often (at least show it has been reviewed recently).
9. Use restraint with any promotional content (e.g., ads, offers).
10. Avoid errors of all types, no matter how small they seem.

Others have extended the list of web credibility guidelines up to 39 items (http://conversionxl.com/website-credibility-checklist-factors/#), such as showing staff bios and photos, client lists, testimonials, and trust marks. A workshop devoted to web credibility contains a set of early helpful papers (http://projects.ischool.washington.edu/credibility/).

These are valuable points of departure, but the open nature of social media communities presents far greater challenges for researchers, community leaders, and community members who seek credible content. Research on trust in social media communities [9] is a growing topic, which deserves further attention.

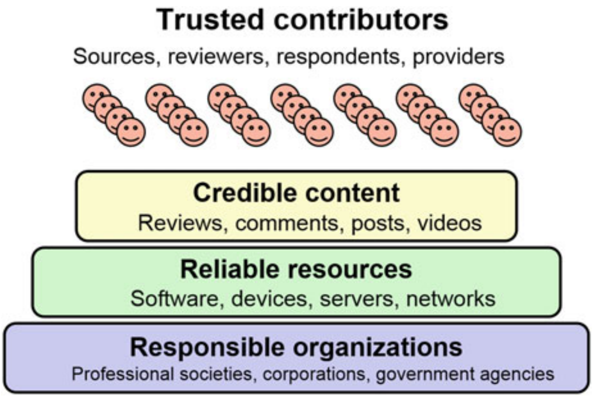


Fig. 2.1 A framework for analysis of social media communities. Ideally, trusted contributors provide credible content that is delivered by reliable resources, guided by responsible organizations. However, contributors may be misinformed, biased, or malicious, so their content is not credible. Similarly, physical resources can be undermined and organizations may be subverted or become corrupt

Framework for Credible Communities

Like the proverbial elephant, there are many ways to think of social media communities. Sociologists may focus on the bounded nature of community members and seek to ensure that only trusted contributors participate [30]. Natural language researchers may study the inherent sentiment or linguistic patterns in the millions of posts, looking for indicators of credible content. Privacy and security analysts want to certify the software, control the devices, restrict access to servers, and protect their networks, while social theorists focus on responsible organizations such as professional societies, corporations, and government agencies.

There are undoubtedly more ways of thinking about credible communities, but these four components (Fig. 2.1) already constitute a large and complex socio-technical system that provides a plethora of research opportunities. At the same time, this four-component framework gives community leaders and members a way to organize their discussions and actions so as to raise their credibility. Each component suggests research tasks, the need for operational tools, and the development of guidelines for community leaders and members. For management effectiveness, quality metrics will be needed to monitor changes and assess the impact of systematic interventions.

A few initial thoughts may trigger deeper thinking and constructive work on (1) trusted contributors, (2) credible content, (3) reliable resources, and (4) responsible organizations.

1. Trusted Contributors

Every community would like to have only trusted contributors, but the rough reality is that many contributors are misinformed even if they are well-intentioned. They can give misleading medical advice or incomplete financial information, which could have devastating effects. Second, contributors may be biased, so they present only favorable book reviews or report only good restaurant experiences. Third, contributors may be maliciously seeking to undermine a competitor's products or a political opponent's reputation.

Many strategies are being tried to ensure that only trusted contributors participate, such as raising the barriers to entry for contributors by requiring a log-in (no anonymous contributions), identity verification, background check, probation periods, and public performance histories. Greater transparency about who the contributors are and what their past is has the potential to increase trust in their future contributions.

Ancient social processes are finding new instantiations in online communities to help ensure trusted contributors. Some communities require recommendations from members to admit new members, a waiting period before contributions are accepted, or several stages of membership so that novices have limited privileges, which are increased as positive contributions are made. However, research to validate, measure, and refine these techniques will be necessary to support practice and develop effective social theories.

Network analysis to reveal past histories of troubling relationships with known malefactors could be a powerful approach [9–11, 30]. In some cases, such as with Twitter, follower and following relationships are accessible so deeper understanding of social relationship is possible, but clever users have developed strategies to appear trustworthy or cover troubling histories. Research on advanced network analysis techniques could improve their efficacy and resistance to subversion [23]. Trustworthy contributors are likely to be related to other trustworthy contributors, but developing a metric based on networks would be a helpful strategy.

2. Credible Content

The core of community credibility is credible content: movie reviews, responses to technical questions, blog posts about travel destinations, how-to videos, and much more. Verifying that each content offering is credible is an enormous and impossible task, especially as the volume and pace grows. Even within the range of credible content, there is a wide range in quality [12, 19] of content, ranging from brief notes to detailed commentaries with evidence to support claims. Studies of question-answering websites have shown that those websites that require question askers to pay for answers produce higher-quality answers.

While encouraging high quality is one research goal, another is filtering out off-topic, inappropriate, or unhelpful postings. Spam filters for email have been refined enough to work quite reliably and rapidly, but that experience is only partially applicable to building credible communities. Tracking contributors and comparing content against blacklist databases of names and spam messages are basic approaches, which could be adopted for social media communities [22]. In addition, research on sophisticated text analysis of individual content items and comparisons with similar items can all help to ensure that only credible content is ever made public. However, these filters are imperfect and attackers will become increasingly sophisticated [3, 14]. Therefore, follow-up verifications and retrospective analyses of all content submitted by a contributor can be helpful.

Social processes such as community confirmation by votes or likes and mechanisms for community members to challenge content can also be beneficial. These processes all build awareness of the threats and a greater devotion to building a credible community. Here again, anecdotal evidence is encouraging, but systematic research and innovative interventions will be helpful. For example, changing from simple “Likes” to allow “Respect” could allow community members to make more nuanced comments on political content [27]. Community managers who wish to ensure credible content face additional challenges in dealing with political discussions or debates over controversial subjects such as climate change or abortion. Content may be seen as credible by some readers, but not by others, often leading to hostile debates that cannot be easily resolved.

While social media community designers are increasingly adding features to promote credible content, there are also leadership strategies to motivate community members to participate in credibility-supportive ways, while discouraging malicious actors. Inspirational leaders who express visionary beliefs about their community can encourage members to be more active in ensuring credible content. These leaders can promote social norms by their examples or praising actions of members, possibly tied to motivations such as altruism, egoism, collectivism (commitment to helping a community), and principlism (devotion to doing good deeds) [2, 29]. They can also arrange social processes by which the members adopt and enforce policies about content, with punishments for violators, and dispute resolution processes to deal with naturally emerging differences. A well-managed community with devoted members who care about their community may be able to inoculate itself against threats and show resilience after attacks or damaging episodes. Research that tracks threats, attacks, and resilient responses could provide valuable guidelines for managers and predictive theories.

3. Reliable Resources

A credible community depends on reliable resources, including trustworthy software, dependable devices, well-managed servers, and secure networks. Each of these software and hardware components has large research communities devoted to self-improvement, but since all these components are needed to produce a credible community, there are many paths to failures. Bug-free software, secure devices, non-stop servers, and private networks are all fantasies promoted by many well-intentioned people, but the reality of these complex systems is that they are dangerously vulnerable [13, 14].

Strong privacy protection builds trust and credibility. Users who fear that their identity, personal data, address, or photo will be exposed beyond the range of those who they grant permission will resist participating or provide only partial information. Research on privacy is a vast topic already, with progress being made about enabling users to understand and specify their privacy requirements [1].

The realistic response is to strive for reliable resources, while continuously monitoring performance and repairing problems promptly. Another part of a realistic response is to make honest statements to all stakeholders about the vulnerabilities, report openly about failures, and invite efforts to make improvements. Active research continues on these issues because so much of every country’s national infrastructure depends on reliable resources. Social media communities have some special needs because of the large and rapidly growing numbers of users, the high variance between normal and peak usages, and because malicious actors often target these resources.

4. Responsible Organizations

We all like to believe that our large international, national, or local organizations are responsible, accountable, and even liable for failures. We all like to believe that these organizations are run by informed leaders acting on behalf of their members with integrity and honesty. Once again, the reality falls far behind the expectations, producing organizations that are corrupt, self-serving, or incompetent.

While there are no guaranteed methods to ensure responsible organizations, the goal is an important one that needs discussion and research. Internal audits, transparent processes, and open reporting of performance are good starts. However, independent oversight by trusted external organizations is still a valuable approach. Better Business Bureau Online, eTrust.org, and truste.com offer some approaches that could help build more credible communities, but research on still newer approaches will be beneficial.

Independent oversight can occur in many ways. Continuous oversight by trusted individuals or organizations is effective but expensive. A less costly approach is annual reviews, such as corporate audits, which are commonly done, but vary in their effectiveness. Strong annual reviews by informed panels who have open access to historical records can lead to valuable reports and recommendations, but the follow-up to ensure that recommendations are followed is vital. Finally, review panels when disasters occur, such as in airline crashes, can lead to recommendations to reduce future threats, but only if conducted in an open environment with full disclosure of reports [18].

Conclusion

The promise of social media communities is that they lower barriers to participation so as to create valuable resources, give assistance where needed, and promote more informed decisions among billions of users. However, the reality is more troubling. Misinformed, biased, and malicious contributors could produce harmful content that would undermine trust enough to destroy the value of these communities. Other threats such as corrupt leaders and internal strife can also undermine otherwise credible communities.

A substantial research effort will be needed to raise the possibility that outcomes will be positive. The research agenda offers rich possibilities for many disciplines and interdisciplines. Multiple research methods, including novel ones, will be needed because of the tightly interrelated nature of social media communities, which defy reductionist approaches. Carefully monitored interventions and rigorous case studies are likely to be more valuable than controlled experiments. Furthermore, research projects that combine basic and applied goals, practical and theoretical approaches, and mission-driven and curiosity-driven aspirations seem more promising than fragmentary efforts [24].

At the same time, designers of social media communities will have to work diligently to produce effective user interfaces, supported by reliable resources, so that community leaders and members can contribute credible content while they help raise the quality of everyone's contributions. There is also research to be done by software, hardware, and network designers, as well as by organizational designers. Responsible organizations can have powerful impacts, especially when their actions encourage every individual contributor to produce credible content.

Acknowledgements Thanks to the conference organizers, SorinMatei and Elisa Bertino, for inviting this keynote and encouraging my work on credibility. I appreciate helpful comments from Brian Butler, B.J.Fogg, Jennifer Golbeck, Michael Hicks, ItaiHimelboim, Jonathan Lazar, Alan Mislove, and Jennifer Preece.

References

Baden, R., Bender, A., Spring, N., Bhattacharjee, B., & Starin, D. (2009). Persona: An online social network with user-defined privacy. *ACM SIGCOMM Computer Communications Review*, *39*(4)*,* 135–146.

Batson, C.D., Ahmad, N., & Tsang, J. (2002). Four motives for community involvement, *Journal of Social Issues, 3*(58), 429–445.

Beutel, A., Xu, W., Guruswami, V., Palow, C., & Faloutsos, C. (2013, May). CopyCatch: Stopping group attacks by spotting lockstep behavior in social networks. Proceedings of the 22nd International World Wide Web Conference (WWW'13), Rio de Janeiro, Brazil (pp. 119–130).

Ceaparu, I., Demner, D., Hung, E., Zhao, H., & Shneiderman, B. (2002). "In Web We Trust": Establishing strategic trust among online customers. In Rust, R. & Kannan, P. K. (Eds.), *E-service* (pp. 90–107). Armonk: M. E. Sharpe Pubs.

Fogg, B. J. (2002). *Persuasive technology: using computers to change what we think and do*. San Francisco: Morgan Kaufmann.

Fogg, B. J., Marshall, J., Osipovich, A., Varma, C, Laraki, O., Fang, N., Paul, J., Rangnekar, A., Shon, J., Swani, P., & Treinen, M. (2000). Elements that affect web credibility: early results from a self-report study. In Proceedings of the CHI '00 Extended Abstracts on Human Factors in Computing Systems. doi:10.1145/633292.633460.

Fogg, B. J., & Tseng, H. (1999). The elements of computer credibility. In Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (pp. 80–87).

Friedman, B., Kahn, P. H. Jr., & Howe, D. C. (2000, Dec.). Trust online, *Communications of the ACM*, *43*(12), 34–40. doi:10.1145/355112.355120.

Golbeck, J. (2009). *Computing with trust*. Berlin: Springer.

Golbeck, J. (2013). *Analyzing the social web*. Boston: Morgan Kaufmann.

Hansen, D., Shneiderman, B., & Smith, M. (2010). *Analyzing social media networks with NodeXL*. Boston: Morgan Kaufmann.

Harper, F. M., Raban, D., Rafaeli, S., & Konstan, J. A. (2008). Predictors of answer quality in online QA sites. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM (pp. 865–874).

Jim, T., Swamy, N., & Hicks, M. (2007). Defeating scripting attacks with browser-enforced embedded policies. In Proceedings of the International World Wide Web Conference (WWW) (pp. 601–610).

Kakhki, A. M., Kliman-Silver, C., & Mislove, A. (2013). Iolaus: Securing online content rating systems. In Proceedings of the 22nd International World Wide Web Conference (WWW'13), Rio de Janeiro, Brazil (pp. 919–930).

Kittur, A., Suh, B., & Chi, E. H. (2008). Can you ever trust a wiki?: Impacting perceived trustworthiness in wikipedia. In Proceedings of the ACM 2008 conference on Computer supported cooperative work.

Lazar, J., Meiselwitz, G., & Feng, J. (2007). Understanding web credibility: A synthesis of the research literature. *Foundations and Trends in HCI 1, 2,* 139–202.

McKnight, D. H., & Kacmar, C. (2007). Factors and effects of information credibility. In Proceedings of the 9th International Conference on Electronic Commerce, Minneapolis, MN, USA

National Academies Committee on Technical and Privacy Dimensions of Information for Terrorism Prevention and Other National Goals. (2008). *Protecting individual privacy in the struggle against terrorists: A framework for program assessment*. Washington, DC: National Academies Press. http://www.nap.edu/catalog.php?record_id=12452

Nichols, J., Zhou, M., Yang, H., Kang, J.-H., & Sun, X. H. (2013). Analyzing the quality of information solicited from targeted strangers on social media. In Proceedings of the Conference on Computer Supported Cooperative Work (pp. 967–976).

Pirolli, P., Wollny, E., & Suh, B. (2009). So you know you're getting the best possible information: A tool that increases Wikipedia credibility. In Proceedings of ACM CHI (pp. 1505–1508).

Preece, J., & Shneiderman, B. (2009). The Reader-to-Leader Framework: Motivating technology-mediated social participation. *AIS Transactions on Human-Computer Interaction 1, 1,* 13–32. http://aisel.aisnet.org/thci/vol1/iss1/5/.

Ratkiewicz, J., Conover, M., Meiss, M., Gonçalves, B., Patil, S., & Menczer, F. (2011). Detecting and tracking political abuse in social media. In IEEE International Conference on Web Search and Social Media (ICWSM).

Ratkiewicz, J., Conover, M., Meiss, M., Gonçalves, B., Patil, S., Flammini, A., & Menczer, F. (2011). Truthy: Mapping the spread of astroturf in microblog streams. In Proceedings of the International World Wide Web Conference (Companion Volume) (pp. 249–252).

Rubin, V. L. (2010). On deception and deception detection: content analysis of computer-mediated stated beliefs. In Proceedings of the 73rd ASIST Annual Meeting.

Shneiderman, B. (2000). Designing trust into online experiences. *ACM Communication, 43*(12), 57–59.

Shneiderman, B. (2013). Toward an ecological model of research and development. The Atlantic. http://www.theatlantic.com/technology/archive/2013/04/toward-an-ecological-model-of-research-and-development/275187/

Stroud, N. J., Muddiman, A., & Scacco, J. (2013). Framing comments in social media. National Communication Association, Political Communication Division, Washington DC.

Vega, L. C., Sun, Y.-T., McCrickard, D. S., & Harrison, S. (2010). Time: A method of detecting the dynamic variances of trust. In Proceedings of the 4th Workshop on Information Credibility (pp. 760–765).

Violi, N., Shneiderman, B., Hanson, A., & Rey, P. (2011, Oct.). Motivation for participation in online neighborhood watch communities: An empirical study involving invitation letters. In Proceedings of the IEEE Conference on Social Computing, IEEE Press, Piscataway, NJ.

Westerman, D., Spence, P. R., & Van Der Heide, B. (2012). A social network as information: The effect of system generated reports of connectedness on credibility on Twitter. *Computers in Human Behavior, 28*(1), 199–206.

Wu, J-J., & Tsang, A. S. L. (2008). Factors affecting members' trust belief and behaviour intention in virtual communities. *Behaviour and Information Technology, 27*(2), 115–125.

Part II
Methods for Researching Trust and Credibility

Chapter 3
Semantic and Social Spaces: Identifying Keyword Similarity with Relations

Yun Huang, Cindy Weng, Baozhen Lee and Noshir Contractor

Introduction

With the development of Web 2.0 technologies, semantic contents and social interactions have become tightly integrated in online social networks and social media such as blogs, micro blogs, and social tagging. Moreover, user generated content on Wikipedia and citizen science sites (e.g., Scitable at Nature Publishing Group), has begun to organize and even generate knowledge from crowd participations.

Most content generated by users is noisy, ambiguous, and unstructured because of the voluntary nature of contributors and varying reliability of information resources. On the other hand, complex human interactions can provide the rich information to reveal the expertise and credibility of users and in turn optimize the process of information retrieval on the Web. A key problem in constructing and mining semantic spaces is how to utilize users' preference information in the process of extracting information structures from unstructured data sources and construct a relevant concept similarity network (Steyvers and Tenenbaum 2005). By jointly considering text and relational data, we propose to analyze multiple dimensions of human expertise and behavior.

This chapter proposes a three-layer framework to integrate semantic and social networks and to reveal people's expertise based on their words and relations. To demonstrate the value of user preference in semantic analysis, we use social tagging activities on CiteUlike as an example to illustrate the potential of utilizing social relations on identifying similar concepts.

Y. Huang (✉) · C. Weng · N. Contractor
The Science of Networks in Communities (SONIC) research group,
Northwestern University, Evanston, IL, USA
e-mail: yun@northwestern.edu

B. Lee
School of Economics and Management, Jiangsu University of Science
and Technology, Zhenjiang, China

E. Bertino, S. A. Matei (eds.), *Roles, Trust, and Reputation in Social Media Knowledge Markets,* Computational Social Sciences, DOI 10.1007/978-3-319-05467-4_3,

Semantics Meets Social Networks

Semantic networks represent semantic relations among concepts using linguistic information, such as documents and keywords. There are a few approaches in constructing concept similarity networks. Formal concept analysis (FCA) is a principled way to derive formal ontology from a collection of objects and their properties (Ganter et al. 2005). The binary relations between objects and attributes reveal formal concepts and concept hierarchy. Similarly, latent semantic analysis (LSA) applies the singular value decomposition (SVD) method to reduce dimensions and construct concept networks based on the frequency relations between documents and keywords (Deerwester et al. 1990). Centering resonance analysis (CRA) (Corman et al. 2002) uses keyword adjacency relations in sentences to understand how the keywords are being used in a specific document. Whereas keyword frequency methods create insights based on a "pile of words," CRA mainly adopts the betweenness concept of social networks. These approaches mostly focus on content of documents and neglect the socio-technical context such as user preference and why and how frequently people use the keywords.

To reflect the influence of user preference information on the process of constructing concept similarity networks, tripartite network models such as high-order singular value decomposition (HOSVD) (Omberg et al. 2007) utilize three types of elements, e.g. authors, keywords, and documents. While incorporating an additional dimension with content information, the users are treated as an additional type of nodes providing more association information but not as a human agents whose preferences, expertise, and relations change the use of keywords and documents.

There are some recent technical approaches to integrate semantic and social networks. The semantic social network (SSN) (Downes 2004) has extended the ontology of Semantic Web to online social networks. Using the Resource Description Framework (RDF), a conceptual description of information in triples and XML, every type of entity and relation is defined by descriptive vocabularies and the ontology provides a new view of Web information space connecting people and resources. As a formal method, SSN utilizes reasoning and deduction to retrieve information among people with related interests. In a more generic setting, heterogeneous information network (Sun et al. 2009) considered the collection of social and semantic networks as a hybrid network of multiple types of entities and multiple relations. Mining a particular meta path, i.e. a sequence of relations among different entity types, reveals the potential similarity structures in a complex network. Both approaches take a symmetric and abstract view of entity types. For example, a user is not different from a keyword unless they are specifically characterized by ontology or meta-paths. Without a conceptual framework, these approaches are limited in their ability to establish a systematic way to combine social and semantic networks.

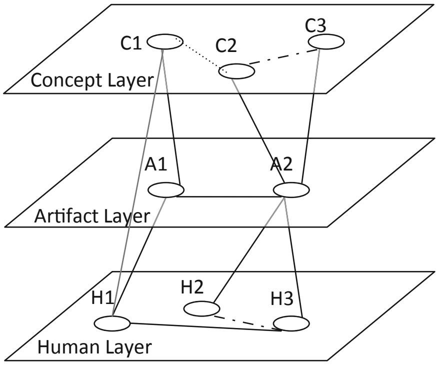


Fig. 3.1 Three-layer multi-dimensional framework

Three-Layer Framework for Multidimensional Networks

In order to reveal the inherent influences in large complex networks, we classify entities in a multi-dimensional network using three categories: human, artifacts, and concepts; and thus construct a three-layer framework representing heterogeneous knowledge and social networks (Contractor et al. 2011). Figure 3.1 describes the three layers and the relations within and between layers.

The *concept layer* represents the content domain of a knowledge and semantic network and consists of all knowledge entities including keywords/tags, properties, classes, topics. The links between entities are the logic relations defined by their ontologies. The semantic networks are either directed (e.g. semantic trees characterizing the concept hierarchy) or non-directed (e.g. concept similarity/sibling networks).

The human layer represents social networks and consists of human agents who can make decisions and actions. Each agent has a certain profile (for example status, preference, and expertise) which potentially affects its behavior. Agents could be individuals or aggregates of individuals such as groups, organizations, countries, etc. The links between two agents are their social relations and interactions such as friendship and communication. The structural tendencies of these social relations reflect the underlying motivations for creating and maintaining links such as homophily and proximity (Monge and Contractor 2003).

The artifact layer represents a collection of physical and information artifacts created by human agents—web pages, articles, products, and events. Artifacts are linked by various connections based on the content or usage such as web page hyperlinks, article citations, and product promotion events. Artifacts also act as intermediaries connect concepts with human agents. Some artifacts are associated with concepts (e.g. document-keywords and product-properties), while others link to human activities and transactions (e.g. user's web page access and product purchase).

The association relations and transactions can be projected to the concept and human layers and be used to generate derived relations. The combination of different relations can produce more information about user behavior and the knowledge domain. For example, suppose person H1 is an editor of a journal (A1) that focuses on “social network analysis” (C1). Therefore the information entity A1 establishes a relation between H1 and C1. Suppose A2 is a research paper, with keywords “friendship” (C2) and “recommender systems” (C3) in the journal A1 coauthored by H2 and H3. The artifact A2 generates many derived relations such as co-authorship between H2 and H3, keyword co-occurrence between C2 and C3, H2 and H3’s expertise indicated by keywords C2 and C3, potential interests of keyword C2 for the journal A1, etc.

The three-layer framework is very flexible in preserving various types of relations: semantic networks in the concept layer, social networks in the human layer, and association and transaction relations in the artifact layer and between layers. The concept and artifact layers represent the application scenario of content and semantic analysis, and the human and artifact layers represent the scenario of social interactions and sociomateriality (Contractor et al. 2011). Based on this framework, a multi-dimensional network can be represented as a tensor (e.g. a 3-dimensional array) and simplified through dimensionality reduction for higher-order factor analysis. For example, the higher-order generalization of SVD (HOSVD) for tensors (Lathauwer et al. 2000) is used efficiently in independent component analysis (ICA) and converts a given N-dimensional tensor into a full orthonormal system in a special ordering of singular values. It is capable of extracting clear and unique structures underlying the given multi-dimensional network (Lu et al. 2011).

Contributions of Social Relations in Identifying Concept Similarity

To demonstrate the utility of incorporating information about social relations to identify similar topic words, we use a sample data set in the social tagging website CiteULike. We compare different LSA methods. In CiteULike’s interest group “Blog and Wiki Research,” there were 2961 tagged documents between November 2004 and August 2010. From these documents, we selected 69 research papers that had been tagged by at least ten users as the test case. These documents were tagged with 169 different tags by 145 users.

We evaluated the performance of three approaches, SVD, HOSVD, user-oriented SVD (UoSVD), to demonstrate the contribution of information at different layers in the three layered framework in semantic analysis.

Latent semantic analysis (LSA) takes a sample of documents as a term-by-document matrix where each cell indicates the frequency with which each term (rows) occurs in each document (columns). Using SVD, the matrix is reduced into a low dimensional vector space in which each term and document is identified by a vector. Thus, the distance between a pair of term vectors provides a similarity measure between the two terms. In the case of social tagging, users choose some tags to annotate relevant information resources. Similarly higher-order SVD (HOSVD) uses a 3-dimensional array (term-by-document-by-user) to include user information with terms and documents and the relations among the three types of entities are used to construct concept similarity in a reduced vector space.

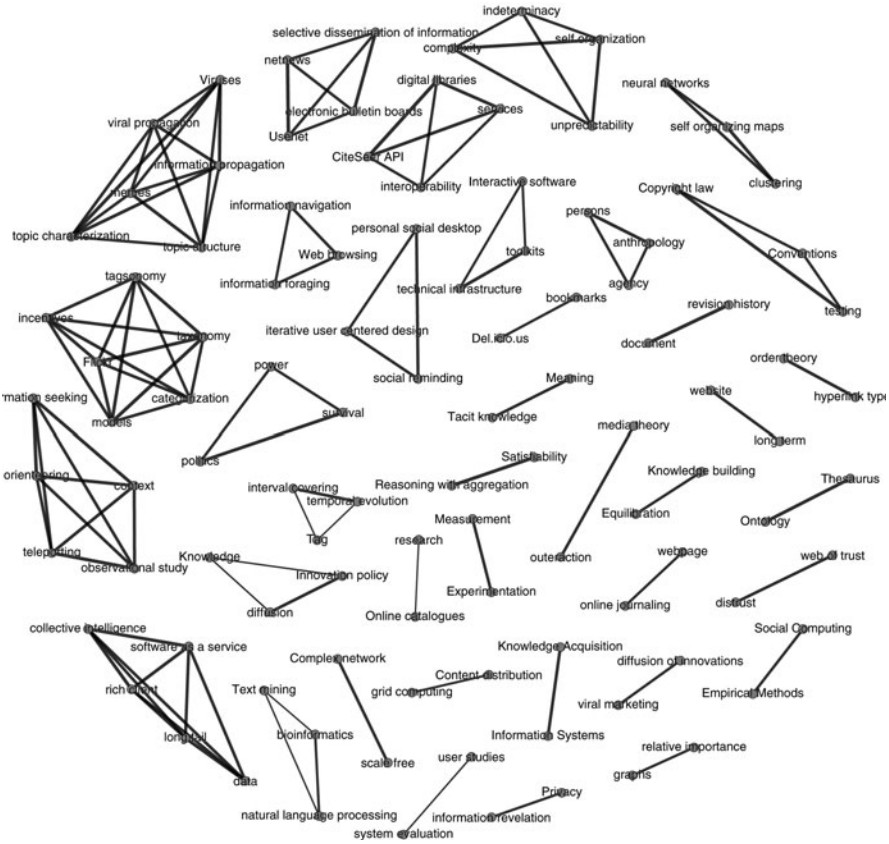


Fig. 3.2 Concept similarity network of sample data based on UoSVD (similarity > 0.4, and isolates removed)

Since the concepts depend on not only their relations to relevant documents but also the types of users who use it, the similarity between tags should consider preference information of corresponding taggers. Therefore we construct user-oriented information explicitly through a term-user matrix in which each cell is a term frequency–inverse document frequency (tf-idf) measuring the relative frequency a tagger used in the collection of the documents. In this user-oriented approach, we use both term-document and term-user matrices via traditional SVD to estimate similarity of the concepts.

Using the sample dataset, the three approaches generated keyword similarity networks with similar network structures. Figure 3.2 visualizes the concept similarity network produced by the UoSVD model. Only the links with similarity values larger than 0.4 are included and the graph shows clusters of tags with their similarity indicated by link width. No isolated tags are included.

In actuality, the link weights in the concept similarity network should be bigger between tags conceptually similar, and smaller among different tags. Therefore, the evaluation criteria should consider not only the consistency of concept similarity in

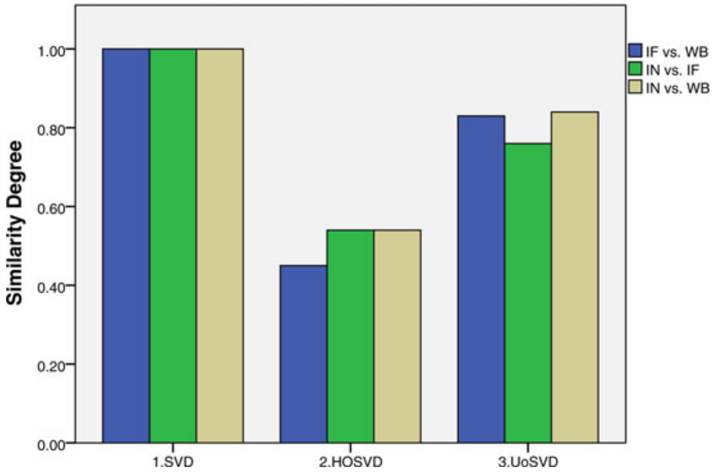


Fig. 3.3 Similarity scores among three related tags

this sample set extracted from one interest group, but also consider the discrimination power to separate different concept clusters in the set.

Figure 3.3 shows the similarity scores between each pair of elements in the cluster—information navigation (IN), information foraging (IF), web browsing (WB)—as an example. The similarity scores among the three tags based on the UoSVD have a larger power to discriminate the tags from each other, whereas the traditional SVD model produce almost the same similarity scores for each pair of tags. The UoSVD model that considers the influences of user preference can be used to discriminate the relations among relevant concepts more accurately in intra-cluster.

As a comparison, Fig. 3.4 shows similarity scores between each element in the cluster of the three tags and an unrelated term "peer to peer" (P2P). Again, we find that the UoSVD is more effective to detect the differences among concepts in different clusters than the traditional SVD and HOSVD models. Furthermore, the UoSVD has a much higher signal ratio of similarity scores of related tags vs. similarity scores of irrelevant tags and therefore has a much better performance to identify true associations among tags. On the other hand, using the same amount of information, the HOSVD achieves the worst results in discriminating tags. This suggests that user behavior and social relations provide a different mechanism in influencing semantic networks, and user information does not reveal the association of tags directly.

Based on accuracy evaluation methods in the literature (Breese et al. 1998), we evaluate the consistency of concept similarity in the concept similarity network using the variance scoring metric. The expected variance of concept similarity between tag i and other tags is defined as

$$V_i^2 = \frac{1}{m-1}\sum_{j}^{m-1}(S_{ij} - \frac{1}{m-1}\sum_{j=1}^{m-1} S_{ij})$$

where S_{ij} is the similarity score between tags i and j. The overall consistency of all tags equals to one minus the average variance of all tags.

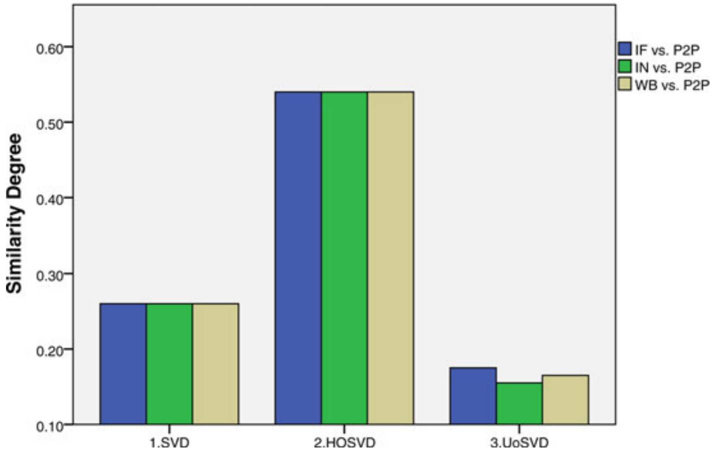


Fig. 3.4 Similarity scores between each of the three tags and an irrelevant tag "peer to peer (*P2P*)"

We evaluate the discrimination of concept similarity in the concept network using the range scoring metric. The expected range of similarity for tag i is defined as

$$R_i = \frac{S_{ij}^{\max} - S_{ij}^{\min}}{S_{ij}^{\max}} \forall j$$

where S_{ij} is the similarity score between tags i and j, and $S_{ij}^{\max}$ is the maximum possible similarity score between tag i and other tags, and $S_{ij}^{\min}$is the minimum similarity degree.

We consider the whole dataset as one cluster and calculate the average variance of similarity for each tag in concept similarity networks generated by different algorithms. Figure 3.5 illustrates the overall consistency and average discrimination scores of concept similarity in the three approaches. The results show that the UoSVD, which utilizes user preference information, produces more consistent similarity scores and has a slightly better discrimination power compared to the SVD model. The size of the differences is very small because we calculated the overall consistency and discrimination scores using all possible pairs of tags. Since many tags are not related and most similarity degrees between tags are zero, the two evaluation scores are highly inflated. More detailed comparisons among different clusters, like the three tag case we illustrated above, would show further significant differences.

Conclusion and Future Work

This chapter discusses the importance of integrating social relations and semantic networks in the discovery of topics and similar concepts in social media. Instead of using a generic network model and treating all types of nodes and relations equally in a heterogeneous network, we propose a three-layer model to characterize the uniqueness of semantic and relational information as well as their interconnections. Using

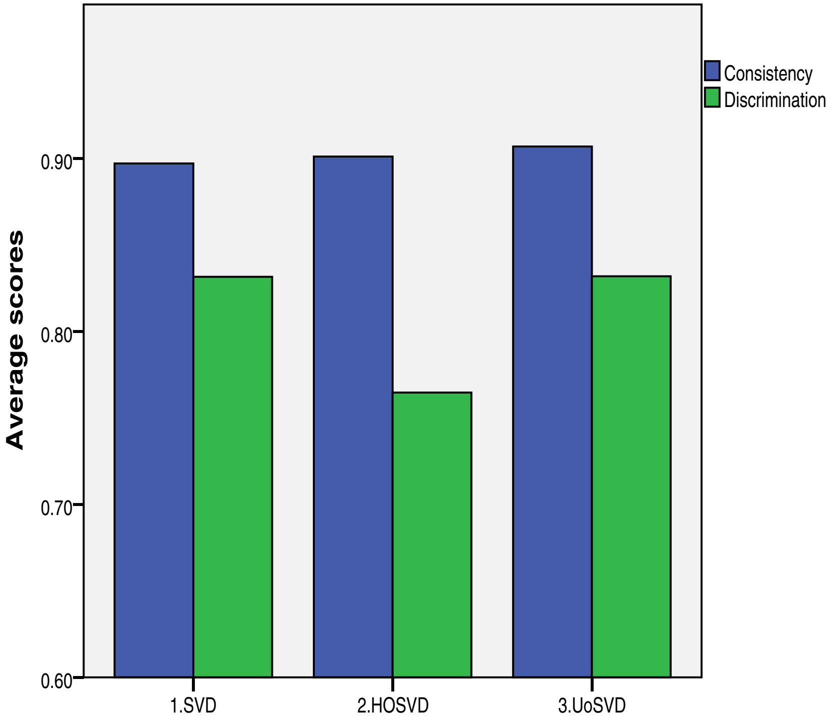


Fig. 3.5 Overall consistency and average discrimination scores of concept similarity in the three approaches

a sample case with social tagging in CiteULike, we demonstrate that the correct use of human preference and relational information can help identify similar concepts and topics. User-oriented information captures people's expertise, motivation, and preference in participating information consumption and production online, and therefore, it potentially affects the structure of the semantic networks.

The three-layer framework we introduce captures the essential structure of many application scenarios, such as scientific publication with authors, documents and keywords, Wikipedia with contributors, wiki articles and concept items, and marketing with consumers, products and features. The increasing practice of social networking makes the semantic space of online topics and content more dynamic. Further advanced methods to discover the latent social space underneath social relation networks have become a key challenge to further integrate the social and semantic networks.

Acknowledgements This study is supported by National Science Foundation (No. CNS-1010904, OCI-0904356, & IIS-0841583), Army Research Lab (W911NF-09-02-0053), the National Natural Science Foundation of China (No. 71273121/G031401).

References

Breese, J. S., Heckerman, D., & Kadie, C. (1998). *Empirical analysis of predictive algorithms for collaborative filtering*. Paper presented at the the fourteenth annual conference on uncertainty in artificial intelligence.

Contractor, N., Monge, P., & Leonardi, P. (2011). Multidimensional networks and the dynamics of sociomateriality: Bringing technology inside the network. *International Journal of Communication, 5,* 682–720.

Corman, S. R., Kuhn, T., McPhee, R. D., & Dooley, K. J. (2002). Studying complex discursive systems: Centering resonance analysis of communication. *Human Communication Research, 28*(2), 157–206.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science and Technology, 41,* 391–407.

Downes, S. (2004). The semantic social network. http://www.downes.ca/post/46. Accessed 24 Oct 2013.

Ganter, B., Stumme, G., & Wille, R. (Eds.). (2005). *Formal concept analysis: Foundations and applications*. Lecture notes in artificial intelligence (Vol. 3626). Berlin: Springer.

Lathauwer, L. D., Moor, B. D., & Vandewalle, J. (2000). A multilinear singular value decomposition. *SIAM Journal on Matrix Analysis and Applications, 21*(4), 1253–1278.

Lu, H., Plataniotis, K. N., & Venetsanopoulos, A. N. (2011). A survey of multilinear subspace learning for tensor data. *Pattern Recognition, 44*(7), 1540–1551.

Monge, P., & Contractor, N. (2003). *Theories of communication networks*. Oxford: Oxford University Press.

Omberg, L., Golub, G. H., & Alter, O. (2007). A tensor higher-order singular value decomposition for integrative analysis of DNA microarray data from different studies. *Proceedings of the National Academy of Sciences, 104*(47), 18371–18376.

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science, 29*(1), 41–78.

Sun, Y., Yu, Y., & Han, J. (2009). *Ranking-based clustering of heterogeneous information networks with star network schema*. Paper presented at the ACM SIGKDD international conference on Knowledge Discovery and Data Mining (KDD'09).

Chapter 4
Emergent Social Roles in Wikipedia's Breaking News Collaborations

Brian C. Keegan

Introduction

Disasters and accidents are endemic to social life, but so are the unique forms of social behavior and organization that emerge following them. The "improvisation of order out of chaos," equanimity of victims, emergence of serendipitous and egalitarian social ties, and redemptive moments of solidarity have characterized postcatastrophe communities for centuries, but are also intrinsically ephemeral and recede as the most acute phase passes (Quarantelli and Dynes 1977; Solnit 2010). Following unexpected and traumatic news events such as a major natural disaster, transportation accident, or mass shooting, familiar reference sources such as Wikipedia become the focus of many people seeking information to help them share information, learn about the response, and make sense of the event (Keegan 2013).

However, the vast majority of Wikipedia contributors are personally unaffected by the immediate consequences of these events, and may not have the most up-to-date information about these events. This should inhibit their motivation to devote their time and expertise to topics so remote from their interests. Furthermore, Wikipedia's policies repeatedly emphasize that the content of its articles should take a historical perspective and rely upon neutral and reliable secondary sources; prerequisites that are obviously absent in the coverage immediately following a breaking news event. In addition to these barriers, developing a collaborative account of a breaking news event on a site where "anyone can edit" would seem to inhibit, rather than promote, the generation of a reliable account. Editors' diverse motivations and skills, their lack of experience working together, no expectation of collaborating in the future, and their volition to contribute as much as they prefer should lead to major breakdowns in the process of collaborating together. The responsibilities for integrating and updating content, reverting vandalism, formatting citations, and mediating disputes are likewise diffused among all editors. This lack of clear roles or strong ties to bind participants together undermines crucial, but unstated, assumptions in many

B. C. Keegan (✉)
Northeastern University, Boston, USA
e-mail: b.keegan@neu.edu

E. Bertino, S. A. Matei (eds.), *Roles, Trust, and Reputation in Social Media Knowledge Markets,* Computational Social Sciences, DOI 10.1007/978-3-319-05467-4_4,

theoretical approaches for understanding online communities and organizational behavior. Furthermore, the volatile information environment, lack of a central authority to assign tasks, make decisions, or enforce rules, and need to sustain attention to developments over long periods of time and across broad topical areas likewise should be a recipe for profound organizational dysfunction. Yet, the top 25 Wikipedia articles by contributors every month since 2003 consist exclusively of articles pertinent to current events. Similarly, articles receiving the most unique edits and pageviews in any given week or month likewise demonstrate a substantial bias toward articles about current events. Wikipedia's coverage appears to thrive in spite of the serious challenges for organizing and coordinating responses to breaking news events on an open and large-scale collaboration system (Keegan et al. 2013).

How is Wikipedia able to cover breaking news events in spite of itself? I argue Wikipedia's ability to manage the complexities of breaking news collaborations derives from the ability of its contributors to improvise and regenerate organizational resources such as interactional roles, routines, and resources developed in previous collaborations. This would imply that breaking news collaborations involve editors who have repeatedly worked together or even specialized in editing content about breaking news articles. Analyzing these patterns requires data that can capture the interactions of Wikipedians with each other as well as changes in these interactions over time. Empirical analyses of Wikipedia collaboration structure use event logs that archive records of changes editors have made to artifacts. Event logs generally contain information about the agent, artifact, order, and action taken such as a Wikipedia editor (agent) making an edit (action) to an article (artifact) at a specific time (order). The event logs of multiple articles are often combined to extract relationships about which editors modified which articles. The resulting networks reveal large-scale patterns of collaboration around who edits which articles (Keegan et al. 2011a) and how these editing patterns are distinct from typical forms of collaboration on Wikipedia (Keegan et al. 2012a, 2013).

However, these analyses usually examine patterns of editor collaboration *across* articles rather than the evolution of editor behavior occurring *within* an editor's contribution history. The temporal ordering of sequential contributions with a single editor's event logs also encodes relationships reflecting the editor's shifting interests and attention. Looking at these records of what an editor modified over time can provide a new perspective on the structure and evolution of their role within collaborations. A "user sociotechnical trajectories" reflects the time evolution of how a single editor's behavior changed through his or her contributions to articles (Iba et al. 2010; Keegan et al. 2012b). These implicit, indirect, and latent interactions of editors' sequential modifications potentially capture unique social roles and collaboration processes that have been overlooked before.

This chapter reviews prior work that has examined relationships and social roles within Wikipedia, provides methodological detail about the construction of sociotechnical trajectories, and explores the concept with a case study of several users who edited articles related to the 2011 Japanese earthquake and tsunami. These collaborations bring together a unique cast of characters with disparate backgrounds that fulfill distinct roles in these collaborations. This analysis suggests that breaking news article collaborations rely to a great extent on interactional roles of motivated editors self-selecting into these collaborations rather than structural roles such as news editors wholly dedicated to editing breaking news articles. Editors exhibit considerable variability in the structure of their editing trajectories reflecting their diverse backgrounds. The emergence and expansion of collaborative infrastructure on these breaking articles employ more improvisational features like disaster response rather than the regeneration of collaborative infrastructures like emergency room care. I conclude by outlining a research agenda for how researchers can employ the sociotechnical trajectories of editors to understand social roles, organizational routines, and behavioral patterns that lead to more reliable user-generated content, and emergence of leadership within self-organizing systems.

Background

Networks on Wikipedia

Wikipedia is not only the "encyclopedia that anyone can edit," but the accessibility of its databases has also made it the "dataset that anyone can analyze." There are a variety of user-to-user, user-to-artifact, and artifact-to-artifact relationships that can be explored within Wikipedia (Keegan et al. 2013). Prior work on Wikipedia has analyzed the structure of editors contributing to articles (Capocci et al. 2006; Jesus et al. 2009; Laniado and Tasso 2011; Keegan et al. 2012a), articles linking to other articles (Kamps and Koolen 2009; Kane 2009; West et al. 2009), editors modifying other editors' contributions (Brandes et al. 2009; Turek et al. 2010; Keegan et al. 2012b), editors' discussions with other editors (Laniado et al. 2011; Leskovec et al. 2010; Massa 2011), and changes in these structures over time (Buriol et al. 2006; Iba et al. 2010; Scripps et al. 2009). In addition to characterizing the structure of the networks of collaborators and hyperlinks among articles, researchers have also examined how these structures influence the quality of articles (Ransbotham et al. 2012; Wilkinson and Huberman 2007; Kittur and Kraut 2008; Hu et al. 2007) and the relationships between concepts across languages (Hecht and Gergle 2010; Bao et al. 2012). However, the network structure of an editors' changing interests and roles is more difficult to capture with static network analyses—which articles did she edit first and which has she contributed to most recently? These shifts in topic and type of page over time are strong behavioral signatures of social roles yet ignored in most empirical network analyses of Wikipedia and other peer production platforms.

Social Roles on Wikipedia

Social roles describe the positions individuals hold within social structures and the expectations individuals have for their own and others' behaviors. Theories of social roles abound, but two dominant theories merit discussion. Interactionists perceive roles as focused on the individual and his or her subjective perceptions, negotiations, contextual demands, and informal interactions. Structuralists perceive roles as focused on the social environment and the cultural or institutional processes that generate patterns of behavior and relationships that individuals occupy (Biddle 1986). However, roles are not stable, but can change in accessibility (barriers to entry), prestige (social and cultural value), and contingency (relevance to specific contexts) (Callero. 1994). Gleave et al. (2009) provide a detailed theoretical and operational definition of social roles in online communities as emerging from behavioral regularities, network attributes, social actions, self-identification, and formal classifications. Social roles may also be defined as an "ecology" in which one role operates in relation to others such as antivandals acting to revert the damage done by vandals (Welser et al. 2007; Geiger and Ribes 2010).

Several previous studies have employed a social role framework to examine knowledge collaboration in Wikipedia and provided diverse findings. Although Wikipedia has some formally credentialed roles such as administrator and bureaucrat, these are a tiny minority of the editor population. The majority of editors inhabit emergent roles organized around practices such as vandal fighting, copyediting, new page patrolling, content standardization, administration, article evaluation, tool development, and new editor welcoming. Gaved et al. (2006) gave one of the earliest examinations of role ecologies on a Wiki identifying "locators" who identify specific information on a topic, "explorers" who gather general information on a topic, "grazers" who move between topics, "monitors" who check known sources, and "sharers" who make information more accessible. Kane et al. (2009) identified "flitterers" who place ideas then leave, "idea champions" who ensure the kernel of idea is maintained and evolved, and "defenders" who use technology to respond to adverse changes in the content. Yates et al. (2010) identified "placeholders," "completers," "housekeepers," and "shapers" who contribute, integrate, and synthesize content on Wikipedia. Welser et al. (2011) identify four distinct social roles: technical editors correcting small style and formatting errors, vandal fighters reverting vandalism and sanctioning norm violators, substantive experts who specialize in improving articles within a particular domain, and social networkers who use the Wiki as a platform for interpersonal relations rather than substantive contributions to content or administration. While these analyses of social roles in Wikipedia are instructive for identifying general behavioral regularities and interactions, they do not examine the roles used for high tempo knowledge collaboration that operate under very different coordination conditions.

Social Roles for High-Tempo Collaboration

Social roles also play an important part in the operation of organizations that must respond to unpredictable and urgent tasks such as disaster response (Majchrzak et al. 2007), emergency medicine (Faraj and Xiao 2006), aircraft carrier flight decks (Weick and Roberts 1993), or breaking news journalism (Berkowitz 1992). Highly differentiated and formalized roles such as attending doctor versus nurse allow individuals to adopt a swift and depersonalized trust based on arbitrary category membership heuristics alone (Meyerson et al. 1996). The roles in these systems are often stable and endure through successive temporary organizations (Bechky 2006; Bechky and Okhuysen 2011; Klein et al. 2006). However, some temporary organizations like disaster response teams lack the role clarity or group stability of other temporary organizations like emergency room teams. The former have diverse motivations, mixed perspectives, varied resources to contribute, and substantial volition to come and go as they please. Factors like these contribute to unstable task definitions and the pursuit of multiple and potentially conflicting goals. These *emergent response groups* are characterized by participants orienting to what is known about the situation, the history of actions already taken, developing "swift trust," and focusing on relationships between people and tasks rather than people and expertise (Majchrzak et al. 2007). Even these theoretical approaches assume colocation of group members and material or physical tasks, neither of which apply to distributed online Wikipedia collaborations. However, this approach emphasizes the ability for Wikipedians to step in and assume roles without prior qualifications, which is appealing for modeling Wikipedia's "anyone can edit" ethos. However, these interactionist roles have problematic implications as it suggests that editors need to "learn the ropes" and improvise the necessary social roles and behaviors rather than regenerating previously effective roles and behaviors.

Other scholars criticize approaches emphasizing temporary organizations' management of ephemerality through improvisation and "swift trust." Coordination and self-organization in temporary teams can also proceed by participants regenerating, adapting, and improvising roles and routines used in previous projects and collaborations (Klein et al. 2006; Bechky 2006; Bakker 2010; Bechky and Okhuysen 2011). Temporary organizations can be organized around enduring, structured role systems that are negotiated, reproduced, and reinforced across collaborations within industries characterized by temporary organizing. Entrants to a position find expectations through socialization and interaction, encounter and deploy resources with which to negotiate expectations, and enact the position in response to particular situations. Role expectations guide interpersonal relationships and the execution of tasks, but this role structure simultaneously provided continuity and stability that temporary projects lack (Bechky 2006; Ratcheva and Simpson 2011). This approach is appealing for the study of Wikipedia's breaking news articles because it suggests that editors occupy structural roles that allow them to specialize in particular types of editing. But because they can regenerate and adapt social roles and behaviors from prior work, this may limit their ability to incorporate innovations and best practices learned outside of this community compared to interactionist roles.

Event Logs and Sociotechnical Trajectories

To explore which of these role types prevail in Wikipedia's breaking news collaborations, editors' behavioral histories need to be collected and analyzed. Many sociotechnical systems archive records and other meta-data about changes in the state of the system into event logs. These data are valuable for editors to trace changes across versions of documents, evaluate other editors' contributions, and build additional tools to support collaboration.[1] Wikipedia editors can review the history of every change made to almost any article since the first edit as well as every revision made by any user. A *temporal adjacency* is the relationship from an artifact a user acted upon to the next artifact the user acted upon. Because sociotechnical trajectories are built from temporal adjacencies in event log data, they capture important temporal contexts and dependencies *in the structure of the network itself*. As we review below, these temporal adjacencies are overlooked in traditional network analysis approaches, but nevertheless encode complex behaviors into micro- and macro-level structures denoting distinct behavioral patterns and dispositions.

A *sociotechnical trajectory* of a user traces the path of users "moving through" the artifacts they have interacted with over time. The aggregation of temporal adjacencies in an editor's contribution history reflects the shifting interests, motivations, and roles from his or her first contribution. These contributions may be highly erratic in the case of vandal fighters moving rapidly between articles or they may be highly focused on a single topic. Using an event log archiving the records of a single user's actions to one or more artifacts, a temporal adjacency exists from artifact i to artifact j when a user's actions on artifact j immediately follow an action on artifact i. The final user trajectory ultimately contains the set of artifacts that the user has taken action on and the temporal adjacencies between artifacts based on the user's event log.

The differences in the construction and interpretations of a traditional editor–article collaboration network and sociotechnical trajectory are illustrated in Fig. 4.1. This example is drawn from the event log in Table 4.1 where one editor makes six contributions to four articles. Using the same event log, the traditional method of constructing collaboration networks of editors and articles is illustrated in the left column and the construction of the user's sociotechnical trajectory is illustrated in the right column.

1. At time 1, editor X makes a contribution to artifact A. In Fig. 4.1a, this creates a link between the editor and artifact in the collaboration network, but creates an isolated editor node in the sociotechnical artifact trajectory. Note that editor X does not appear in the user trajectory because the trajectory is unique to this user based solely on her behavior.
2. At time 2, editor X makes a contribution to artifact B and the number of articles in the collaboration grows to two which is reflected in both types of networks. However, the trajectory captures the temporal adjacency $A \rightarrow B$ that is missed in the collaboration network. In other words, the editor can be said to have moved from artifact A to artifact B.
3. At time 3, the early stages of a "chain" begin to form in the artifact trajectory (Fig. 4.1c) as the editor modifies a third article but never returns to the articles she previously edited.

[1] In the remainder of this chapter, I will use the terms "editor" and "user" interchangeably to refer to members of the Wikipedia community who make contributions to the project on articles, discussions, and other pages. However, "users" can refer generally to individuals within other sociotechnical systems while "editors" are specific to Wikipedia.

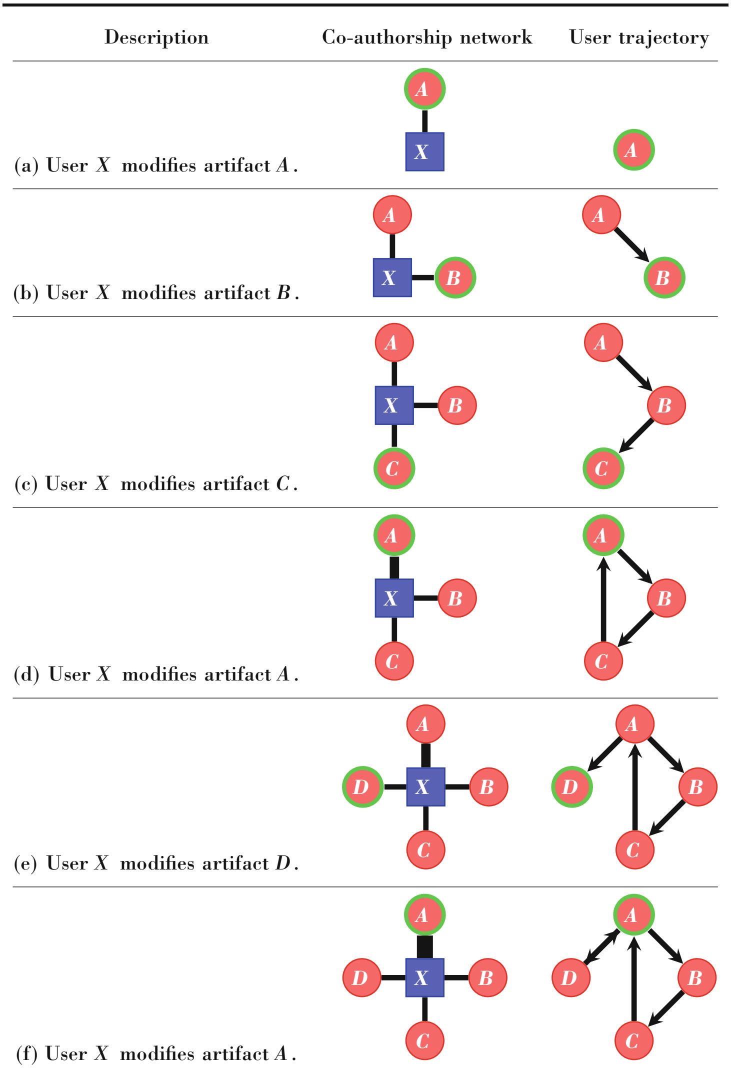


Fig. 4.1 A user sociotechnical trajectory. The user (*blue square*) contributed to 4 pages (*red circles*). Pages outlined in *green* received the most recent contribution. The edge width reflects the number of revisions the user made to the page

4. At time 4, this nascent chain ($A \to B \to C \to A$) is closed and creates a "cluster" or "cycle" where the editor returns back to editing an article she previously edited. This cycle is a particular structural form that can be detected with traditional social network metrics.
5. At time 5, that modifies artifact D. This temporal adjacency reveals A's increasing centrality as a place where the editor returns to and departs from that is obscured in the collaboration network.

Table 4.1 Example of an editor's event logs. The activities are all edits and the order are the timestamps of the contributions. The performer is user X and the cases are the set of artifacts $\{A, B, C, D\}$

Activity	Case	Performer	Order
Commit	A	X	1:01
Commit	B	X	2:02
Commit	C	X	3:03
Commit	A	X	4:04
Commit	D	X	5:05
Commit	A	X	6:06

6. At time 6, editor X's sixth contribution modifies artifact A yet again, reenforcing artifact A's centrality in the behavioral repertoire of the editor as well as creating a reciprocated link between A and D that is distinct from the cycle.

Formally, the sociotechnical trajectory of a user is a one-mode directed graph wherein an edge from artifact i to artifact j exists if and only if the user made a contribution to artifact j immediately following a contribution to artifact i in a temporally-sorted event log. Thus, a $A \rightarrow B$ dyad in an article trajectory can be interpreted as "user i contributed to artifact B after artifact A". These graphs are visualized using a combination of spring-embedding algorithms within Gephi to ensure that nodes with similar link patterns cluster together visually while nodes that do not share links tend to be repulsed. While this structural method invites the application of existing network analytic methods to understand positions, the focus here will instead be on qualitatively examining features in these editors trajectories' that predispose them or uniquely qualify them to participate in breaking news article collaborations.

The nodes in these visualizations are colored by their namespace or the page type. There are at least 14 distinct namespaces on Wikipedia, but activity is primarily concentrated in a handful of these. "Main" namespace is where the articles themselves reside, "Talk" namespace is the discussion pages associated with these articles, "User" namespace is where editors post information about themselves, "User talk" is where editors communicate with other editors, "Wikipedia" namespace is for administrative and policy-related content, "Wikipedia talk" is for discussions about these policies and procedures. The remainder about files, MediaWiki, templates, help, categories, and portals is highly specialized and make up a tiny fraction of total contribution to the entire project. Because these patterns of contribution to specific namespaces reflect distinct types of work and varying levels of familiarity with organizational norms, they are important for understanding editors' roles. The extent to which editors' contributions are concentrated in any one of these namespaces reflects some social role or specialization on the part of the editor as a contributor, copywriter, consensus-builder, vandal-fighter, policy-enforcer, or other roles.

The edges in this graph also encode information related to the delay or lag between an editor's consecutive edits. Because an editor can potentially shift from editing article A to article B many times, this edge can contain multiple lag values that can vary dramatically in their values. To simplify this array of lags, only the median value reflecting a central tendency for the editor to wait before editing the next article is used. Some lags may be very short, of the order of seconds or minutes, reflecting a highly engaged editor moving quickly to update several articles in rapid succession while other lags may be very long, of the order of months or years, reflecting an editor who went on hiatus between successive edits. These time lags are reflected in the trajectory by adjusting the darkness or opacity of the edges such that darker lines indicate shorter (median) lags reflecting immediate engagement while fainter or whiter lines indicate longer (median) lags reflecting incidental relationships. These distinctions are especially important in the context of a breaking news collaboration as the rapid engagement of editors across a variety of articles may reflect important coordination work responding to problematic editors, standardizing information across articles, or executing a decision made in discussion with others.

User Trajectories

This section explores the sociotechnical trajectories of editors who were significant contributors to articles around the 2011 Tōhoku earthquake and tsunami such as the "Fukushima Daiichi nuclear disaster" and "Fukushima Daiichi Nuclear Power Plant" (Keegan et al. 2011). These editors' contributions are almost exclusively focused on a single article or handful of articles within this general topic. User L.tak extensively involved several articles related to the "Fukishima Daiichi nuclear disaster" and "Fukushima Daiichi Nuclear Power Plant" articles, talk pages, and related pages beginning on March 11. He (or she) continued editing these articles on a daily basis until April 1, ultimately making more than 211 revisions out of the 6165 revisions on the article. User Flodded edited the main article about the earthquake and tsunami exclusively approximately 14 h into the collaboration and continued to edit daily until March 23 making 542 of the article's approximately 6000 revisions. L.tak's contributions were also wide-ranging and varied. He was the most active editor on the articles for the "power plant" article and talk page as well as the second most active editor on the "nuclear disaster" article, and seventh most active on its talk page. Like Flodded, L.tak's involvement was extensive but temporary and appears to have stopped contributing to either article after early April.

Remarkably, neither of these editors ever crossed paths: they worked on their "own" articles independently of each other despite the similarity and timeliness of their topics. Alternatively, a user like ACSE edited many articles related to this breaking news event, but concentrated attention on a pair of articles, editing the "nuclear disaster" article 160 times, the "earthquake and tsunami article" 83 times, and the other articles no more than 13 times. Thus, highly active editors appear to occupy distinct social roles as either specialists focusing solely on a single article (like Flodded) or highly related topics (like L.tak) or as something like generalists moving between several or articles like ACSE. This specialization of prolific editors contributing to only a single article or subtopic is startling as it suggests substantive coordination or collaboration in coverage proceeds through other channels and mechanisms than coauthorship of articles. These features and these editors' interactions with them will be explored in editor trajectory sections below.

Flodded

User Flodded was the prolific contributor to the "earthquake and tsunami" article, making the most contributions (560) in the corpus and is the first editor trajectory (Fig. 4.2). Flodded's first edit was made in August 2009 to the article "Shellfish" and involved updating and adding citations. S/he edited an article about a failed dot-com company "AboveNet" and then went on a lengthy hiatus until January 2011. Flodded's renewed editing activity was related to another breaking news event, Jared Lee Loughner's assassination attempt against Gabrielle Giffords in Tuscon, Arizona. Flodded edited the articles "2011 Tuscon shooting," "Jared Lee Loughner," "Gabrielle Giffords," and "United States Congressmen killed or wounded in office" in rapid succession over an 11-h period on January 11. Flodded was initially involved in copyediting the articles to remove unverifiable speculation and unencyclopedic content. As is often the case with breaking news articles, this article was "semiprotected" by administrators to limit the changes made by novice or unregistered editors. Unregistered editors or editors who have been active for fewer than 4 days and 10 edits are blocked from editing, but may make requests for edits on the talk page. Flodded was involved in responding to several of these edit requests and then became involved in an intense discussion about whether Loughner identified as an atheist on both the discussion page and "Biographies of living persons" administrative notice board. He continued to perform copyediting duties on the Loughner article, fixing capitalizations, ensuring the consistency of names and styles, and correcting grammatical mistakes as well as remaining involved in the article's discussion page. Despite the marathon 11-h editing session, Flodded abruptly stopped editing the article and did not make another contribution until February 21, performing daily antivandalism work on unrelated articles about "Extremes on earth," "Bell Mobility," "Lowest temperature record on earth," and other topics on a daily basis. However, s/he was not deeply involved in the ongoing maintenance of these articles, but simply made a single contribution and moved on to other topics. In early March 2011, he edited the article "Cheiracanthium," a genus of spiders, to update information implicating them in a recall of Mazda vehicles.

As discussed above, Flodded was a relatively early editor of the "earthquake and tsunami" article, but s/he was not among the first editors. His initial edits focused on removing over-specific information relating to areas where minor tsunami alerts had been issued justifying these edits on the talk page:

> We could list out thousands of places with tsunami warnings or that received a few extra cm of water. Obviously this is not feasible, nor is it encyclopedic. I suggest a good balance would be to only list places that have reported more than minor damage, have reported casualties, have reported large-scale evacuations in mainstream media, or are otherwise notable.

Flodded was also an extremely active editor on the discussion pages, making 257 revisions between March 11 and March 22 on topics like the looming nuclear disasters, finding sources to verify the extent to which the island of Honshu had been displaced, and increasingly on the topic of establishing reliable numbers about the casualty tolls. Flodded went on a remarkable 24-h editing marathon; between 19:35

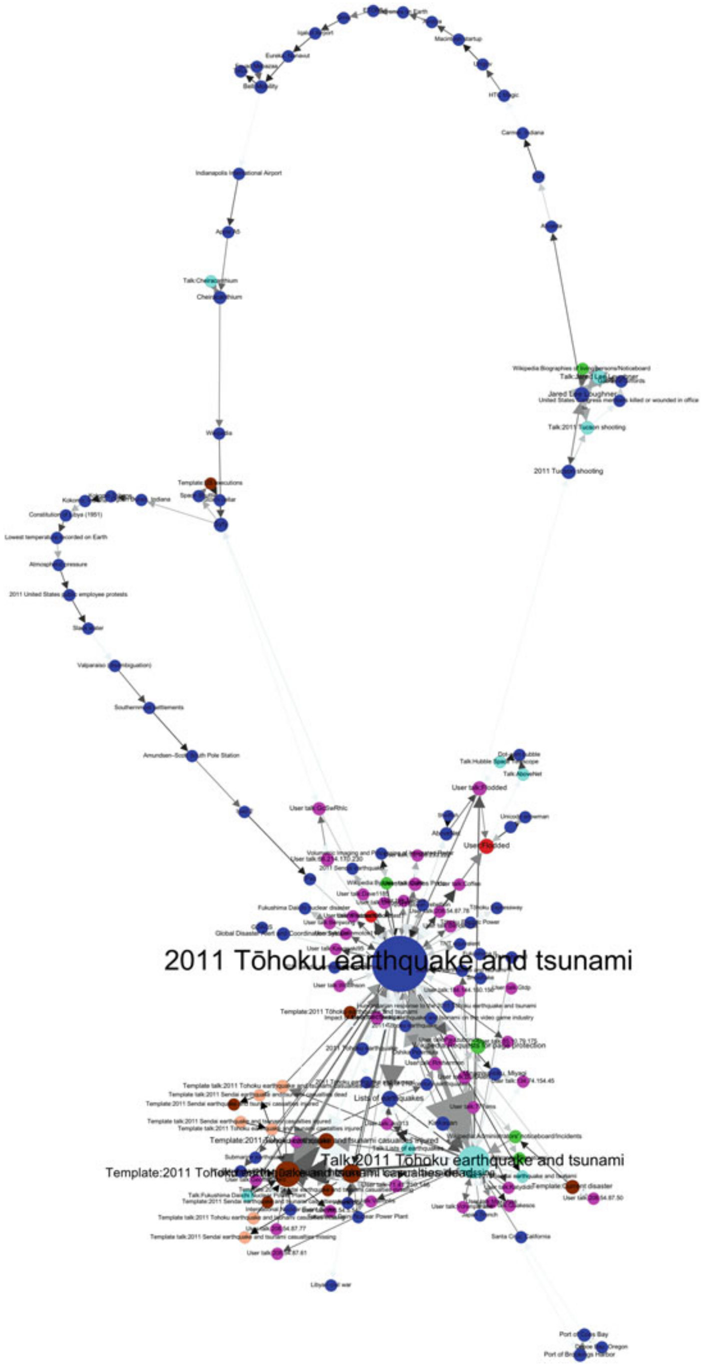


Fig. 4.2 User trajectory for Flodded

UTC on March 11 and 19:47 UTC on March 12. Flodded made several edits per hour presumably precluding the ability to sleep during this time frame. After a 7-h break, he embarked on another 24-h editing marathon stretching from 3:34 UTC on March 13 to 2:47 UTC on March 14 in which he made several changes per hour.

Returning to his editor trajectory, several structural features merit discussion. First, the graph is comparatively small, having only 149 nodes and 285 edges, but very dense (1.27e-2). The halo of light red points around the central “earthquake and tsunami” article represents the talk pages of other users Flodded communicated with about the article, warning them to stop reverting his changes or providing boilerplate welcome messages to new users cautioning them about the norms of editing on Wikipedia. This halo structure of pendants with reciprocated ties to the core article reveals that Flodded would be working on the earthquake and tsunami article, go to these users’ talk pages to warn them, and then return immediately to editing the central article again. Several articles are also present in this halo such as articles with alternative titles for the event (“2011 Sendai earthquake,” “Japanese earthquake and tsunami,” “2011 Tohoku earthquake”) that each redirect to the main article. The strong tie between the main article and the light blue dot reflects that a substantial amount of his total activity involved shuttling between the main namespace article and the article’s talk page in rapid succession, 107 transitions in total with a median edit lag of 4 min and 7 s. Flodded was also involved in a variety of administrative processes related to requesting page protection as well as filing reports related to user misbehavior which are the peripheral green nodes near the central node.

Flodded’s intense editing sessions became shorter and more infrequent and he began to shift attention to editing the casualty templates on March 16. As previously discussed, this is highly specialized and technical work involving knowledge of how to identify and locate templates, format them appropriately so they appear correctly in the rendered pages, and update the information contained within them on a regular basis. As the Japanese authorities released information about casualty numbers at the beginning and end of each day, Flodded would take these reports and update the numbers in the corresponding templates. Despite these contributions to the casualty templates, Flodded remained involved in many other aspects of the article, a “jack of all trades” involved in many discussion threads, communicating with users on their talk pages, performing copyediting, updating information on related articles such as “Lists of earthquakes by magnitude,” and participating in administrative discussions. His final edits on the topic were on March 23, and apart from 3 revisions to the Libyan civil war on April 3, Flodded has not made a single contribution since then.

Flodded fulfills an interesting role as an editor demonstrating a latent interest in not only editing articles about current events throughout his history but also unusually dedicated by contributing for 48 h in a 55-h period of time and making a substantial number of edits in the successive weeks. Although his edits were highly concentrated, he nevertheless played a crucial coordinating role discussing a variety of topics with editors on the talk and their user pages. Despite the apparent lack of an editing history which would qualify him for this type of work, Flodded fluently engaged in a variety of tasks, demonstrating knowledge of Wikipedia policies justifying his editing decisions when challenged by other editors, participating in

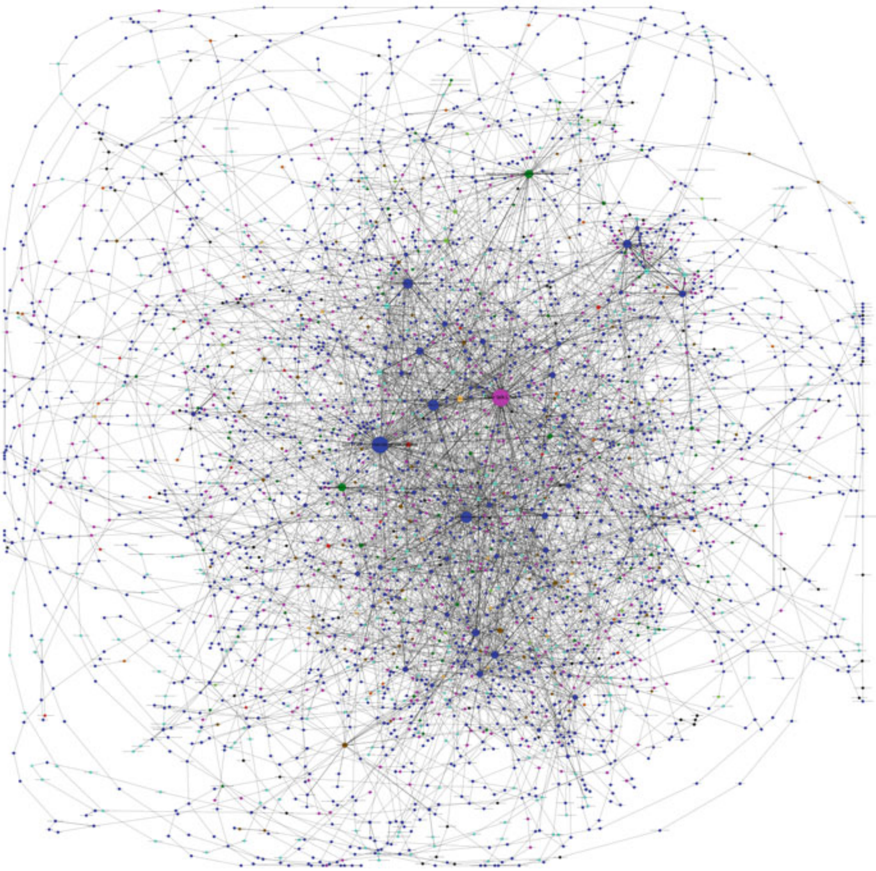


Fig. 4.3 User trajectory for L.Tak

arcane bureaucratic proceedings about protecting pages and notifying administrators of trouble, and actively developing and modifying highly specialized templates about casualty numbers.

L.Tak

User L.Tak was the second most prolific editor in the corpus, the most active editor of the "Fukushima Daiichi Nuclear Power Plant," and the second most active editor of the "nuclear disaster" article with 211 edits (after User Sandpiper's 281 edits). L.Tak's editor trajectory is plotted in Fig. 4.3. This trajectory reveals several significant differences from Flodded's structure that in turn have implications for understanding the role ecology of users responding to breaking news articles. First, it is clear that L.Tak has a substantially deeper and more varied editing history than Flodded, making 9907 revisions since making his first contributions in late October 2007 and then beginning the contribute regularly in May 2009 on the article "European Parliament election, 2009." 63.9 % of L.Tak's contributions are in the "Main" article namespace, 14.5 % in the "User talk" namespace, and 11.0 % in the "Talk" namespace for article discussions. With 3206 unique pages edited and 6105 unique edges, L.Tak has a substantially larger but also less dense (5.93e-4) trajectory than Flodded.

While Flodded had a predilection for contributing to articles about events in the news, L.Tak's extensive editing history is more complex. The most central article is his own talk page which suggests much of his activity involves responding to other editors' queries and concerns. The history of this talk page suggests a problematic debut and struggle with the learning curve of Wikipedia norms and rules initially but more recently becoming a backchannel with other editors soliciting his opinion and asking for elaboration on actions performed elsewhere. Other central articles in his trajectory concern international trade, visa, and labor agreements as well as environmental organizations. L.Tak's intense involvement in and extensive contributions to the "nuclear disaster" article motivating this analysis is, incidentally, very peripheral in his trajectory residing in the dense outlying subgraph at approximately 1 o'clock. The articles preceding his involvement in the nuclear disaster article are a variety of copyediting tasks and linking to other concepts on a variety of outwardly mundane topics like provincial and colonial governance in the Netherlands and the articles following his involvement are about the foreign relations of European countries and nuclear treaties. This trajectory suggests a passing interest in the social and cultural history about nuclear technologies and the environmental movement, information that became relevant in the aftermath of the tsunami-induced nuclear disasters.

The work L.tak performed was initially focused on the "nuclear plant" article copyediting to ensure the consistency of times and timezones, removing alarmist predictions, and plagiarized material. While L.tak did not have the marathon 24 h editing sessions of Flodded, he nevertheless made regular contributions over 6-, 8-, and even 14-h periods of time between March 11 and 15, with contributions slowing thereafter. L.Tak also fulfilled an essential coordinator role, with his contributions shuttling between the article page, discussion page, and user talk pages. The contributions L.Tak made during this time largely involved copyediting and removing duplicate information as well as adding information about the timeline of events and reliable sources.

Sandpiper

User Sandpiper was the sixth most active editor in the corpus and the most active editor of the "nuclear disaster" article and his user trajectory is plotted in Fig. 4.4. Sandpiper made 9240 revisions since starting June 2005, editing articles about Sussex and Harry Potter. Like L.Tak, his editing trajectory is also substantially more complex than Flodded but Sandpiper's trajectory also has distinct subgraphs corresponding

Fig. 4.4 User trajectory for Sandpiper

to distinct phases of his editing history. Dalliances with unrelated topics are also apparent with a burst of editing relating to articles about English radio transmitting station towers, "Cutty Sark," and a large amount of activity on the 1916 "Battle of Jutland." Like L.Tak, Sandpiper's participation in the "nuclear disaster" is not embedded within a larger subgraph of breaking news events, but a tangent from his typical edits. This trajectory is emblematic of an editor who focuses on a particular topic and works extensively on a variety of articles within it but then moves on to an entirely new topic. The diversity of the colors also reflects a diversity of activity in making changes to articles, participating in discussions, and talking to other users. This user is a generalist who specializes in both time and topic, unlike L.Tak who is a generalist, who also edits a diverse set of articles but returns back to earlier articles throughout.

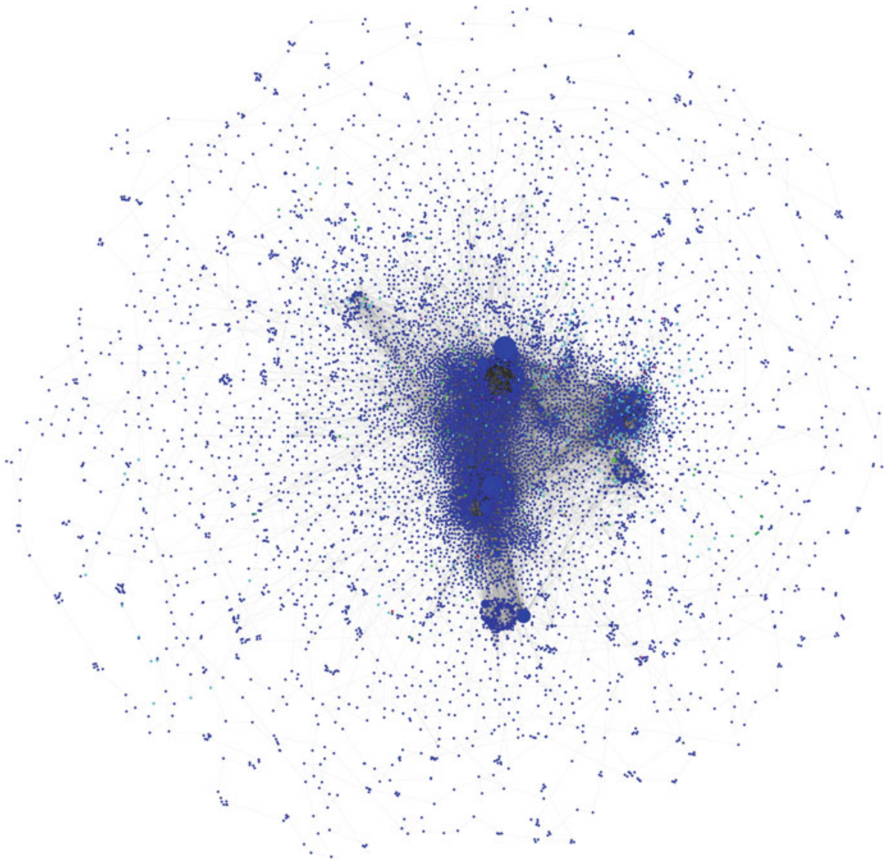


Fig. 4.5 User trajectory for ACSE

ACSE

User ACSE's trajectory is plotted in Fig. 4.5. His 41,778 revision editing history focused predominately on a strange pair of topics, Japanese pop, and Japanese serial killers. But ACSE was also the editor who contributed to 34 articles in the Tohoku corpus, updating information on many of the preexisting articles about towns, villages, and other points of interest that had been affected by the tsunami as well as editing the "earthquake and tsunami" and "nuclear disaster" articles extensively. This lack of embeddedness in a larger context of current events editing occurs in many other editor trajectories as well. Although he is not a regular editor of breaking news articles, this editing trajectory reveals a specific and important types of expertise about Japanese culture and geography. The preponderance of blue in this graph reflects the fact that ACSE engages minimally with discussions on article discussions or user talk pages—his contributions are almost exclusively audience-facing. This may reflect preferences to eschew these discussions and move on to other topics, or could also reflect the inherent credibility of his edits. The fact he edits article namespaces almost exclusively suggests his contributions may have high levels of credibility because few editors are reverting him or attempting to draw him into discussions.

Discussion

A characteristic feature of breaking news article collaborations is shifting attention across articles as collective effort initially focused on a central article (e.g., the earthquake and tsunami article) but then diffused to other articles and recentralized again on another related breaking event (the nuclear disasters) (Keegan et al. 2011a). Despite the opportunity for a single editor to make substantial contributions to each of the articles about parallel breaking news events, the most prolific editors on many articles like the nuclear disasters had negligible activity on others like the earthquake and tsunami. Examining the user trajectories of several top contributors suggests that prolific editors' investments in breaking news articles are at once novel but also reflect a latent interest or expertise in the topic. Editors of the articles about the nuclear disasters are drawn not from a cohort of editors dedicated to editing breaking news events, but rather editors like L.Tak with a background in international trade or ACSE's familiarity with Japanese pop culture. These editors' backgrounds conferred the collaborative competence, editing skills, and norm familiarity to extend and expand their repertoire of practices and routines necessary to manage a complex collaboration even if they had limited or no prior experience working on breaking news articles. This suggests that the capacity to engage in the intense coordination demanded on these articles can be acquired and learned in situ rather than developed from peripheral participation on prior breaking news articles or reliance on other editors with whom they have previously collaborated.

Wikipedia's collaborations on articles about current and breaking news events bring together a unique cast of characters with disparate backgrounds who fulfill distinct roles in these collaborations. This analysis suggests that breaking news article collaborations rely, to a great extent, on interactionist roles of motivated editors self-selecting into these articles rather than structural roles such as news editors wholly dedicated to editing breaking news articles. While editors exhibited considerable variability in the structure of their editing trajectories reflecting their diverse backgrounds, trajectories within breaking articles follow regular structural patterns reflecting the presence of a highly centralized coordinators and substantial churn in contributor cohorts. Across breaking articles, these central coordinators appear to be unique as well as otherwise inexperienced breaking news collaborators. This complicates attempts to frame these collaborations as communities of practice because they lack the deference to tenure and peripheral participation and instead appear to embody the improvisation and adaptation found in other high tempo and emergent response groups. The social roles that emerged on these breaking articles reflect more of the interactionist dimension of disaster response teams rather than the regeneration of collaborative infrastructures found in ER teams.

These findings have theoretical implications for understanding the origins and transformation of social roles and structures. As other authors have noted, roles in Wikipedia are highly informal but these breaking news articles appear especially flexible given the variance in participants' backgrounds. Breaking news articles about major news events will inevitably attract a large number of editors making only passing contributions. The responsibility for synthesizing, copyediting, and integrating these contributions fall to everyone in an open peer-production system, yet editors with some contextual background but wholly lacking the experience of working on other high tempo articles nevertheless appear to thrive and invest themselves heavily. As Bechky (2006) found in her study of role adoption, roles are not a consequent of position in a structure but resources that are claimed, negotiated, and enacted. Editors do not operate in a vacuum but continually encounter collaborations in the midst of their unfolding development complete with dependencies on synthesizing content across articles, copyediting new content, and explicitly coordinating efforts with other editors working in parallel. These overlapping dependencies constitute a dynamic environment of opportunities and resources which results in an ecology of roles which editors adopt and negotiate in response to others' actions as well as their own background.

Future Research Agenda

The cases above are illustrative of the types of analyses that can be conducted by condensing large and complex event log data into sociotechnical trajectories. Given the fluidity with which editors inhabit and shed roles in breaking news article collaborations, further analysis and methodological development is needed. In particular, the method for extracting and interpreting users' sociotechnical trajectories outlined here can be expanded into a larger research agenda to examine how users' trajectories interact with each other and overlap. The trajectory analogy can be extended in several ways to reveal temporal patterns ("velocity"), pervasive forces ("fields"), recurring patterns of actions ("orbits"), and actions preceding abrupt changes ("collisions") within sociotechnical systems:

Velocity The edges which link the nodes in artifact and user trajectories reflect the time elapsed or the delay between actions. Because some actions occur in quick succession (e.g., an antivandal bot reverting changes made by a troll), while other actions are prolonged (e.g., months passing between a editor's edits), these temporal lags can be called "velocities" to reflect the rapidity with which a user or artifact moved from one state to another. The distribution of velocities within a user suggests the intensity of work that he or she engages in. These low velocity transitions can be potentially highlighted as transitions or discounted as boundaries.

Collisions Mapping the trajectories of multiple users together provides an opportunity to analyze a trajectory's "field." Again borrowing from classical mechanics, collisions occur when two trajectories intersect. If two editors edit the same article, their respective trajectories will collide at that article (albeit at different positions along their own trajectories) and these editors may exhibit similar behavior thereafter, such as continuing to edit similar articles. If two articles are edited by the same editor, again these articles' trajectories will intersect. The position of this collision in each article trajectory might reveal whether the editor has a tendency to work on articles at certain stages of their development. The number of collisions between different users' trajectories may reveal shared latent interests or even emergent communities of practice.

Orbit Highly regular or periodic action sequences observed across many user trajectories are "orbits." An orbit might be a sequence of articles which always have a tendency to be edited in succession. For example, a user responding to a vandal would first revert the damage to the article itself, warn the user on his talk page, and finally notify administrators on a notice board to take action against the vandal. These types of orbits capture organizational routines, many of which have been automated within Wikipedia (Geiger and Ribes 2010).

Researchers can employ the sociotechnical trajectories of users to not only understand social roles as I did here but also to examine organizational routines that generate credibility, behavioral patterns that lead to more reliable user-generated content, and emergence of leadership within self-organizing systems. Trajectories were only computed for four out of the hundreds of users who contributed to these articles, but trajectories could also be computed and compared across all these editors as well to look for similarities in their behavioral patterns.

This type of comparative analysis could begin to unpack whether particular types of sequences or structures are associated with editors becoming socialized into the community and learning to making valuable and high-quality contributions. Take for example an editor who wants to add new information across many articles. This editor could make the changes herself, editing each article individually and creating a "chain" within her sociotechnical trajectory. But these changes may also lack consensus within the community and lead to them being reverted and her then having to make appeals on discussion boards *afterwards* for others to adopt the changes. This would manifest as a high number of "collisions" with other editors across articles. Alternatively, we might imagine her canvassing editors and discussion boards ahead of time to develop consensus, creating a dense web of connections in her trajectory rather than a chain as she diplomatically shuttles between them. This pattern of collaboration might lead to higher quality edits that are more accepted by the community or may mobilize other editors to make the changes themselves. This thought experiment thus also documents behavioral patterns that lead to more reliable user-generated content and the emergence of a leader within a self-organized system.

Researchers also might employ user trajectories to understand the dispositions and evolution of behavioral patterns that predict being elected to administrative roles.

Wikipedia administrators, for example, are granted a variety of tools that allow them to delete pages, ban editors, or protect articles from being edited after passing through an intensive screening process. Comparing the trajectories of these editors may reveal similarities in their behavior as they migrate toward particular editing patterns around antivandalism efforts or new content monitoring. For example, behavioral regularities in reporting vandalism might involve reverting changes on the vandalised page, warning the responsible user on her talk page, and then notifying other users on administrative notice boards that would lead to characteristic cycles in a user's trajectory of moving from articles to user talk pages to administrative boards repeatedly. Users' trajectories that are characterized by high levels of cyclicality and reciprocity (consider again the example in Fig. 4.1) demonstrate higher levels of repeat engagement and monitoring of articles. Thus the user's trajectory capturing the "velocity" of edits and number of "orbits" can serve as a proxy for her commitment and may forecast her effectiveness as a potential administrator.

The sociotechnical trajectory method outlined here opens up new domains for inquiry into latent relationships that have been heretofore ignored in previous network analyses of Wikipedia. More than graphs of who edited what, these trajectories can be read as a narrative of editors inhabiting, discarding, and sampling different social identities over their history. But more than inhabiting a particular social role, the differences between trajectories may also reveal the extent to which authors protect valuable content that does not require them to litigate it in other forums and forecast their leadership and influence as they actively move between domains within the system. Thus, sociotechnical trajectories allow the researcher to mix quantitative metrics for sampling or deductive inference with qualitative interpretations for contextualization and inductive inference, making them superlative tools for mixed methods research.

Conclusion

Wikipedia's coverage of breaking news events challenges traditional theoretical conceptions of organizational behavior and social roles. Despite being a radically open platform for participation that attracts hundreds of editors with mixed motives and expertise, the resulting articles are nevertheless exemplars of timeliness, depth, and style. Drawing on theories of both social roles in online communities as well as high-tempo organizing, this analysis examined whether the most active editors of articles related to a breaking news event performed social roles characterized by a regeneration of prior structural forms or improvisation of new interactional forms. Examination of several prominent editors' sociotechnical trajectories revealed that few possessed expertise specific to editing breaking news articles. However, these editors' histories revealed editors migrated very credible local reputations from other domains to these breaking articles. Editors improvised on their prior social roles as dispute mediators or experts in Japanese culture and emerged as central coordinators—sometimes even leaders—in the efforts to coordinate work on breaking news articles. These findings suggest that rather than demanding explicit credentials to engage in some types of knowledge work or occupy certain social roles, editors focus on the task and trust each other to leverage their existing competencies or adapt to the needs at hand.

Acknowledgements I want to thank Darren Gergle, Noshir Contractor, and Daniel Gruber for feedback on earlier versions of this chapter.

References

Bakker, R. M. (2010). Taking stock of temporary organizational forms: A systematic review and research agenda. *International Journal of Management Reviews, 12*(4), 466–486.

Bao, P., Hecht, B., Carton, S., Quaderi, M., Horn, M., & Gergle, D. (2012). Omnipedia: bridging the Wikipedia language gap. Proceedings of the 2012 ACM annual conference on human factors in computing systems, pp. 1075–1084. New York: ACM. http://doi.acm.org/10.1145/2208516.2208553. doi:10.1145/2208516.2208553.

Bechky, B. A. (2006). Gaffers, gofers, and grips: Role-based coordination in temporary organizations. *Organization Science, 17*(1), 3–21.

Bechky, B. A., & Okhuysen, G. A. (2011). Expecting the unexpected? How swat officers and film crews handle surprises. *Academy of Management Journal, 54*(2), 239–261.

Berkowitz, D. (1992). Non-routine news and newswork: Exploring a what-a-story. *Journal of Communication, 42*(1), 82–94.

Biddle, B. J. (1986). Recent development in role theory. *Annual review of sociology, 12,* 67–92.

Brandes, U., Kenis, P., Lerner, J., & van Raaij, D. (2009). *Network analysis of collaboration structure in Wikipedia*. Proceedings of the 18th international conference on World wide web, pp. 731–740. New York: ACM. http://doi.acm.org/10.1145/1526709.1526808. doi:10.1145/1526709.1526808.

Buriol, L. S., Castillo, C., Donato, D., Leonardi, S., & Millozzi, S. (2006). Temporal analysis of the Wikigraph. IEEE/WIC/ACM international conference on web intelligence, pp. 45–51.

Callero, P. L. (1994). From role-playing to role-using: Understanding role as resource. *Social Psychology Quarterly, 57,* 228–243.

Capocci, A., Servedio, V. D. P., Colaiori, F., Buriol, L. S., Donato, D., Leonardi, S., & Caldarelli, G. (2006). Preferential attachment in the growth of social networks: The internet encyclopedia Wikipedia. *Physical Review E, 74*(3), 036116.

Faraj, S., & Xiao, Y. (2006). Coordination in fast-response organizations. *Management science, 52*(8), 1155–1169.

Gaved, M., Heath, T., & Eisenstadt, M. (2006). Wikis of locality: Insights from the open guides. Proceedings of the 2006 international symposium on Wikis, pp. 119–126. New York: ACM. http://doi.acm.org/10.1145/1149453.1149475. doi:10.1145/1149453.1149475.

Geiger, R. S., & Ribes, D. (2010). The work of sustaining order in Wikipedia: The banning of a vandal. Proceedings of the 2010 ACM conference on computer supported cooperative work, pp. 117–126. New York: ACM. http://doi.acm.org/10.1145/1718918.1718941. doi:10.1145/1718918.1718941.

Gleave, E., Welser, H. T., Lento, T. M., & Smith, M. A. (2009). *A conceptual and operational definition of 'social role' in online community*. System Sciences, 2009. HICSS'09 42nd Hawaii international conference on, pp. 1–11.

Hecht, B., & Gergle, D. (2010). The tower of babel meets web 2.0: User-generated content and its applications in a multilingual context. Proceedings of the SIGCHI conference on human factors in computing systems, pp. 291–300. New York: ACM. http://doi.acm.org/10.1145/1753326.1753370. doi:10.1145/1753326.1753370.

Hu, M., Lim, E.-P., Sun, A., Lauw, H. W., & Vuong, B.-Q. (2007). Measuring article quality in Wikipedia: Models and evaluation. Proceedings of the sixteenth ACM conference on information and knowledge management, pp. 243–252. New York: ACM. http://doi.acm.org/10.1145/1321440.1321476. doi:10.1145/1321440.1321476.

Iba, T., Nemoto, K., Peters, B., & Gloor, P. A. (2010). Analyzing the creative editing behavior of Wikipedia editors: Through dynamic social network analysis. *Procedia-Social and Behavioral Sciences, 2*(4), 6441–6456.

Jesus, R., Schwartz, M., & Lehmann, S. (2009). Bipartite networks of Wikipedia's articles and authors: A meso-level approach. Proceedings of the 5th international symposium on Wikis and open collaboration, p. 5.

Kamps, J., & Koolen, M. (2009). Is Wikipedia link structure different? Proceedings of the second ACM international conference on Web search and data mining, pp. 232–241. New York: ACM. http://doi.acm.org/10.1145/1498759.1498831. doi:10.1145/1498759.1498831.

Kane, G. C. (2009). It's a network, not an encyclopedia: A social network perspective on Wikipedia collaboration. *Academy of Management Proceedings, 2009,* 1–6.

Kane, G. C., Majchrzak, A., Johnson, J., & Chenisern, L. (2009). A longitudinal model of perspective making and perspective taking within fluid online collectives. ICIS, Phoenix, p. 10.

Keegan, B. C. (2013). A history of newswork on Wikipedia. Proceedings of the 9th international symposium on open collaboration, pp. 7:1–7:10. New York: ACM. http://doi.acm.org/10.1145/2491055.2491062. doi:10.1145/2491055.2491062.

Keegan, B. C., Gergle, D., & Contractor, N. (2011a). Hot off the Wiki: Dynamics, practices, and structures in Wikipedia's coverage of the Tōhoku catastrophes. Proceedings of the 7th international symposium on Wikis and open collaboration, pp. 105–113. New York: ACM. http://doi.acm.org/10.1145/2038558.2038577. doi:10.1145/2038558.2038577.

Keegan, B. C., Gergle, D., & Contractor, N. (2011b). Hot off the Wiki: Dynamics, practices, and structures in Wikipedia's coverage of the tōhoku catastrophes. Proceedings of the 7th international symposium on Wikis and open collaboration, pp. 105–113. New York: ACM. http://doi.acm.org/10.1145/2038558.2038577. doi:10.1145/2038558.2038577.

Keegan, B. C., Gergle, D., & Contractor, N. (2012a). Do editors or articles drive collaboration?: Multilevel statistical network analysis of Wikipedia coauthorship. Proceedings of the ACM 2012 conference on computer supported cooperative work, pp. 427–436.

Keegan, B. C., Gergle, D., & Contractor, N. (2012b). Staying in the loop: Structure and dynamics of Wikipedia's breaking news collaborations. Proceedings of the 8th international symposium on Wikis and open collaboration.

Keegan, B. C., Ceni, A., & Smith, M. A. (2013a). Analyzing multi-dimensional networks within mediawikis. Proceedings of the 9th international symposium on open collaboration, pp. 1:1–1:10. New York: ACM. http://doi.acm.org/10.1145/2491055.2491056. doi:10.1145/2491055.2491056.

Keegan, B. C., Gergle, D., & Contractor, N. (2013b). Hot off the Wiki structures and dynamics of Wikipedia's coverage of breaking news events. *American Behavioral Scientist, 57*(5), 595–622.

Kittur, A., & Kraut, R. E. (2008). Harnessing the wisdom of crowds in Wikipedia: Quality through coordination. Proceedings of the 2008 ACM conference on computer supported cooperative work, pp. 37–46. New York: ACM. http://doi.acm.org/10.1145/1460563.1460572. doi:10.1145/1460563.1460572.

Klein, K. J., Ziegert, J. C., Knight, A. P., & Xiao, Y. (2006). Dynamic delegation: Shared, hierarchical, and deindividualized leadership in extreme action teams. *Administrative Science Quarterly, 51*(4), 590–621.

Laniado, D., & Tasso, R. (2011). Co-authorship 2.0: Patterns of collaboration in Wikipedia. Proceedings of the 22nd ACM conference on hypertext and hypermedia, pp. 201–210. New York: ACM. http://doi.acm.org/10.1145/1995966.1995994. doi:10.1145/1995966.1995994.

Laniado, D., Tasso, R., Volkovich, Y., & Kaltenbrunner, A. (2011). When the Wikipedians talk: Network and tree structure of Wikipedia discussion pages. ICWSM'11-Proceedings of the fifth international AAAI conference on weblogs and social media.

Leskovec, J., Huttenlocher, D., & Kleinberg, J. (2010). Signed networks in social media. Proceedings of the SIGCHI conference on human factors in computing systems, pp. 1361–1370. New York: ACM. http://doi.acm.org/10.1145/1753326.1753532. doi:10.1145/1753326.1753532.

Majchrzak, A., Jarvenpaa, S. L., & Hollingshead, A. B. (2007). Coordinating expertise among emergent groups responding to disasters. *Organization Science, 18*(1), 147–161.

Massa, P. (2011). Social networks of Wikipedia. Proceedings of the 22nd ACM conference on hypertext and hypermedia, pp. 221–230.

Meyerson, D., Weick, K. E., & Kramer, R. M. (1996). Swift trust and temporary groups. *Trust in organizations: Frontiers of theory and research, 166,* 195.

Quarantelli, E. L., & Dynes, R. R. (1977). Response to social crisis and disaster. *Annual review of sociology, 3,* 23–49.

Ransbotham, S., Kane, G. C., & Lurie, N. H. (2012). Network characteristics and the value of collaborative user-generated content. *Marketing Science, 31*(3), 387–405.

Ratcheva, V., & Simpson, M. (2011). Temporary organisations: Implications for knowledge work. Available at SSRN 1853546.

Scripps, J., Tan, P.-N., & Esfahanian, A.-H. (2009). Measuring the effects of preprocessing decisions and network forces in dynamic network analysis. Proceedings of the 15th ACM SIGKDD international conference on knowledge discovery and data mining, pp. 747–756. New York: ACM. http://doi.acm.org/10.1145/1557019.1557102. doi:10.1145/1557019.1557102.

Solnit, R. (2010). *A paradise built in hell: The extraordinary communities that arise in disaster*. New York: Penguin Books.

Turek, P., Wierzbicki, A., Nielek, R., Hupa, A., & Datta, A. (2010). Learning about the quality of teamwork from Wikiteams. Proceedings of the second international conference on social computing (SocialCom), pp. 17–24. doi:10.1109/SocialCom.2010.13.

Weick, K. E., & Roberts, K. H. (1993). Collective mind in organizations: Heedful interrelating on flight decks. *Administrative science quarterly, 38,* 357–381.

Welser, H. T., Gleave, E., Fisher, D., & Smith, M. (2007). Visualizing the signatures of social roles in online discussion groups. *Journal of social structure, 8*(2), 1–32.

Welser, H. T., Cosley, D., Kossinets, G., Lin, A., Dokshin, F., Gay, G., & Smith, M. (2011). Finding social roles in Wikipedia. Proceedings of the 2011 iconference, pp. 122–129. New York: ACM. http://doi.acm.org/10.1145/1940761.1940778. doi:10.1145/1940761.1940778.

West, R., Precup, D., & Pineau, J. (2009). Completing Wikipedia's hyperlink structure through dimensionality reduction. Proceedings of the 18th ACM conference on information and knowledge management, pp. 1097–1106. New York: ACM. http://doi.acm.org/10.1145/1645953.1646093. doi:10.1145/1645953.1646093.

Wilkinson, D. M., & Huberman, B. A. (2007). Cooperation and quality in Wikipedia. Proceedings of the 2007 international symposium on Wikis, pp. 157–164. New York: ACM. http://doi.acm.org/10.1145/1296951.1296968. doi:10.1145/1296951.1296968.

Yates, D., Wagner, C., & Majchrzak, A. (2010). Factors affecting shapers of organizational Wikis. *Journal of the American Society for Information Science and Technology, 61*(3), 543–554.

Chapter 5
Words and Networks: How Reliable Are Network Data Constructed from Text Data?

Jana Diesner

Introduction

Social network data as well as the information produced or shared by network participants are prominent sources for studying reputation and authority in social media. Research studies on this topic often start with one or more network datasets and bring relevant substantive questions about socio-technical concepts such as the evolution of credibility to the data. This chapter deals with the reliability of network data itself and aims to shed some light on the following question: How reliable or accurate are network data depending on the data construction method for cases where text data are used as an input to this process? I provide a concise overview on some of the most common methods for constructing network data from text data sources, report on our findings from applying these methods to three corpora from different domains and genres, and derive implications and suggestions for theoretical and practical work.

Basically, network data can be collected or constructed in two ways: First, it might be explicitly available. For example, based on information about network participants, i.e. individuals or organizations who get represented as nodes in a graph, their connections, e.g. other social agents who they have friended or whose content they have commented on or replied to, and the content that network members provide or disseminate, such as their posts and tweets. In this case, existing application programming interfaces (APIs) and tools can be used to download and prepare these network data for analysis. For example, Facebook, Twitter, and YouTube provide such APIs, and network analysis tools such as NodeXL (Hansen et al. 2010) and ConText (Diesner et al. 2013) provide respective data import options.

Alternatively, network data can be constructed or inferred from textual data and metadata that are generated, authored, or disseminated by network participants. These data typically occur in the form of semi-structured or unstructured natural language text data (Corman et al. 2002; Danowski 1993; Diesner and Carley 2005).

J. Diesner (✉)
GSLIS/UIUC, 501 E Daniel Str, 61820 Champaign, IL, USA
e-mail: jdiesner@illinois.edu

E. Bertino, S. A. Matei (eds.), *Roles, Trust, and Reputation in Social Media Knowledge Markets,* Computational Social Sciences, DOI 10.1007/978-3-319-05467-4_5,

In computing, this process is also known as relation extraction (Bunescu and Mooney 2005; Culotta et al. 2006; Roth and Yih 2002).

Besides distilling network data from text data, text data can also be used to enhance explicitly given social network data with the information authored or disseminated by network members. This can be done, for instance, by linking nodes representing agents to nodes representing highly salient information associated with these agents. The resulting networks are typically referred to as socio-semantic networks (Diesner 2012; Gloor and Zhao 2006; Roth and Cointet 2010). One of the main advantages of considering text data for network analysis is that this approach allows for studying the interplay and coevolution of information and social networks. This includes the transformative role that language can play in networks and vice versa (Milroy 1987).

Overall, constructing or enhancing network data based on text data involves a plethora of decisions that have to be made. For example, how to identify nodes and linking them into edges. These decisions can majorly impact the understanding that end users gain about a network and any conclusions they draw from that. The problem here is that the impact of these choices on the resulting relational data is insufficiently understood. This chapter focuses on the different views of a network that one can get when using different relation extraction methods. Who cares about this knowledge? I argue that an empirically grounded understanding of the impact of choices made for text analysis on the derived networks structures contributes to an improved comparability and generalizability of respective methods and tools. Furthermore, such knowledge helps researchers and practitioners to draw valid and reasonable conclusions from analysis results. This is particularly important in cases when validating network data against ground truth data is hard to infeasible, e.g. in the case of covert or historic networks.

From Words to Networks: Methods for Constructing Network Data from Text Data

Network Construction Based on Text Data

In the (computational) social sciences and (digital) humanities, textual data are often converted or coded into networks by developing and applying a codebook (Abello et al. 2012; Gerner et al. 1994; Roberts 1997). Codebooks contain rules for translating relevant pieces of text data into code. These codes represent relevant categories for studying a certain topic, domain, or corpus. Applicable categories can be identified in a top-down fashion from theory and/or in a bottom-up or empirical fashion from the underlying data (Bernard and Ryan 1998; Glaser and Strauss 1967). Node classes can also serve as codes, e.g. “agents”, “organizations” and “locations” (Diesner and Carley 2008). Multi-columned tables that associate text terms with codes are also referred to as thesauri or dictionaries. Traditionally, codebooks and thesauri were created in a manual or semi-automated fashion (Bernard and Ryan 1998), which allows for incorporating human expertise, manual verification of the term to code assignments, and the creation of a controlled vocabulary at the cost of scalability and generalizability (Diesner 2012). Alternatively, techniques from natural language processing and/or machine learning can be applied to create codebooks and thesauri (Cohen and Sarawagi 2004; Diesner and Carley 2008; Roth and Yih 2002), which enable the efficient coding of vast amounts of text data sources (Abello et al. 2012).

The identified instances of relevant entity classes can further be used as nodes for constructing networks. Common approaches for linking nodes into edges rely on (a mixture of) co-occurrence or proximity as well as semantic, syntactic, and statistical features of the text data. While proximity-based approaches have been criticized for their arbitrariness (Corman et al. 2002) and potentially high ratio of false positives (Diesner 2012), it is the most common technique for linking codes or nodes into edges. Technically speaking, proximity-based node linkages result in association networks; a very common type of relational structures extracted from text data. The considered node classes determine the type of network that gets constructed: for example, when identifying social agents (people and organizations) from text data, the resulting graphs represent social networks. When retrieving instances of knowledge and information and the connections between them, the resulting networks can represent semantic networks (Diesner and Carley 2011; Woods 1975). We are taking a more humble approach herein by referring to networks where nodes represent instances of knowledge and information referenced in the text data as knowledge networks.

Network Construction Based on Metadata

While codebook applications operate on the content level, metadata associated with text corpora can serve as another or supplemental source of information for constructing network data. For example, when using LexisNexis—a provider of large collections of data from various sources and genres—to search for documents, the retrieved articles can be downloaded along with metadata. These metadata concisely index the content of the underlying text bodies along various categories. For the case of news wire data, for example, these categories entail “person” and “organization” (social agents), “geographic” (locations), and “subject” (themes). Furthermore, in LexisNexis, each metadata entry is associated with a relevance score that indicates the strength of the association of an article with an index term. Resembling the idea of proximity-based link formation as discussed above, indexed keywords can be linked into edges if they co-occur for the same article. The link weight can be increased accordingly when the same pair of index terms is observed for multiple articles. Another prominent source for metadata are keywords for research proposals and publications that authors select when submitting a paper. Such keywords can be based on a predefined catalogue of eligible terms (controlled vocabulary) and/or identified by the authors given the content of their documents.

Building (multi-modal) network data from metadata is a highly efficient process: Once the metadata are organized, e.g. in a database, the network construction process becomes basically a search and retrieval routine. The ConText software for example supports the construction of metadata databases from previously downloaded LexisNexis files, and the construction of one- and multi-modal network data from these databases (Diesner et al. 2013). The limitation with this approach is that the assignment of metadata entries and relevance scores to articles is not always transparent. For LexisNexis, for example, there is no publicly available documentation on the algorithms or methods used for this process.

Ground Truth Network Data

One way to assess the accuracy of relation extraction techniques and network construction based on text data and metadata is to compare the obtained results against ground truth data, which are also referred to as gold standard data. Ground truth data are typically generated by humans who are specifically trained for this task. Humans can construct ground truth network data in two ways: first, by performing relation extraction based on some text corpus by hand, typically in a computer-supported fashion, and second, by denoting network data to the best of their knowledge, generally also in some computer-assisted way. Both processes are assumed to result in reliable or validated data at the expense of costs and scalability. In other words, given the time-consuming nature of this process, it is often not possible to generate ground truth data for a large-scale dataset or networks. This fact hinders the validation of relation extraction techniques, including the evaluation of the performance of prediction models beyond accuracy rates (Diesner 2012).

Overall, the whole process of going from texts to networks and validating the resulting data is only needed or applicable if one cannot ask network members directly about their relationships or their views of a network (Krackhardt 1987). This applies, for instance, to the case of hidden or historic networks (Diesner and Carley 2005; Sparrow 1991).

Problem Statements

Given these different approaches to network construction and validation, the following research questions with high impact for practical applications are eminent yet heavily under-researched: First, given a corpus, how closely do the results from various content-based and metadata-based network construction techniques resemble ground truth data? And second, how do the outcomes of these methods compare to each other? In other words, what different views of a network do we gain when choosing one method over another? We have conducted several of these comparisons in a series of empirical experiments and report on our findings in the results section (Diesner 2012).

Data

We used three datasets for our analyses: First, our curated version of the Enron email dataset (herein referred to as Enron). This particular version contains 58,266 emails from employees of the former Enron corporation (Diesner et al. 2005). Second, a corpus of news articles about the Sudan (herein referred to as Sudan). This corpus is a curated collection of 79,388 news wire articles released between 2003 and 2008 about the Sudan (Diesner 2012). We collected these data from LexisNexis. Third, a corpus of 55,972 proposals accepted for funding through the European Framework Programmes between 1988 and 2010 (herein referred to as Funding) (Diesner 2012).

While these datasets differ with respect to genre (social media, news articles, scientific writing), domain (business, politics, science), target audience (from internal or private to public), and time span, they are comparable in that they entail text bodies plus metadata: For Enron, we used the email bodies and social agents denoted in the email headers. For Sudan, we worked with the content of the articles and the index terms assigned by LexisNexis. For Funding, we used the project title plus description and predefined index terms selected by the people who submitted the proposals. For details on these data see also Diesner (2012).

Methods

For extracting network data from text data, we built codebooks and thesauri, applied them to the text data, and linked any matches based on their proximity (for details see Diesner 2012). For each dataset, two different thesauri were constructed, which enables the comparison of the impact of different approaches to this step:

First, we used text mining techniques to identify salient terms, e.g. based on (weighted) term frequency metrics, and leveraged existing external and internal dictionaries. We manually verified, consolidated, and disambiguated every entry. This process took between two days (Enron) and six weeks (Sudan) where the time costs mainly depend on the quality and compatibility of leveraged existing material. I refer to this process as relation extraction based on classic codebook construction (CCC). For all three datasets, we aggregated networks per year (Sudan), funding period (funding) and stages of the organizational crisis (Enron) into cumulative graphs per time chunk. The same procedure was also used for the next two methods.

Second, we ran prediction models for entity extraction on each corpus. I refer to this process as entity extraction-based codebook construction (EECC). We had built these models by using conditional random fields, a supervised machine learning technique particularly suited for learning from sparse, sequential data where it is highly beneficial to exploit long-range dependencies (Diesner 2012). Our models go beyond the classic set of named entities (people, organizations, locations) by also detecting other entity classes that are relevant for modeling socio-technical systems, such as resources, tasks, events, knowledge, and attributes, as well as instances of entities that are referred to by a name (e.g. Barack Obama) or not (e.g. politician).

While the models achieved accuracy rates (F scores) of 87.5–88.8 % during the k-fold cross-validation of the machine learning process, applying them to our datasets and again manually verifying their fitness showed that thesauri built this way also need some post-processing in the form of reference resolution and cleaning. Still, the EECC approach outperforms the alternative CCC process in terms of time costs, with this process taking seconds to a few minutes for generating a thesaurus per corpus and up to 2 days for post-processing it. Moreover, the prediction models generalize with known accuracy while a thesaurus built in the classic way for one dataset cannot be assumed to generalize well to corpora from other domains, genres, or points in time due to the deterministic nature of thesauri.

For constructing metadata networks, for Sudan, we linked any two entities occurring in the metadata that represent people, organizations, locations, or knowledge per article into bidirectional, weighted graphs. The weights were identified by computing the average of the lowest-relevance scores for any two linked entities. For Funding, we coded all index terms as knowledge and linked any such pairs per proposal into edges. For Enron, we connected senders and receivers (to, cc, bcc) into directed social networks that were weighted by the cumulative frequency per entity pair. Note that this approach defines a classic, explicitly given social network; the way it is often constructed from social media data. The resulting network can then be compared against social networks extracted from the text data. Each of these operations was a matter of minutes once we had curated the data and organized them in relational databases.

As for ground truth networks generated by human experts, we were only able to construct such data for Sudan. This was possible through a collaboration with Dr. Richard Lobban, a leading expert on the Sudan, and his team. More specifically, we went through a qualitative, computer-supported, iterative process of building expert-verified networks of tribal affiliations in the Sudan for each calendar year of 2003–2008. We started by applying a list of all tribes in the Sudan, which was provided by Dr. Lobban's team, to our Sudan corpus, creating a first visualization of the tribal network and sending that to Dr. Lobban for verification, i.e. annotating false positives and false negatives in terms of nodes and edges. Once we received their modified maps, we adjusted our coding scheme and regenerated the network data. We repeated this process until Dr. Lobban's teams assessed the networks as representative of the ground truth based on their expertise. The time costs for this process are comparable to building codebooks without leveraging machine learning methods. Since this process cannot be expected to scale up, it can only be used for small to moderately sized networks.

Once we had constructed these networks, we compared them within and across datasets and methods. More specifically, we identified the structural overlap of nodes based on their node names or labels and the edges between them.

Results and Conclusions

How much do network data constructed from text data or metadata resemble ground truth data? It depends, but overall very little, as our results suggest: Out of the social network data built in collaboration with subject matter experts, 53 % of the nodes and 20 % of the links also appeared in networks distilled from text bodies when using the classic thesaurus construction (CCC) approach. These values drop to 11 % for nodes and 5 % for edges for relation extraction based on automatically built thesauri (EECC), and to flat zeros for metadata-based networks. What accounts for these differences? The main reason for the overlap between networks based on ground truth data and the CCC method is that we reused the same list of tribes as input for the human experts and thesaurus construction while the EECC method finds any applicable matches purely based on the underlying machine learning techniques. We observed the same effect for other methods: CCC-based networks resemble metadata networks more closely than the EECC-based networks, mainly because we enhanced the classic thesauri with data that we also used for defining nodes for metadata networks, e.g. lists of index terms. At the same time, EECC-based networks and metadata networks are constructed from different data, namely text bodies and metadata; with the differences in terminology and scope leading to different views of the networks. In summary, reconstructing social networks by applying text mining techniques to corpora, including metadata, will lead to largely incomplete and biased incomplete results. This limitation could be alleviated by switching from proximity-based node linkage to alternative methods, such as approaches based on syntax, semantics, and machine learning techniques (Roth and Yih 2002; Zelenko et al. 2003). In our studies, the structural agreement between any pair of networks was consistently higher on the node level than on the edge level. This effect might also change by using different node linkage strategies.

Another factor that we observed to strongly impact the agreement in networks structure is the network size: Larger networks have a higher chance to resemble parts of networks constructed with alternative methods that lead to smaller networks, both in number of nodes and edges. This fact is of methodological and practical relevance since various network metrics have been shown to correlate with network size (Anderson et al. 1999; Friedkin 1981).

Comparing networks not on a structural but a substantive level leads to different findings depending on the domain and network construction method: For corpora of news articles, social networks created from metadata feature major international key entities and their connections, while social networks distilled from text bodies provide to a more fine-grained and localized understanding of important actors and their links. In contrast to that, when looking at knowledge networks across the genres and domains considered, text-based networks give a high-level overview on salient terms and their connections for a given domain, while metadata networks drill down to more specific pieces of knowledge and information per domain. The reason for this effect is that the keywords and index terms from which the metadata networks are constructed are already highly condensed and often carefully selected mini-summaries of the underlying text bodies while these concepts are being elaborated on in a more detailed fashion in the actual text bodies. This explanation also partially accounts for another observation: Looking at ambiguity issues in the generated networks, we found that metadata networks are less limited by co-reference resolution issues than methods that operate on the content level. Co-reference resolution here means disambiguating terms with the same surface form but different meaning and consolidating terms with different surface forms but the same meaning.

Synthesizing our findings, we recommend fusing text-based and metadata-based networks in an informed fashion: using machine-learning-based entity extraction to build a thesaurus, refining and enhancing it based on subject matter expertise if available, and linking up nodes based on methods that are more advanced than co-occurrence is best suited for generating social networks. These networks can be combined with knowledge networks derived from metadata. Based on our findings, the resulting networks will allow for a broad and deep look into social and knowledge networks.

Acknowledgements This work was supported by the National Science Foundation (NSF) IGERT 9972762, the Army Research Institute (ARI) W91WAW07C0063, the Army Research Laboratory (ARL/CTA) DAAD19-01-2-0009, the Air Force Office of Scientific Research (AFOSR) MURI FA9550-05-1-0388 and the Office of Naval Research (ONR) MURI N00014-08-11186. Additional support was provided by CASOS, Carnegie Mellon University. The views and conclusions contained in this chapter are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the NSF, ARI, ARL, AFOSR, ONR, or the United States Government.

References

Abello, J., Broadwell, P., & Tangherlini, T. R. (2012). Computational folkloristics. *Communications of the ACM, 55*(7), 60–70.

Anderson, B. S., Butts, C., & Carley, K. M. (1999). The interaction of size and density with graph-level indices. *Social Networks, 21*(3), 239–268.

Bernard, H., & Ryan, G. (1998). Text analysis: Qualitative and quantitative methods. In H. Bernard (Ed.), *Handbook of methods in cultural anthropology* (pp. 595–646). Walnut Creek: Altamire Press.

Bunescu, R., & Mooney, R. (2005). Subsequence kernels for relation extraction. *Advances in Neural Information Processing Systems* (NIPS), Vancouver, BC.

Cohen, W. W., & Sarawagi, S. (2004). *Exploiting dictionaries in named entity extraction: combining semi-Markov extraction processes and data integration methods*. Paper presented at the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Seattle, WA.

Corman, S. R., Kuhn, T., McPhee, R. D., & Dooley, K. J. (2002). Studying complex discursive systems: Centering resonance analysis of communication. *Human Communication Research, 28*(2), 157–206.

Culotta, A., McCallum, A., & Betz, J. (2006). *Integrating probabilistic extraction models and data mining to discover relations and patterns in text*. Paper presented at the Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics. New York, NY.

Danowski, J. A. (1993). Network analysis of message content. *Progress in Communication Sciences, 12,* 198–221.

Diesner, J. (2012). *Uncovering and managing the impact of methodological choices for the computational construction of socio-technical networks from texts*. Carnegie Mellon University. (CMU-ISR-12–101, PhD Thesis).

Diesner, J., & Carley, K. M. (2005). Revealing social structure from texts: Meta-Matrix text analysis as a novel method for network text analysis. In V. K. Narayanan & D. J. Armstrong (Eds.), *Causal mapping for information systems and technology research: Approaches, advances, and illustrations* (pp. 81–108). Harrisburg: Idea Group Publishing.

Diesner, J., & Carley, K. M. (2008). Conditional random fields for entity extraction and ontological text coding. *Journal of Computational and Mathematical Organization Theory, 14,* 248–262.

Diesner, J., & Carley, K. M. (2011). Semantic networks. In G. Barnett & J. G. Golson (Eds.), *Encyclopedia of social networking* (pp. 766–769). Sage.

Diesner, J., Frantz, T. L., & Carley, K. M. (2005). Communication networks from the Enron Email Corpus. It's always about the people. Enron is no different. *Computational Mathematical Organization Theory, 11*(3), 201–228.

Diesner, J., Aleyasen, A., Kim, J., Mishra, S., & Soltani, S. (2013). *Using socio-semantic network analysis for assessing the impact of documentaries*. Paper presented at the WIN (Workshop on Information in Networks). New York, NY.

Friedkin, N. E. (1981). The development of structure in random networks: An analysis of the effects of increasing network density on five measures of structure. *Social Networks, 3*(1), 41–52.

Gerner, D., Schrodt, P., Francisco, R., & Weddle, J. (1994). Machine coding of event data using regional and international sources. *International Studies Quarterly, 38*(1), 91–119.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine.

Gloor, P., & Zhao, Y. (July 2006). *Analyzing actors and their discussion topics by semantic social network analysis*. Paper presented at the 10th IEEE International Conference on Information Visualisation. London, UK.

Hansen, D., Shneiderman, B., & Smith, M. A. (2010). *Analyzing social media networks with NodeXL*. Morgan Kaufmann.

Krackhardt, D. (1987). Cognitive social structures. *Social Networks, 9*(2), 109–134.

Milroy, L. (1987). *Language and social networks* (2nd ed.). Oxford: Blackwell.

Roberts, C. W. (1997). *Text analysis for the social sciences: Methods for drawing statistical inferences from texts and transcripts*. Mahwah: Lawrence Erlbaum Associates, Inc.

Roth, C., & Cointet, J. (2010). Social and semantic coevolution in knowledge networks. *Social Networks, 32*(1), 16–29.

Roth, D., & Yih, W. (2002). *Probabilistic reasoning for entity and relation recognition*. Paper presented at the International Conference on Computational Linguistics (COLING), Taipei, Taiwan.

Sparrow, M. (1991). The application of network analysis to criminal intelligence: An assessment of the prospects. *Social Networks, 13*(3), 251–274.

Woods, W. (1975). What's in a link: Foundations for semantic networks. In D. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science* (pp. 35–82). New York: Academic Press.

Zelenko, D., Aone, C., & Richardella, A. (2003). Kernel methods for relation extraction. *The Journal of Machine Learning Research, 3,* 1083–1106.

Chapter 6
Predicting Low-Quality Wikipedia Articles Using User's Judgements

Ning Zhang, Lingyun Ruan and Luo Si

Introduction

As the largest free and collaboratively edited on-line encyclopedia, Wikipedia has more than 4 million articles and about 120,000 active contributors from all over the world[1]. In the meantime, it faces a great challenge to ensure article's quality because anyone can act as a contributor to create or revise an article. These contributors, who may not be professional editors, have different levels of writing skills and may even be malicious. To claim the criteria of high-quality, a set of *Featured Articles*[2] are voted by Wikipedia editors. An article is selected as *featured article* only if it fulfills a series of quality requirements on accuracy, neutrality, completeness, and style. The criteria are so strict that only 3825 out of 4,191,250 articles are featured on English Wikipedia, which also indicates that most articles still have sufficient space for improvement. To help reduce the workload of selecting featured articles, researchers proposed a problem of automatically judging whether an article was featured and formulated it as a binary classification problem. Different approaches are applied and good results are gained to predict featured articles from non-featured ones (Blumenstock 2008; Wilkinson and Huberman 2007; Hu et al. 2007; Stvilia et al. 2005).

Featured articles give perfect examples of high-quality writings and provide a very good quality measurement. Nevertheless, having all of the non-featured articles

[1] https://en.wikipedia.org/wiki/Wikipedia:Statistics#Page_views.

[2] http://en.wikipedia.org/wiki/Wikipedia:Featured_articles.

N. Zhang (✉)
Department of Computer Science, Purdue University, West Lafayette, IN, USA
e-mail: lemonustc@gmail.com

L. Ruan
Google, Mountain View, CA, USA

L. Si
Department of Computer Science, Purdue University, West Lafayette, IN, USA

E. Bertino, S. A. Matei (eds.), *Roles, Trust, and Reputation in Social Media Knowledge Markets,* Computational Social Sciences, DOI 10.1007/978-3-319-05467-4_6,

revised to achieve this standard is impractical since that it brings about unaffordable workload and that even the most experienced editors cannot guarantee that all of their writings are able to meet those strict quality requirements. To best improve reading experience with limited resources, the community has to start from finding out those low-quality contents which affect the reading experience most significantly and have urgent needs for improvement.

There does not exist a precise definition for “low-quality”, but apparently, not all non-featured articles are low-quality.

Vandalism detection is among those practices to find out low-quality contents, and tools to automatically detect vandalism have been studied for a long time. For instance, Wikipedia has deployed a group of rule-based bots to automatically identify and eliminate vandalism[3]. Also, researchers casted a binary classification problem to do vandalism detection (Potthast et al. 2008; Smets et al. 2008). In Adler et al. (2010, 2011), reputation features based on WikiTrust[4] project are utilized, which builds a reputation (authors and articles) system for Wikipedia.

Another attempt is article flaw prediction. Instead of defining low-quality, researchers aimed at specific quality flaws and developed automatic tools to identify them. Anderka et al. (2012) made use of “cleanup tags” (each kind of tags represents a corresponding quality flaw) to investigate whether an article has quality flaws. They formulated a set of one-class classification problems and achieved reasonable result in predicting common quality flaws.

Vandalism definitely affects reading experience a lot, but it only counts for a small part of low-quality contents. Besides, even though cleanup tags are quite closely related to low-quality, they still cannot explicitly indicate user’s reading experience. As the influences on reading experience differ from flaws, it is hard to tell whether an article’s quality is below user’s expectation, even when we know exactly what flaws it has.

In this paper, we utilize article ratings collected by Wikipedia Article Feedback Tool to judge article quality from user’s point of view. As an user-oriented representation of article quality, article ratings indicate user’s reading experiences and feedbacks toward articles directly and intuitively. We formulate a set of binary classification problems to find out low-quality articles and we believe that this approach provides a promising solution to predict low-quality Wikipedia articles.

The remainder of this paper is organized as follows. The data description is given in “Revision Feature Extraction and Article Rating Processing”. “Low-Quality Revision Prediction” provides the problem formulation. Experiment setting and results evaluation are shown in “Experiment and Result Evaluation”. “Conclusion and Future Work” concludes our work and provides our sight on future work.

[3] http://en.wikipedia.org/wiki/Category:Wikipedia_anti-vandal_bots.
[4] http://www.wikitrust.net/.

Table 6.1 Revision feature table

Feature	Description
Language count	Number of versions in different languages
Category count	Number of categories that the article belongs to
Expression count	Number of mathematic expressions
Picture count	Number of pictures (including tables)
Reference count	Number of referenced resources
External link count	Number of external links (toward pages outside Wikipedia)
Internal link count	Number of internal links (toward pages inside Wikipedia)
Sixth subsection count	Number of sub-sub-sub-sub-sub-sub-sections
Fifth subsection count	Number of sub-sub-sub-sub-sub-sections
Fourth subsection count	Number of sub-sub-sub-sub-sections
Third subsection count	Number of sub-sub-sub-sections
Second subsection count	Number of sub-sub-sections
Section depth	Depth of the section structure
Section length	Section length in average
Words count	Number of words
Word length	Word length in average
Sentence count	Number of sentences
Sentence length	Sentence length in average
Paragraph count	Number of paragraphs
Paragraph length	Paragraph length in average
Author contribution	Number of revisions this author contributed
Registered author	Whether the author is registered: 1 if yes, 0 if not
Author reputation	Average rating this author got over all of his/her contributions
Parent revision rating	Average rating the revision's parent revision got

Revision Feature Extraction and Article Rating Processing

Revision Feature Extraction

When we read an "article" in Wikipedia, we are actually reading the newest version of the article. Ever since the creation of an article, Wikipedia generates a new version for it whenever there comes an editing behavior. A term "revision" is used to represent each of these historical versions. As time goes by, an article may have thousands of different revisions edited by hundreds of different users. Wikipedia keeps all historical revisions and indexes them by "revision ID".

Unlike some other research that treats each "article" (the newest revision) as a data point, we count different revisions of an article as different data points and choose revision as the unit in our experiment. We use a backup of English Wikipedia[5] which was dumped in August 2012 with complete historical revisions. As an article may have thousands of revisions, we parse more than 10 million revisions and randomly sample 1/50 of them. We write a Python parser and extract 22 features (the first 22 features shown in Table 6.1) from the text and meta information of these sampled

[5] http://en.wikipedia.org/wiki/Wikipedia:Database_download.

Table 6.2 Statistic information of ratings

Dimension	Total number of ratings	Average number of ratings revisions got	Average rating each revision got	Average variance of ratings each revision got
1, Trustworthiness	702,308	8.22	3.70	1.75
2, Objectivity	660,388	7.73	3.69	1.71
3, Completeness	685,959	8.03	3.51	1.72
4, Well-written	736,666	8.62	3.78	1.63

revisions. Those features are related to article structure, content, network, and editing history. Pages with less than 30 revisions and revisions that do not have valid contributor information are filtered out.

Article Rating Processing

In order to achieve better quality assessment and to boost reader engagement, Wikipedia launched Article Feedback Tool (AFT) in 2010[6]. From July 2011 to July 2012, the fourth version of AFT was deployed on the entire English Wikipedia and was shown as a "rating this article" box on the bottom of article pages. Users can rate an article (the "current" revision at that specific time) in four dimensions of quality: trustworthiness, objectivity, completeness and well-written, each from one star to five star (the larger, the better). We make use of this 1-year anonymous rating data[7]. Revisions which have less than three user rating records are filtered out. We calculate the average rating and rating variance for each revision in each of the four quality dimensions.

Based on the rating information, we add two important revision features (the last two features shown in Table 6.1). The first One is author reputation, which is indicated by the average rating that the revision's author got over all of the revisions he/she has edited before. The second feature is parent revision rating showing the quality of the former revision on which the current one is based. Every revision must come out from a specific author's editing behavior on a parent revision, so its quality must be closely related to the author's writing skill and the parent revision's quality. As indicators of author's writing skill and parent revision's quality, the above two features will act as very important roles in our experiment.

Combining ratings together with revision features, we get a data set that consists of 85,417 revisions with their features and ratings. Some statistics of the ratings are shown in Table 6.2.

As it can be seen from the table, each revision in our data set has about eight rating records in average. In different quality dimensions, the average ratings range from

[6] http://en.wikipedia.org/wiki/Wikipedia:Article_Feedback_Tool.
[7] http://datahub.io/dataset/wikipedia-article-ratings.

Table 6.3 Percentages of low-quality revisions

Dimension	1, Trustworthiness	2, Objectivity	3, Completeness	4, Well-written
Number of low-quality revisions	6633	6565	10,173	4871
Percentage of low-quality revisions	0.078	0.077	0.119	0.057

3.5 to 3.8. We also find that the rating variances are relatively large as the average variance is about 1.75. We will make further discussion about it in "Experiment and Result Evaluation".

Low-Quality Revision Prediction

Regarding each article revision as a data point, we propose the low-quality article prediction problem as follows: given a revision with extracted feature values, decide whether it is low-quality in each of the four quality dimensions. Here "low-quality" is defined as that its average rating in the specific dimension is lower than a threshold h.

Let R be the set of article revisions (represented as feature vectors) and let Q be the set of four quality dimensions. We ignore the possible correlation between different quality dimensions and aim at finding a classifier c which can solve the following multi-labeling problem:

$$c : \mathbf{R} \to 2^Q,$$

where 2^Q denotes the power set of Q. To keep the problem simplified, we formulate this multi-labeling problem by multiple binary classifications instead of by multi-class classification:

$$c_i : \mathbf{R} \to 0, 1, i = 1, 2, 3, 4$$

where in each of the four dimensions, a revision with feature vector $\mathbf{x}$ is labeled with $t = 1$ if its average user rating is lower than h and with $t = 0$ if its average rating is higher than h.

Experiment and Result Evaluation

In experiment, we heuristically set h to be 2.5 and revisions with average rating lower than 2.5 are labeled as low-quality. The percentages of low-quality revisions (85,417 revisions in total) are shown in Table 6.3. Our data set is quite imbalanced: in the four dimensions, even the highest percentage of positive samples is below 0.12. But it also reveals the nature of low-quality prediction problem: even though most articles

Table 6.4 F1 scores of baseline algorithms

Dimension	Baseline 1 (author reputation)	Baseline 2 (parent revision rating)
1, Trustworthiness	0.0214	0.2290
2, Objectivity	0.0221	0.2247
3, Completeness	0.0457	0.3258
4, Well-written	0.0193	0.2061

are not perfect, only a small part of them are significantly below reader's expectation and are in urgent need for improvement.

We apply Logistic Regression algorithm and compare it with another two simple algorithms on the data set.

As a discriminative classification algorithm, Logistic Regression models the class posterior probability with logistic sigmoid function. In our case, it is the probability of a revision to be low-quality given all its feature values. We apply BFGS method to maximize the regularized log-likelihood function, which is shown as follows:

$$L = \ln P(\mathbf{t}|\mathbf{X}, \mathbf{w}) = \sum_n (t_n \ln y_n + (1 - t_n) \ln (1 - y_n)) - \beta \mathbf{w}^T \mathbf{w},$$

where $y_n = P(t_n = 1|\mathbf{w}, \mathbf{x}_n) = \sigma(\mathbf{w}^T \mathbf{x}_n)$ denotes the posterior probability ($\sigma(\cdot)$ is logistic sigmoid function) and $\mathbf{w}$ is the coefficient vector. Please refer to related books (Hosmer and Lemeshow 2004) for more details about Logistic Regression algorithm.

The two baseline algorithms are derived from the two important features which we mentioned above. The first simple baseline algorithm directly uses author reputation to predict the revision's rating: if its author got an historical average rating which is lower than 2.5, we predict the revision to be low-quality. In the second baseline algorithm, we utilize parent revision rating to make prediction: if a revision's parent revision is low-quality (get an average rating lower than 2.5), we predict it also to be low-quality.

We choose F1 score to measure the results. Since F1 score can be interpreted as a weighted average of precision and recall, it provides us an objective and comprehensive view angle to compare the algorithms. The formal definition of F1 score is as follows:

$$F1 = 2 * \frac{precision * recall}{precision + recall}.$$

We split the data set into two part. Half of it is used for training purpose and the other half serve as testing set.

Table 6.4 contains the results of the two baseline algorithms. The F1 values of the second baseline algorithm are acceptable but those of the first baseline algorithm are surprisingly low. After checking precision values and recall values, we find that the recall of the first baseline algorithm drastically pulls down the F1 value, which

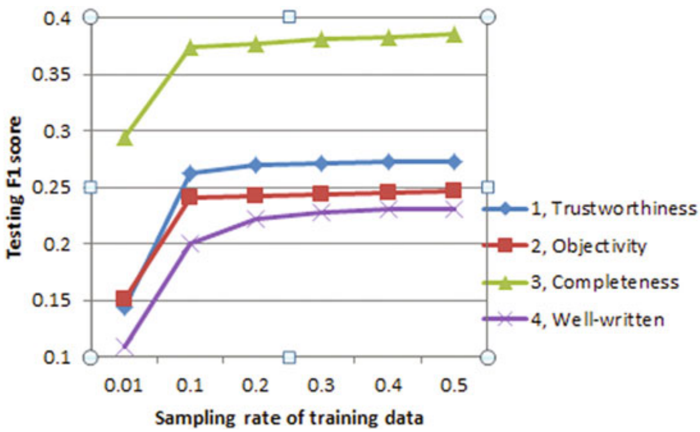


Fig. 6.1 Learning curves of Logistic Regression algorithm

Table 6.5 Testing F1 scores of Logistic Regression algorithm

Sampling rate of training data	0.01	0.10	0.20	0.3	0.40	0.5
1, Trustworthiness	0.1434	0.2630	0.2703	0.2715	0.2720	0.2727
2, Objectivity	0.1510	0.2410	0.2419	0.2429	0.2454	0.2460
3, Completeness	0.2940	0.3743	0.3775	0.3811	0.3831	0.3855
4, Well-written	0.1090	0.2003	0.2224	0.2273	0.2304	0.2312

indicates that bad editors do write some of the low-quality revisions but there are still many other low-quality contents are composed by normal or even good editors. Moreover, the cooperative nature of Wikipedia makes the parent revision's rating more markedly than author's reputation in rating prediction task since editors just modify a relatively small part of the article in most cases.

Compared to the baselines, Logistic Regression algorithm gets more reasonable results. Figure 6.1 shows the learning curves of Logistic Regression algorithm and the testing F1 values are listed in Table 6.5. In the most imbalanced dimension of "well-written" (only 5.7 % of the points are labeled as 1), the F1 score for testing set is about 0.23. In the more balanced dimension of "completeness" (about 12 % of the points are labeled as 1), the F1 score for testing set reaches 0.38. Here we can tell that the distribution of data has distinct influence on Logistic regression algorithm's effectiveness. Moreover, most features we use are article statistics. Compared to the other high-level quality dimensions such as "Trustworthiness" which is harder to be interpreted without understanding of natural language, the dimension of "completeness" is more predictable by those naive features.

As mentioned in "Article Rating Processing", article rating variances are non-negligible. To investigate how rating variances can influence the performance of the learning algorithm, we filter out revisions which has large rating variance and only

Table 6.6 Testing F1 scores after variance filtering

Variance threshold	1	0.5
1, Trustworthiness	0.3733	0.4549
2, Objectivity	0.3248	0.445
3, Completeness	0.5160	0.5937
4, Well-written	0.2656	0.3544

Table 6.7 Number of revisions left

Variance threshold	1	0.5
1, Trustworthiness	27,405	16,268
2, Objectivity	28,343	16,749
3, Completeness	25,941	13,875
4, Well-written	29,567	16,808

keep those revisions whose rating variance is lower than a threshold v. Testing results and number of revisions left after doing such filtering are listed in Tables 6.6 and 6.7.

We can see that the results improve quite a lot after doing variance filtering: the lower the threshold, the better the results. When $v = 0.5$, Logistic Regression algorithm reaches a highest F1 value of 0.5937 in the third quality dimension. Since the data distribution shifts only a little bit after variance filtering, we rule out the possibility that it is caused by higher positive data percentages. In fact, when human readers cannot achieve an agreement about the revision's quality and give the revision very biased ratings, it is difficult for automatic algorithms to judge the quality, too. Correspondingly, by doing variance filtering, we remove the revisions that have too biased ratings and the improvements are expected. However, the filtered data set becomes much smaller and this side-effect brings about a problem to strike a balance between performance and applicability.

Conclusion and Future Work

This paper proposes a novel low-quality Wikipedia article prediction task by making use of Wikipedia reader feedback data. We formulate it as a set of binary classification problems and obtain reasonable predicting results by applying Logistic Regression algorithm and adding variance filtering. Though our results are far from perfection, it provides a new method to automatically identify potential low-quality articles and achieves fairly good effectiveness, especially given the complicated nature of human writing, readers' various tastes, and the extremely imbalanced data distribution.

For future work, at least three problems are noticeable as mentioned earlier. First, features that we currently use are not serviceable enough to interpret the quality dimensions and specific requirements well. Content-based analysis such as Natural Language Processing techniques might help extract more useful features. Second, biased user ratings bring noise into the data and need to be handled properly, either by more advanced technique to process those ratings or more robust learning algorithms.

Finally, the imbalanced data distribution limits the classifier a lot and leaves large space for improvement.

Wikipedia is an important knowledge market. This research work has the potential to build an effective and efficient tool for ensuring the quality of Wikipedia and other knowledge markets by identifying low quality contents.

References

Adler, B., de Alfaro, L., & Pye, I. (2010). Detecting Wikipedia vandalism using WikiTrust. Notebook papers of CLEF, pp. 22–23.

Adler, B. T., de Alfaro, L., Mola-Velasco, S. M., Rosso, P., & West, A. G. (2011). Wikipedia vandalism detection: Combining natural language, metadata, and reputation features. In *Computational linguistics and intelligent text processing*, pp. 277–288. (Springer).

Anderka, M., Stein, B., & Lipka, N. (2012). Predicting quality flaws in user-generated content: The case of Wikipedia. Proceedings of the 35th international ACM SIGIR conference on research and development in information retrieval, pp. 981–990. (ACM).

Blumenstock, J. E. (2008). Size matters: Word count as a measure of quality on Wikipedia. Proceedings of the 17th international conference on World Wide Web, pp. 1095–1096. (ACM).

Hosmer, D. W. and Lemeshow, S. (2004). *Applied logistic regression* (Vol. 354). New York: Wiley.

Hu, M., Lim, E.-P., Sun, A., Lauw, H. W., & Vuong, B.-Q. (2007). Measuring article quality in Wikipedia: Models and evaluation. Proceedings of the sixteenth ACM conference on information and knowledge management, pp. 243–252. (ACM).

Potthast, M., Stein, B., & Gerling, R. (2008). Automatic vandalism detection in Wikipedia. In *Advances in information retrieval*, pp. 663–668. (Springer).

Smets, K., Goethals, B., & Verdonk, B. (2008). Automatic vandalism detection in Wikipedia: Towards a machine learning approach. AAAI workshop on Wikipedia and artificial intelligence: An Evolving Synergy, pp. 43–48.

Stvilia, B., Twidale, M. B., Smith, L. C., & Gasser, L. (2005). Assessing information quality of a community-based encyclopedia. Proceedings of the international conference on information quality, pp. 442–454.

Wilkinson, D. M., & Huberman, B. A. (2007). Cooperation and quality in Wikipedia. Proceedings of the 2007 international symposium on Wikis, pp. 157–164. (ACM).

Part III
Tools for Enhancing Trust and Transparency

Chapter 7
From Invisible Algorithms to Interactive Affordances: Data After the Ideology of Machine Learning

Bernie Hogan

Introduction

Humans are limited. We cannot fly, shoot lasers from our eyes or leap tall buildings in a single bound. Nor can we reasonably comprehend even a fraction of the knowledge about the external world that is presented to us. Unfortunately, solutions to our superhero problems remain in the realm of science fiction. Fortunately, our ability to organize and present information has expanded rapidly since the advent of the digital era. In general, the digital revolution has brought with it the profound new capacities for the *codification, sorting, selection* and *visualization/presentation* of information about the external world. Yet, these capacities neither work nor evolve in synchrony. Technologists, designers and engineers can focus primarily on one or more capacities. Research breakthroughs on one (such as force-directed layouts for visualization, BigTable for selection or collaborative filtering for sorting) do not necessarily alter our capacities for another. To that end, it is possible and plausible that certain capacities are given particular emphasis, while others remain understudied and poorly implemented. When one capacity, such as our capacity to visualize information, is seen as particularly important or self-evident, work on this capacity is driven by ideology.

In this chapter, I posit that the dominant ideology of information management is one of sorting, especially personalized or relevance-based sorting, and infused with faith in machine learning. This dominance is concerning because it is based on invisible algorithms. That is, the specific algorithm used to rank order a series of elements, whether it is content from friends, web pages or goods, includes many elements that are either unknown to the person consuming the information, or worse, unknowable. The triumph of such invisible algorithms poses serious challenges to any study of reputation or credibility. We may intuitively accept that a certain ordering "makes sense", but without an ability to assess this ordering we are at the mercy of

B. Hogan (✉)
Oxford Internet Institute, University of Oxford,
1 St Giles, Oxford OX1 3JS, UK
e-mail: Bernie.hogan@oii.ox.ac.uk

E. Bertino, S. A. Matei (eds.), *Roles, Trust, and Reputation in Social Media Knowledge Markets,* Computational Social Sciences, DOI 10.1007/978-3-319-05467-4_7,

those building the algorithms. Worse, to the extent that we consider this ideology as necessary, we restrict our ability to imagine alternative means for the management of information and concede that the judgment of the algorithm designers is inherently better than our own.

I contend that a focus on invisible algorithms is not mere necessity. It is partially a product of a historical configuration of technologies present at this point in time. At the moment, data are messy, processors are fast, and "Big Data" means parallelizable algorithms that reduce large columns of data into discernible metrics. Also, at the moment, browser toolkits for visualization are evolving slowly while still placing significant demands on the personal computer; graph databases are emerging as an alternative and increasingly scalable means for organizing and querying data; and the web of linked data is becoming an increasingly plausible and coherent future for the Internet. Yet, a focus on invisible algorithms permeates computer science from the most basic foundational courses to some of the most sophisticated technological offerings of our age.

I begin this chapter with a definition, overview and brief discussion of invisible algorithms. I then briefly discuss alternative approaches that focus on interactive affordances via selection and visualization. I give an overview of three domains where both approaches have been applied: music, email and friending. I conclude with a reflection on why invisible algorithms make practical sense in many cases at the present but ultimately remain limited by the very features that have made them so useful: single column ordering, the reduction of highly dimensional data and simplicity for the end user.

Defining the Invisible Algorithm

An algorithm, at its most basic, is a mechanism for taking in some form of data and outputting some other form of data. It is not dissimilar to a mathematical function, $f(x) = y$. That is, an algorithm is a finite, discrete series of instructions that receive an input and produce an output. The earliest documented algorithms by al-Khwārizmī (approximately 800 CE) were designed to solve linear and quadratic equations. In the twentieth century, the concept was given its most thorough treatment when Alan Turing described what is now known as the universal Turing machine (1936). One of the charms of the Turing machine is that it describes a minimalistic set of ingredients to which all algorithms must adhere. If a function can be computed, it can be computed using a Turing machine (although not all functions can be computed) and vice versa. Turing machines are not actual machines, of course, but a minimal set of capacities needed to sequentially arrive at a solution using a specific input. What is important for this discussion is that they can be treated as "black boxes". As such, one does not need to know how the computation is performed in order to use it. It is sufficient to know that an algorithm receives input and presents an output.

The notion of an algorithm as a black box permeates computer science. I recall an early exposure to this notion in my second computer science course as the "precondition-postcondition contract", stating that an algorithm will give the correct output as long as it receives the correct form of input (Main 2002). We were facetiously supposed to "sign" this contract as symbolizing our focus on creating reliable modular code. The black box idea embeds itself in some of computer science's most vexing thought experiments, such as Searle's Chinese Room. Particularly in object-oriented programming, good code is modular code that merely assumes other algorithms and packages produce output as specified. The internals of any object or function need not, and typically should not, be accessible to external programs.

That an algorithm itself can be a black box is a source of consternation for scientists, policy makers, lawyers, marketers and hackers. Is a specific cryptography algorithm really unbreakable? Are search results unnecessarily privileging a particular product or excluding a specific term? How can I increase my ranking in Google's search results? If the algorithm is a black box, it makes the answer to many such questions notably challenging. Unfortunately, herein lies an important distinction that has previously gone unmentioned: Some of these questions concern an algorithm because it is a black box, but others are more challenging—they concern an algorithm because it is invisible.

Any algorithm can be treated as a black box where one is interested in the output of that algorithm and not its instruction set. An *invisible algorithm* is one where both the instruction set of the algorithm *and at least some of the inputs* to the algorithm are designed to be hidden from the recipient of the output. In this regard, the question about whether a cryptographic algorithm is unbreakable is actually qualitatively different from a question about increasing one's search results. If one knew the cryptographic algorithm, a user could follow along from the input to arrive at the output. If one knew Google's ranking algorithm and followed along from what is, ostensibly, the input, the output is still unclear. For example, when I go to a web page that computes a trigonometry function, from my point of view, I am using a black box algorithm. When I go to a search engine to look for "trigonometry", from my point of view, I am using an invisible algorithm. In the calculation case, all the inputs it needs come from me or are available to me (sometimes as "defaults"). In the latter case, the algorithm ranks web pages using both the search term "trigonometry calculator" and a host of other hidden factors to define what is presented. Is it using my location? My grammar? My past history?

Presently, it is said that Google's search rankings include approximate personalization signals (Pariser 2011). The obvious criticism is that it is not known which signals it uses, thereby leading to certain pages being unnecessarily hidden. A well-diffused critique of this situation is Pariser's *Filter Bubble* (2011). The argument suggests that as we train algorithms to know what we like, it will hide things we find objectionable, problematic or inappropriate, leading to users tacitly reinforcing their own worldview. Pariser's solution to this problem still remains within the algorithmic paradigm: One ought to actively seek out novel and challenging information to introduce noise to the algorithms. This criticism falls flat in two ways: First, it places the burden on the user when it was not the user who hid the information or rendered the filter bubble in the first place. Second, it suggests a lack of imagination about how we are to seek out, evaluate and consume information. The single ranked list remains intact, it is only the ranking itself that ought to change.

Instead of seeking to retune the invisible algorithm, I follow Gillespie (2014) in interrogating the entire notion of an algorithmic logic to the presentation of information. Gillespie's critique primarily concerns what he calls "public knowledge algorithms". These invisible algorithms simultaneously project an air of objectivity while rendering opaque the many features that went into such ranking. They do not merely determine what the appropriate answers are but also what the appropriate questions are. In that sense, they become an epistemology, or way of knowing, in their own right.

When considering the notion of social reputation, whether it is an eBay seller rating, a Google PageRank or Twitter Klout score, we tacitly assume that it is possible to discern a single ranked list of features—and that the challenge is in finding out how to tune the formula that creates this single list. Ironically, the very forms of data that have given rise to this ideology can demonstrate the inadequacy of this approach. Relational data (i.e. data that can be described as a graph) necessarily resist such abstract ordering—both mathematically and philosophically.

Philosophically, there are a host of reasons why relational data resist an ordering. I discuss these at length with regards to the specific domains of music, email and friending. But in general, nodes in relational data are categorical elements. One cannot say that jazz is better than rock music or polka music. One can only use proxies such as sales, listenership or recency. However, one can say jazz is more related to rock than polka by noting that many albums are labelled as both rock and jazz and few as jazz and polka. Thus, one can create a relational topology that signifies associations between categorical elements without necessarily implying a rank ordering.

Mathematically, we can think about the problems of ordering via the notion of a "monotonic function". A monotonic function preserves the ordering of a set such that if $x >= y$ and $F(x) >= F(y)$ then F is a monotonically increasing function. x is bigger than y and will always be that way when function F is applied. Many social network algorithms are not actually monotonic functions, however. One core non-monotonic social network metric is "betweenness" (Freeman 1979). If we want to know who is the "best connected" in a graph, it is common to use betweenness to report who is best connected across different parts of the network. Thus, if "Doug" has the highest betweenness in a graph, it means that the shortest paths run through Doug. But what about "Charlie", who has the second-highest betweenness? If you remove Doug from the network and recalculate the betweenness of the graph, Charlie might be in first place or he might be in third place. Without Doug, the shortest paths might reroute to include Charlie or reroute through someone else.[1] In short, networks have "dependency" issues; we cannot talk about the position of one node without taking into account the position of the other nodes.

If relational data are not necessarily orderable, then why do such orderings persist? There is no single answer, but many proximate ones: Practicality, technology and ideology all play a role.

[1] I want to thank Mason Porter (personal communication) for this insight.

In ***practical*** terms, such orderings exist because people find them at least sufficiently useful. In 2012, Google processed approximately 5.1 billion queries every day (Statistic Brain 2013). Facebook rank orders socially relevant news and over half a billion people use it every day. Putting information in a list is straightforward (literally straight and forward). To present something non-linearly requires something more than a single line. Should one present results in multiple columns? In a word cloud? On a globe? All alternative ways of presenting information require considerable decision-making about the aesthetics, layout and design. Another practical feature is that with a single ordering, sites with rankings can appeal to a scientific method for determining effectiveness: Alter the algorithm for half the users and see if they do more of the preferred outcome (such as "clicks per impression" or time on site). With alternate layouts that change both the layout and elements, it may be more challenging to change a single feature and identify its impact.

In ***technological*** terms, such orderings place limited demands on the users' hardware. If I am searching on a mobile phone or an old computer, I do not need a WebGL-enhanced browser to search for the number of a local pizza place. While being able to see these places on a map and evaluate their nearness, prices, storefront, reviews, etc. would be pleasant, it might not be necessary for the query.

That rankings persist because of ***ideology*** is perhaps the most challenging assertion. Ideology is a set of axioms or foundational statements that inform one's point of view (Fairclough 1995). For example, to have a Marxist ideology is to assume that class conflict and exploitation permeates all instances where capital is used to create value. Ranked ordered lists are emblematic of a *machine learning ideology*. This ideology assumes that as long as the system can learn from the user whether the ordering is effective, the ordering will get ever more effective (Segaran 2007). Everything is data, so training on ever more data will provide ever more effective results. Every click on a second result, every time someone does not "like" the top story on Facebook, every time someone purchases another product listed on a page are all signals to be reabsorbed into a ranking algorithm. Within this ideology, feeding data back into the system makes it possible to get increasingly close to the most accurate results.

One pernicious aspect of ideology for invisible algorithms is that it makes the collection and analysis of personal data a virtue. Individuals might not realize what data of theirs are being used, which can include everything from typing speed to how quickly the mouse moves on the page. However, within this ideology, such collection is still acceptable from the data provider's perspective. It is not a private matter, it is simply more data. It helps to train the algorithms that then make life easier by providing more effective results. In essence, it is an unknown social contract with a voracious voyeur who believes that this voyeurism will by its nature make itself more useful to the user.

Below are examples of how the logic of invisible algorithms as a manifestation of a machine learning ideology permeates the organization and presentation of multiple forms of data: music, email and friends. In each case, we can see that the data can be (and often are) relational, and that alternative forms of presentation and ranking have emerged.

Invisible Algorithms and Their Alternatives in Practice

Music

The digital revolution enabled a reconsideration of how music is discovered, distributed and organized. The distribution of music has always been bound up with socially constructed notions of taste as well as technical constraints (such as FM, cassette players, radio, size of an amphitheatre, etc.). However, when music became widely available digitally, it became possible to access heretofore unimaginable volumes of music with relative ease and convenience, prompting entirely new means for discovery and organization. Now that music can be integrated into databases that signify and calculate the relatedness of music, one does not need other media to discover new music. Virtually every streaming music service, whether it is Google Play, Spotify, iTunes or Pandora, has some means to impose a machine learning logic onto music. Thus, one can become acquainted with new music through the Internet without getting an explicit recommendation from a specific other person but from a *representation* of music as a *data structure*.

Representing music as data poses a number of significant challenges for either invisible algorithms or interactive affordances. This is because music exists within a highly tiered multilevel structure. A **song** is on an **album** that is by a **band** who are on a **label**. Each song might be more or less popular and include certain guest artists. At each level (song, band, etc.) one might also ascribe a genre (e.g. dance music), a subgenre (house music) or even a subsubgenre (electrohouse). While such features would seem to present an explosion of potential ways to explore new music, in most cases, these features are not represented structurally. Instead, music listening is reduced either to sorting by single rational signals (such as most downloaded, alphabetical, most recent) within any given single column, or via an invisible algorithm based on some amalgam of co-listening and other factors. One of the challenges of representing this relationally is not the paucity of possible relevant features for the graph, but the abundance.

In addition to the invisible algorithms of iTunes Genius, Amazon, Pandora and Google Play, interactive affordances are meant to provide an overview of music. These have yet to be adopted on any broad scale. I contend that one reason for this is the complexity of providing a Graph Search-like capability using database systems designed for sorting and ranking. Thus, we are left with user-curated networks of music, such as *Iskur's Guide to Electronic Music,*[2] a flash-based app that shows subgenres networked and linked in larger genres, all subjectively assessed. Other approaches include eMusic's "*Infinite Explorations*", based on user co-downloading, and *MusicPlasma,* based on Amazon co-buying. In most cases, however, these systems do not provide a categorical system based on sets and filters. Instead, the edges linking different music (typically artists) are themselves based on invisible algorithms.

---

[2] http://techno.org/electronic-music-guide/.

This leads to an interesting if somewhat underwhelming situation. The emergent maps do not tend to show macroscale features such as clusters. Instead, they merely project invisible algorithms into a different space. For instance, the music app Discovr operates as a music discovery service that draws in streaming from other services. It includes features for following bands and visualizing one's artists as a "music map". The map shows approximately six nodes based on a search for a specific artist. These nodes are provided by Echonest, a music tagging service that specializes in the tagging and organization of music. Echonest's clients include top music vendors such as Spotify, MTV and EMI, and they specialize in categorizing and tagging music according to a host of proprietary formats. In Discovr, one can click on any node and six more appear with links to the selected artist. Thus, one grows outwards a selection of nodes that are linked by an invisible algorithm queried by Echonest in order to presumably discover new music. What is not provided, however, are ways to determine what constitutes an edge, what sorts of community detection features are available, any notion of the edge weight or ways to filter. Thus, the network is merely a projection of ranked lists into two dimensions rather than one. How one arrives at the edges typically remains opaque even if the user can now browse non-linearly through such links. While the ranking algorithm has been given a high degree of polish (as is seen elsewhere), the interactive features remain thin.

Email

Email remains a dominant medium for communication. Yet it is overloaded as a communication medium, a to-do list, a means for registration and a host of other uses (Whittiker and Sidner 1996). The dominant way to organize email is simple: recency. This ordering makes a great deal of sense for most individuals. However, there are a small number of individuals who are utterly besieged by hundreds of emails a day. High-tech workers, journalists, celebrities and often academics routinely get hundreds of emails a day. This has led to the notion of triaging email. The term "triage" comes from the medical field where patients must first be triaged to determine their severity. In email, one might, by analogy, get an extremely important update as well as many offers for new jewellery, male enhancement and lucrative offers from wealthy princes abroad.

There has been much work on classification systems designed to sift out spam from relevant messages. This is exactly the sort of task for which machine learning shines. Using augmented Bayesian classifiers and other related techniques, modern email systems have been relatively successful at keeping spammers at bay. However, once one gets rid of the utterly inessential mail, how exactly ought we to organize what remains?

In 1999, Eric Horvitz designed PRIORITIES at Microsoft Research, a machine learning-based system for determining which emails to display to a user based on passive cues. The system would use visual cues such as bold or larger font to indicate a mail message that the system believed was particularly important. It used continual feedback from the user in order to better train the system which emails one ought to prioritize, essentially working within the ideology of machine learning. Within this framework, Horvitz et al. (1999) sought to train on criticality. That is, some messages and some senders have information that is known to be specifically time critical. The purpose of PRIORITIES was re-organizing mail in terms of its criticality so that the user would only receive alerts from messages when the cost of interruption was less than of ignoring the message. Like Google's 40-odd signals of personalization, PRIORITIES used a bevy of signals in ways that it considered to be appropriate, such as monitoring the acoustical environment, an individual's calendar and past responsiveness to individuals in one's inbox. While the algorithm that classified critical messages correlated at 0.9 with one user's own evaluation of which messages were critical, the system was never fully implemented into Microsoft Outlook. While there are no published records of why this occurred (in fairness, this is common at Microsoft Research), co-author and PRIORITIES developer Jacobs noted that when the system failed to assess a critical message, users would be at a loss to determine why and to fix this issue.[3] For users, this represented a significant problem. If one depends on a system for assessing critical messages, even a one-in-ten probability that a message is misclassified represents a seriously high level of uncertainty for everyday workflows.

More recently, Google has introduced priority inbox to their Gmail software. This system is not as explicitly oriented towards criticality, but like PRIORITIES uses machine learning techniques in order to differentiate mail that it suspects is important for the user. As their supporting documentation says, priority inbox uses cues such as sender, recency, whether it is sent to the user, etc. There is currently no public statistics on how many people adopt priority inbox; so it is challenging to evaluate its effectiveness. However, it is worth noting that like PRIORITIES it does not actually present the user with the specific reasons why a mail would or not be in the priority inbox. However, by keeping the priority inbox small and presenting the remainder of the mail on the same page, Gmail avoids the nagging suspicion about whether important mail is getting lost. Nevertheless, Google maintains their focus on invisible algorithms in presenting priority inbox. As stated on the help page "Gmail uses a *variety of signals* to prioritize your incoming messages, including who you've emailed and chatted with most and which keywords appear frequently in the messages you opened recently. If Priority Inbox mistakes an email as important or doesn't flag one that's important to you, *you can teach it* to make better selections" (Google 2013, emphasis mine).

The many signals embedded in an email, including cc, sender, recency and content, allow mail to be considered in ways other than merely criticality or priority. Following several years after PRIORITIES, Carman Neustaedter and the Community Technologies Group (also at Microsoft Research) took a more augmented than algorithmic approach to mail. Their program SNARF (Social Networks And Relationship Finder) was another experimental approach to mail that sought to augment

[3] Personal communication, July 2005.

human decision-making rather than operate as a substitute for it (Neustaedter et al. 2005). SNARF presented multiple views of one's email that filtered and sorted the senders based on the relationship to the user. Every view of email was based on a customizable filter that could be tweaked by the user. For example, one view showed the senders ranked by how frequently the user sends messages to the sender (implying that the senders with whom the user has regular contact ought to be ranked at the top). Much like PRIORITIES, SNARF was a research project and never released as a commercial product. However, it is a clear example of how social network information can be used in direct ways under the control of the user, rather than in indirect ways and reduced to a single rank order. That said, SNARF was probably too radical a departure from traditional email practices (in particular by showing the sender and a count of messages, rather than the messages themselves).

Like PRIORITIES being the spiritual ancestor of Gmail's priority inbox, the ideas embedded in SNARF also existed in a commercial product (although it is not clear that one influenced the other). Xobni (or "inbox" spelt backwards) was an add-on panel for Microsoft Outlook that similarly allowed significant customization of how email was sorted, paying specific attention to the relationship between the user and the sender. Xobni is still relatively unique in enabling a view of the user as well as a view of the messages. Most email products currently have the ability to search by individual or email address, but not a dedicated view of the individual and his/her relationship to the sender. Although Xobni is not likely to be considered a household name, it was acquired by Yahoo in late 2013 for 48 million in cash, suggesting that there is still significant interest in providing novel interfaces that augment human decision-making rather than merely new invisible algorithms that re-rank email within the dominant single list paradigm.

Friends

Relative to the management of messages and music, the notion of articulating friends online is perhaps the most recent and novel way in which invisible algorithms have been applied to data structures. Friendship is unambiguously multidimensional (Marsden and Campbell 1984). It combines recency, locality, taste, reciprocity, homophily and a host of other complex attributes into a single state—"the friend" (Gilbert and Karahalios 2009). As noted in qualitative research, however, friendship online is much maligned by the need to signify a host of different relationship types using a relatively user-friendly label. As boyd notes, there are many reasons to signify someone as a friend on a social network site, such as the social cost of saying no or the wish to look popular (2006).

The high-dimensionality of friends has led to the application of both augmented and algorithmic approaches, with Facebook at the vanguard of, ironically, both approaches. Here I mention Facebook's top news feed and the recent Graph Search.

The News Feed

When Facebook introduced a news feed in 2006, many users were in shock that the site would so aggressively repurpose content (Sanghvi 2006). Facebook had provided a novel affordance for consolidating and viewing content about others. But in doing so, it exposed people to a flood of information about their friends, some of which was not always warranted. The group "Students Against Facebook News Feed" (Schmidt 2006) gained over a quarter million members in less than a week. Interestingly, this first step was not an invisible algorithm but an interactive affordance. The news feed was a brilliant case of how to consolidate and reorganize content in such a way that it allowed users to see things they would not otherwise.

Over several years, the news feed included ever more information from other individuals to the point where there were substantial complaints about its capacity to provide information overload. Thus, in 2009, Facebook leveraged a machine learning logic of invisible algorithms in order to simultaneously minimize the volume of new information and the burden on users to configure what information was available. This was not the only direction Facebook could have taken. One (ostensibly failed) alternative to Facebook, Diaspora, explicitly sought to embed "aspects" or specific views based on categorically defined lists of friends. It was not unlike the groups of friends that Facebook originally tried. Not long after, Orkut rolled out a way to categorize individuals into lists automatically. In 2011, Google rolled out their alternative to Facebook, Google+, and explicitly embedded the logic of different spaces for different individuals under the notion of "social circles". Ironically, it did not include Orkut's capacity for automatic group creation based on social signals. Instead, Google provides users with a single list of friends and asks users to organize and create social circles with the significant tedium of dragging each user one by one into separate social circles.

Graph Search

The machine learning logic for presenting information and potential friends will always be limited by its reliance on invisible algorithms. While Facebook can record huge volumes of signals from the users, it also has a mass of highly structured data. Although Facebook are not using the semantic web explicitly, they have unambiguously embedded semantics within their own Open Graph. Graph Search is in many ways a culmination of a drive to use this information in visible rather than invisible ways. Like the "knowledge" engine, Wolfram Alpha, it follows a categorical ideology rather than a machine learning ideology. Within this ideology, it is assumed that a user can ask for very specific well-defined objects and that it is the user, not the system, which is to be trained over time. In both the cases of Wolfram Alpha and Facebook Graph Search, the interface indicates how queries are supposed to be contracted using a syntax that is similar to, if not exactly the same as, everyday speech.

In the case of Graph Search, filtering comes first and ranking second. One can filter to "my friends" or "friends of my friends". This can be chained with a series of attributes such as “who like bacon”. In fact, one of the early memes of Graph Search was a set of queries that exposed unexpected lists of people, such as “people who like white supremacy and hip hop music”.

What is peculiar about Graph Search, however, is that at present there is still no means for sorting within a set of people. Rather, Facebook’s preferred ordering is there. A cursory inspection of friends indicates the sort to be based on a vaguely definable logic of recency and activity. Given the necessity of ordering in some fashion, some measure of relevance would appear to have the advantage of not being entirely stable (and thus predictable and dull), but again it implies an invisible algorithm even behind what is otherwise a very user-centred set of interactive affordances.

Interactive Affordances of Friends as Networks

Particularly in the wake of significant optimizations in JavaScript and dynamic web pages vis-à-vis HTML5, WebGL and AJAX technologies, there has been a veritable explosion of new interactive affordances for the presentation of information. This is the first crucial step in moving beyond the single ranked list and the typically invisible algorithms that order this list. Google, for example, now provides extensive customization filters on their Google Maps product. Hotel booking sites provide a huge number of ways to pivot on salient details such as price and nearness. Travel booker Hipmunk presents users with a pleasant means for seeing flights on a timeline that can be sorted by user-defined details, including a clever sortable scale “agony” that weights by the price and the number of changes that one needs to make.

For relational data, the obvious next step is the presentation of data within a social network format. Indeed, I have been working with many colleagues in order to present Facebook data as a social network for several years. The latest incarnation, http://CollegeConnect.us/ (Fig. 7.1), is an attempt to provide users with a means for viewing their network holistically as well as exposing which people went to a college or university. The idea is to show individuals peers that have been to college or university so they can ask questions. We do not rank users based on personalization, but instead provide a traversable topology and a means for allowing users to explore the network themselves.

When viewing a social network from Facebook on CollegeConnect, users are presented with an open-ended means for searching. It thus serves as a form of discovery and a means to place information in context. That is to say, the position of an individual is indeed influenced by the individuals connected to that person. In a sociogram, a co-worker is adjacent to other co-workers and parents next to other family members. This form of contextualization enables individuals to now see related information in a wider context, again not unlike the use of networks to categorize music. By contrast, when search results, top news stories and collaborative filtering algorithms rank by invisible algorithms, they rank by the similarity of the elements to the user or the query. But these elements are rarely, if ever, positioned because of their relation to each other. On Facebook’s top news, one might see a story from a co-worker next to a story from one’s parents and one’s ex-girlfriend.

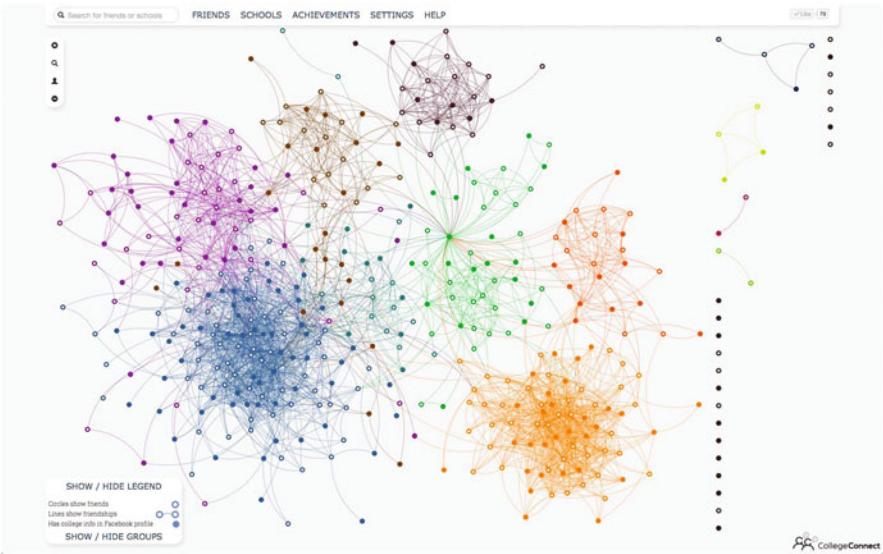


Fig. 7.1 CollegeConnect.us interactive interface for visualizing social networks and surfacing college information

The Challenges for Interactive Affordances

In the cases of web surfing, music consumption, email use and social media friending, there are different reasons for a focus on either invisible algorithms or interactive affordances. However, in all cases, it appears from cursory inspection that greater attention is being paid to the approaches that help users input less and receive more elements that are similar. That is, in all cases, information curators eschew a visual, often two or three-dimensional approach to information management for one that is one-dimensional for the user and n-dimensional for the server. The server gets as many dimensions of data as it needs (such as time spent on site, frequency of replies, distance to sender) and then using some algorithm reduces these dimensions to a single factor. It rank orders the users based on this factor and presents this information to a user.

Yet in all cases, I was able to demonstrate that much of this n-dimensional data can be described as a graph, and to that extent can be presented or arranged based on relations rather than just vectors of similarity. In the case of both music and friends, I showed instances where these relations are visualized as a network. So if all of these data can be described as a topology, then why are such topologies not ubiquitous in the user experience of information spaces? Here I present a list of reasons, rather than solutions. If interactive approaches are to seriously compete with ranked lists from invisible algorithms, then at least in the case of network visualizations, they must confront the following challenges:

1. ***Networks are sciency.*** From marketing materials to bestselling books, networks abound as complicated dot-and-line figures that show the world as interconnected and complex... and virtually unintelligible. These graphics commit a host of visual errors such as unnecessary edge crossings, too much chart junk or a poor layout. That they look complicated *is* the point. One of the core websites archiving such information design is deliberately called "visual complexity" (Lima 2011). Attempts to reduce the clutter of these networks also tend to require additional effort from the user to understand motifs or to "read" the graph in a novel way (as is done with hive plots or substrate plots). This creates the network as specialized work for professionals.
2. ***Toolkits for the visual presentation of network data are still lacking and configuration is difficult.*** It is only within the last couple of years that browser-based JavaScript toolkits for networks have emerged that can be easily embedded into websites. These tools still either require massive processing power to render the entire graph or are inherently limited in what they can show. A layout that works well to show everyone who tweeted *Kony 2012* will not work so well if one wants to dynamically filter out the peripheral tweeters from those who are central.
3. ***Centrality over community***. The triumph of the single ranked list harkens to a focus on key individuals. That is, often we want to focus on or discover the key individual (whether it is the most important person in the inbox, the exemplar of a genre or the most interesting tweeter). Being able to situate that artist, colleague or personality within their respective social context is not as prevalent. Networks permit us to see communities in conversation, but to do so might take the focus away from the singular individual.
4. ***What counts as "as simple as possible" is unclear***. There are narratives, conflicts and factions to be found in networks. Yet, it is virtually impossible to present all of this in a static image. While contemporary network graphics allow filtering and highlighting, it is still not obvious how much configuration is too much. An entire dashboard of controls to change a node's size or colour based on a variety of factors suggests that designers do not know which factors work best as defaults and why. This, then, hints at the central tension between invisible algorithms and interactive affordances—we wish to augment human intelligence through presentation of information, but still must trust the user to be able to know how to look for interesting data.

Conclusion

The criticisms against networks and interactive networks are significant and substantial. Where then does one turn? We can accept that a network is not always best presented as a sociogram, but it is still a relational structure. As such, we need not necessarily project a network in its totality. Many programs employ a graph but do not visualize it. Facebook Graph Search is one such example of a forward-thinking way to manage the complexity of graphs with the user rather than instead of the user. But in doing so, it had to step outside of a machine learning ideology that treats all data as columns to be weighted along the way to a perfect sorting. Instead, it had to combine highly scalable (and novel) graph databases along with named entities that are of specific types or classes and natural language processing. This sort of work is clearly interdisciplinary and demanding, but it does move the needle away from the notion of the perfect search through increased optimization.

The machine learning approach employing invisible algorithms reduces a large vector of measurements to a single scale and then presents a rank order of elements in this scale. In doing so, it must weigh the different vectors according to some scale—what is more important: Something that is local? Something that is liked by my friends? Something that is recent? The main concern with aggregation is that the recipe for this secret sauce is hidden from the user and based on the ideology that training is most important.

Information retrieval in a single column list is almost entirely dominated by some measure of relevance. The calculation of this relevance is typically based on an invisible algorithm. Such algorithms do not have to be deliberately manipulative or underhanded. However, to the extent that they are based on some amalgam of multiple measurements, such as cosine similarity in topic space or recency, they remain opaque. This opacity has given rise to both mundane complaints, such as the cheeky column “TiVo thinks I’m gay” (Zaslow 2002), as well as serious concerns, such as the legal case against Google. In the latter matter, a European commission has questioned whether Google had unnecessarily privileged its own social media product, Google + , in search rankings (Manne and Wright 2010). As of publication, Google is still in discussion with the EU as the algorithms behind the search results remain hidden even to the European authorities.

In the case of relational data, social network diagrams and graph navigation operate as an accompaniment and alternative to opaque scores based on relevance. While they do not necessarily eliminate the need for invisible algorithms, they shift at least some of the decision-making power as well as the work back to the user. The user can gain multiple advantages in these cases. First, there is the ability to see relational structures at a macro level. For example, building a network of American political blogs highlights how explicitly blog linking falls along party lines (Adamic and Glance 2005). The second is a clearer social contract—users are viewing output based on input that they can theoretically understand. The third is imagination. Users can now repurpose forms of data and project them in novel ways for new forms of discover.

The notion of reputation online is one that has heretofore been bound to the machine learning ideology. That is, how can we better create a system that ranks people, places or things using whatever data we have on hand? But in interrogating this logic, I am now prompting a new question: How can we maximize the users’ ability to evaluate these people, places and things themselves rather than to let the wisdom of crowds embedded in a machine learning ideology do it for them? It is not a denial of the use of social computing (or even a denial of the utility of invisible algorithms), but a renegotiation. The possibilities emerging from graphs and sociograms alongside invisible ranking algorithms are only now starting to be realized.

References

Adamic, L. A., & Glance, N. (2005). The political blogosphere and the 2004 US election: Divided they blog. In Proceedings of the 3rd international workshop on Link discovery (pp. 36–43).

boyd, d. (2006). Friends, friendsters and top 8: Writing community into being on social network sites. *First Monday, 11*(12). http://firstmonday.org/article/view/1418/1336. Accessed 10 June 2014.

Fairclough, N. (1995). *Critical discourse analysis*. Boston: Addison Wesley.

Freeman, L. C. (1979). Centrality in social networks conceptual clarification. *Social Networks, 1*(3), 215–239.

Gilbert, E., & Karahalios, K. (2009). Predicting tie strength with social media. In CHI '09: Proceeding of the twenty-seventh annual SIGCHI conference on human factors in computing systems. New York: ACM. doi:http://doi.acm.org/10.1145/1518701.1518736.

Gillespie, T. (2014 forthcoming). The relevance of algorithms. In T. Gillespie, P. Boczkowski, & K. Foot (Eds.), *Media technologies*. Cambridge: MIT Press. http://culturedigitally.org/2012/11/the-relevance-of-algorithms/. Accessed 8 Nov 2013.

Google. (2013). Priority inbox. *Google help forum*. https://support.google.com/mail/answer/186531. Accessed 8 Nov 2013.

Horvitz, E., Jacobs, A., & Hovel, D. (1999). Attention-sensitive alerting. In Proceedings of UAI '99, conference on uncertainty and artificial intelligence (pp. 305–313). Stockholm, Sweden. http://dl.acm.org/citation.cfm?id=2073831. Accessed 10 June 2014.

Lima, M. (2011). *Visual complexity: Mapping patterns of information*. New York: Princeton Architectural.

Main, M. (2002). *Data structures and other objects using Java*. Boston: Addison Wesley.

Manne G. A., & Wright J. D. (2010). Google and the limits of antitrust: The case against the antitrust case against Google. *Harvard Journal of Law Public Policy, 34*(1), 171–244.

Marsden, P. V., & Campbell, K. E. (1984). Measuring tie strength. *Social Forces, 63*(2), 482–499.

Neustaedter, C., Brush, A. J., Smith, M. A., & Fisher, D. (2005). The social network and relationship finder: Social sorting for email triage. In Proceedings of the 2005 Conference on Email and Anti-Spam (CEAS). Stanford: Stanford University.

Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. New York: Penguin.

Sanghvi, R. (2006). Facebook gets a facelift. *Facebook Notes*. https://www.facebook.com/notes/facebook/facebook-gets-a-facelift/2207967130. Accessed 8 Nov 2013.

Schmidt, T. S. (2006). Inside the backlash against facebook. *Time Magazine*. 6 September 2006. http://content.time.com/time/nation/article/0,8599,1532225,00.html. Accessed 8 Nov 2013.

Segaran, T. (2007). *Programming collective intelligence*. Sebastopol: O'Reilly.

Statistic Brain. (2013). *Google annual search statistics*. http://www.statisticbrain.com/google-searches/. Accessed 8 Nov 2013.

Turing, A. (1936). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society, 42*(2), 230–265.

Whittiker, S., & Sidner, C. (1996). Email overload: Exploring personal information management of email. In Proceedings of CHI'96 conference on computer human interaction (pp. 276–283). New York: ACM. doi:10.1145/238386.238530.

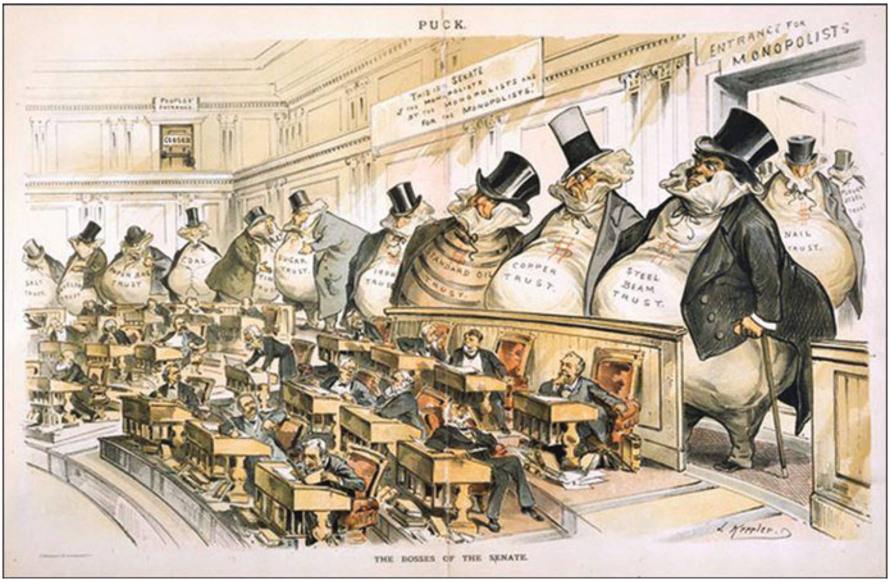
Zaslow, J. (2002). If TiVo thinks you are gay, here's how to set it straight. *Wall Street Journal*. 26 November 2002. http://online.wsj.com/news/articles/SB1038261936872356908. Accessed 8 Nov 2013.

Part IV
Novel Research Directions

Chapter 8
Breaking the Iron Law of Oligarchy: Computational Institutions, Organizational Fidelity, and Distributed Social Control

Howard T. Welser

Corruption: Enabled from the Top Down



(Joseph Keppler 1889: http://www.senate.gov/artandhistory/art/artifact/Ga_Cartoon/Ga_cartoon_38_00392.htm)

An external observer, witness only to the last century of advances in technology, might easily suspect that contemporary society would be unpolluted, egalitarian, and

H. T. Welser (✉)
Department of Sociology and Anthropology, Ohio University,
147 Bentley Annex, Athens, OH 45701, USA
e-mail: welser@ohio.edu

E. Bertino, S. A. Matei (eds.), *Roles, Trust, and Reputation in Social Media Knowledge Markets,* Computational Social Sciences, DOI 10.1007/978-3-319-05467-4_8,

prosperous. We now enjoy access to electric cars, low-cost solar energy, unprecedented communication technology, and dozens of other technological advances. This expectation is reinforced by the noble intentions of our political constitutions and the mission statements of our charitable organizations.

Despite these positive expectations, our world is still one of pollution, inequality, and poverty. Our leaders are charged to work toward collective goods, and yet leaders of organizations often work toward narrow self interests at the expense of the many and contrary to the declared organizational mission. Across all types of organizations, corruption emerges from the top down. Corruption can be understood by combining agency theory with Gottfredson and Hirschi's definition of crime (1990) where corruption includes acts of force and fraud by agents that employ organizational power and resources to further self-interests; these acts will often include undermining or damaging the organizational mission.[1] Accordingly, organizations across all areas of contemporary society suffer from corruption: economic firms, political parties, labor unions, educational institutions, governing bodies, judicial systems, and charitable organizations. Whenever agents exert organizational control, there is the risk that they will be tempted to subvert their authority and act toward their own interests. This problem is not new.

Over 100 years ago, Robert Michels (1911) observed that even the most radically democratic organizations would devolve into oligarchy. Michel saw oligarchy and corruption as the inevitable result of the hierarchical systems necessary for large-scale organizations. As authority and power accumulated at the tops of organizations, the temptations and rewards for corruption would become increasingly powerful, and the organization would make a transition to a system that primarily lines the pockets of the leadership elite and would progress very little toward the avowed goals of the organization. The decision making, influence, and capacity to evade effective monitoring accumulate at the tops of organizations. The systems of control flow primarily from the top down, and the absence of effective controls creates the conditions for corruption. The ideology or beliefs of the organization do not protect the members from the risk of corruption. Exemplary cases of organizationally facilitated crimes are easy to find among charities and churches as well as in government, education, sport, and business (Paltrow 2013; Washburn 2008; Mason et al. 2006; Huther and Shah 2000; Theobald 1990). The notion of a church leader like Pat Robertson using humanitarian aid to fund his private exploitation of a diamond mine seems more like satire than reality (Prophet Motive 1999). And yet, corruption and mismanagement of funds are commonplace in charities and other nonprofit organizations (Hundley and Taggart 2013).

As much as we might want to blame Pat Robertson, Jeff Skilling, or Dick Cheney for their behavior, we should remember that they behaved exactly as we should predict, given the opportunities presented by their organizational position. From the

[1] Corruption is sometimes used in a more limited sense where it is applied either to agents of the state, for instance, "using public office for personal gain" (Shah and Huther 2002) or actions of firms as they relate to the state, such as businesses agents bribing state officials to further the interests of their firm.

perspective of rational choice, and agency theory in particular, the difficulty in explaining corruption is not in understanding why people would steal or cheat, but in identifying the circumstances when they will not. Our organizations continue to be plagued by corruption precisely because they are designed with inherent shortcomings that make it very easy for leaders to serve their personal interests at the expense of the organizational mission. Those shortcomings hinge fundamentally on three factors: (1) hierarchically defined asymmetries in monitoring, sanctioning, and dependence such that high-status leaders in organizations enjoy very weak monitoring are subject to weak sanctions and are often the least dependent upon their organizations for their continued well-being; (2) extant systems of review, evaluation, and social control in organizations are concentrated by hierarchy and the application of these system is inescapably connected to actors identity and their organizational roles; and (3) opportunities for the most lucrative corruption are often correlated with position in the hierarchy such that upper-level managers who are the least constrained by the systems of control have the best resources to exploit. It is instructive to consider the control systems of extant organizations in light of theoretical models of social control, agency, and normative compliance.

The leaders of our organizations are made amoral by the design of those organizations, organizations that have not been redesigned to solve problems of oligarchy and corruption identified over a century ago. Rather than working to solve the most pressing human problem of the last century, we, the social scientists, information scientists, and social computing engineers, have been busy with other tasks. This chapter is an invitation to all who are in a position to study or design organizations to solve our most fundamental social problem.

Embedding Organizations in Systems of Computer-Mediated Interaction

Despite Michels' warning, organizations have continued to foster the concentration of power, organizational control, and inequalities of responsibility that gives rise to corruption and distortion of organizational missions. However, recent developments in large-scale online communities illustrate some of the ways that organizations can overcome the tendency toward oligarchy in organizations. Five examples from computational systems provide partial clues to how organizations can use digital affordances to overcome some of the inequalities and distortions that occur in traditional, large-scale hierarchical institutions.

> Large software-producing firms manage major code-writing projects through a code base that is edited in a context that preserves the content of every edit and the identity of every editor. In principle, these records could become the basis for organizational review and reward. However, such records of code are not typically part of a systematic, automated, and double-blinded peer review process. So, even though data on the quality of every programmers contributions is potentially available for evaluation and reward, that data is only partly used, and the reputations of the coders and the judges are likely to drive the assessment of any particular section of code.

> Wikipedia, and wiki projects in general, demonstrates a radical flattening of organizational hierarchy. In principle, every participant can evaluate the contributions of every other editor and can administer positive and negative sanctions. The asynchronous nature of interaction and the automated recording system allow members to evaluate the content edits as well as the tenor of interactions preserved in the edit record. The project illustrates how large-scale organizations can be flattened and made more democratic. Positive reputational effects are present as are deleterious effects related to personalistic, political, and other capricious motivations. The same opportunities for double-blinded monitoring and sanctioning present in the large software datasets are available in Wikipedia as well. However, the review of contributions remains arbitrary, uneven, and overly influenced by reputation effects.
>
> The online discussion system Reddit presents a large-scale online community where all registered contributors can vote contributions up or down, which in turn influences the visibility of those items and accumulates as an assessment of that member's identity. The system is large scale and has introduced many behind the scenes modifications to try to limit users' capacity to game the system. The system also allows users to create multiple accounts, which in principle allow contributors to contribute in ways that they might otherwise not feel free to. Identifiability of actions with identities is an affordance of computational institutions that can be managed to enhance contribution toward organizational missions.
>
> Google documents allows multiple parties to edit a collaborative document and has the capability to present editors either as anonymized characters or as their login identity. While not currently enacted in this manner, contributions to such documents could be presented as edited by particular anonymized contributors who could be evaluated and sanctioned according to the content of their contributions alone, and those sanctions could be delivered to the identity of the contributor while maintaining the anonymity of both parties. The sanctioning acts themselves could be presented to evaluators who could judge the merit of those acts while maintaining double-blind anonymity.
>
> CrowdGrader is a homework grading system that enlists students as peer reviewers in a double-blinded review system where mechanisms for encouraging accuracy of evaluations and feedback on the evaluation process are integral to the design. After submitting their homework contributors are required to evaluate the work of a random sample of anonymized others. The fundamental procedures of this grading system could be applied to contributions in coding projects, in Wikipedia, or in meetings that combine aspects of Google-documents-shared editing system and chat features.

Computer-mediated systems like these allow organizational actions to be captured as digital events, and they represent major advances in the potential democratization of participation in organizational monitoring and sanctioning. The open structure of reporting all contributions takes an important step toward transparency in democratic organizational control. These systems also operate, partly by leveraging the positive side of informal reputational systems. By linking contributions to semidurable identities, the contributors are encouraged to comply with organizational goals, and this compliance is made less subject to distortion by organizationally powerful individuals because it is a distributed system of monitoring and sanctioning. The association between acts and personally meaningful identities allows norms, identity, and social sanctions to influence behavior. Those positive dimensions are an important focus for the Kredible net conference and research efforts. However, this chapter will highlight the fact that reputation systems can also have counterproductive effects and that computational systems can be further extended with formal systems that double blind the review of organizational contributions and otherwise break the link between organizational actions and capricious implementation.

The digitally mediated contexts discussed above illustrate ways that online systems offer different interactive affordances for solving problems of social order, but extant systems still perpetuate many of the linkages between actor, action, sanction, and sanctioner that have the potential to foster oligarchy. The purpose of this chapter is to spur further attention to mechanisms that promote oligarchy in organizations and to begin the process of developing systems of distributed social control that will ultimately break the iron law of oligarchy.

Theoretical Models for Controlling Corruption in Organizations

Michels (1911) drew attention to several factors that would lead to organizational leaders to use their authority for personalistic ends. In different ways, the factors that undermine compliance with the organizational mission hinge on inequalities of access to and implementation of the systems of organizational control. The fundamental principles of these systems are best described by combining principles of Weberian institutional analysis with models of social control from the rational actor tradition. The following section articulates the key contributions of these and related theories for understanding social control in organizations.

Max Weber (1978) developed an analytic framework for assessing differences between institutional forms. Where Weber focused on developing ideal types of authority systems and classes of institutions this discussion will focus on how models in rational choice theory can help identify features of organizations that contribute to, or impede corruption in organizations. This analysis will focus on dimensions directly related to instrumentally rational action and generally to motives to action linked directly to material cost and benefit. Material incentives and instrumental rationality are not the only important dimensions of action, rather they provide an important minimum threshold for analysis. Future work should consider how charismatic, habitual, and traditional modes of action can both increase and decrease the likelihood of corruption in organizational types.

The rational actor framework (Coleman 1990) combines some version of the following assumptions in order to model the social implications of the decision making of actors in particular institutional contexts. (1) *Instrumental action.* Action is assumed to be instrumental or purposive which means that courses of actions are pursued because of the expected results of those actions. Actions are not therefore assumed to be expressive, habitual, bounded by tradition, or otherwise enacted without attention to the likely results of those actions. (2) *Self interested.* Actors are typically assumed to be individually self-interested, which means that they generally seek to maximize individual benefits while avoiding individual costs. (3) *Rational.* Rational actors use a comparison of costs and benefits to select the path of action that is most likely to maximize their interests. Rational refers to the decision-making process, not to the values held by individuals. (4) *Material values.* Actors typically are assumed to value money or other fungible goods (goods that can be exchanged for other goods). Depending on the context, power and status may serve as important supplementary value assumptions.

Given these first four assumptions models typically call two contextual factors into analysis as well: (5) *constraints of available information,* and (6) *discounting* of future payoffs for given levels of uncertainty. Although actors are assumed to act within the constraints of available information, differences in availability of information help to define important ways that organizations differ in their susceptibility to corruption. Similarly, in situations of organizational uncertainty, access to short-term gains will tend to undercut compliance with organizational rules that are incentivized by distant, uncertain future rewards.

The Problem of Controlling Agents

The best model for framing the problem of corruption in organizations is agency theory. Agency theory or the "problem of agency" refers to situations where one party, the principal, holds the rights to some resource, but needs to entrust another actor, the agent, to act on the behalf of the principal (Ross 1973; Arrow 1984; Eisenhardt 1989; Kiser 1999; Shapiro 2005). In businesses, employees are agents that are contracted to act on the behalf of the owners of the business, the principal. In Robert Michel's example of the Marxist labor union, the union leaders act as agents of the union, which is collectively owned by the membership. The leaders then are contracted to act on the behalf of the union as a whole, but like all agents, they are presented with options that will further their own interests at the expense of the principal.

Kiser uses a sociological variety of agency theory (Kiser 1999) to predict likelihood of corruption and inefficiency in state organizations and to explain variation in the organizational strategies adopted in a variety of premodern circumstances (Kiser and Schneider 1994; Kiser and Cai 2003; Kiser and Sacks 2011). Kiser and coauthors' use of agency theory focused on variations in capacity for monitoring, sanctioning, dependence, and the alignment of interests. Their research allows us, in the historical and comparative setting, to better predict when and where corruption should be more likely based on matching the organizational solutions to the agency problem as enhanced or limited by the technological and practical limitations of each case. In the contemporary period, a shortcoming of Kiser et al.'s research is that it assumes that under conditions of modernity, a bureaucratic system will be optimally efficient. However, we need to extend our standards of evaluation beyond the traditional understanding of standard bureaucratic design because of digital innovations that offer substantially superior alternatives to century old systems of unmediated, interpersonal, hierarchical management.

The Problem of Powerful yet Low-Dependence Agents

The most basic solution to the problem of agency is to align the interests of the agent with that of the principal. When lawyers earn a percentage, or salespersons work on a commission, the contractual relation is written to allow the principal to assign a portion of the proceeds to the agent, which is paid in direct relation to the agent's capacity to succeed. In the simple principal/agent (where the principal is the owner who stands to gain materially from successful completion of agent duties), the systems of monitoring and sanctioning only need to be applied to the agent because the principal is automatically aligned with her or his existing interest in the success of enterprise. The principal, by virtue of ownership can claim exemption from monitoring as well as the right to impose the conditions of the contract. The agent can try to negotiate alternative conditions, but the agent's primary recourse is to simply refuse the work, thereby forfeiting claim to any compensation. This arrangement works as a solution to the agency problem in simple organizations where owner/principals can be clearly identified. However, most large organizations create additional challenges.

Large organizations often employ many levels of management, and those managers often behave both like principals (in terms of power) and like agents in terms of interests. All managers, but especially upper-level managers will often be granted the privileges and status of the owner/principals (relative freedom from monitoring and sanctioning, power to create contracts, high compensation; for a related discussion of overcompensation and corruption see Zyglidopoulos et al. 2009). However, these actors, who take a managerial role, frequently do not hold sufficient interest in the success of the organization. The executive privilege that high-level managers retain would seem to be a traditional carryover of rights from a simpler model of organization where the executive level manager is the owner, and thus a true principal. High-level executives typically will have access to corrupt opportunities that are both highly lucrative and damaging to the interests of the organization. When viewed from outside of an organization, or from the vantage of a true principal of an organization who is entirely dependent on the success of an organization, it seems quite strange that these glorified agents enjoy such power and latitude with far less monitoring than lower-level agents. For any individual agent, the degree of compliance with organizational missions and values depends on the degree to which agents are dependent on the success of the organization for their own current and future well-being. For the organization, cooperation and success depends on the degree to which dependence is universally shared across agents and roles. Corruption will be limited to the extent that hierarchical elites do not hold different organizational interests than rank and file agents and the organization as a whole.

Motivating Contribution to Group Goods Through Monitoring, Sanctioning, and Dependence

Hechter defines solidarity as the proportion of individual goods (time, energy, resources) that are contributed to a group (1988). Obligatory groups, such as communes or other voluntary associations face especially difficult challenges in motivating members to contribute rather than to free ride and enjoy the benefits of group membership without performing their group obligations. The key solutions in the sociological rational choice approach involve the development of socially efficient systems of monitoring, sanctioning, and dependence. While Hechter's (1988, 1990) work focuses on obligatory groups, firms and other compensatory groups still face major challenges in ensuring contribution even with the additional leverage of contracts and pay incentives.

In face-to-face interactions, and especially in large organizations, systems of monitoring and sanctioning are often incomplete, and key organizational members may experience low levels of dependence, seeing plenty of alternative opportunities and low exit costs. The same monitoring challenges confront common pool resources and related communities and industries (Keohane et al. 1993; Coleman and Steed 2009). Improving systems of monitoring and sanctioning should result in higher levels of compliance with organizational values and lower levels of corruption. Theories of common pool resources, group solidarity and agency all suggest that actors who are strongly dependent on the organization will be less susceptible to corruption and, further, that developing effective and efficient systems of monitoring and sanctioning are necessary preconditions for achieving high levels of contribution and low levels of corruption. This chapter offers two key extensions, first, that firms and other compensatory organizations still require major improvements in the design of monitoring, sanctioning, and dependence, and second, that the asymmetries of control and privilege that are traditionally associated with rank in organizations must be replaced by a distributed system of control, or we will not escape from the iron law of oligarchy.

Distributed Social Sanctioning, Norms, and Organizational Controls

Organizational norms can be understood as the informal rules and expectations that control the actions of some organizational agents under certain conditions (Coleman 1990). In Coleman's understanding, proscriptive norms arise as solutions to imbalances of externalities. An externality is a third party cost that arises due to an agent's actions in a given situation. If an agent of an organization acts in ways that create negative externalities for others, those other members have an incentive to exert negative sanctions, and to the degree that they do so, we can say then that a proscriptive norm emerges as an informal means of controlling the proscribed behavior. Similarly, prescriptive norms arise when third parties award positive sanctions for behaviors that generate positive externalities for them. In Coleman's model, when and where norms arise becomes an empirical question, one that depends on the capacity of third parties to impose negative sanctions for violations and positive sanctions for compliance. However, the informal and personalistic nature of sanctioning behavior gives rise to additional challenges, such as the second order free rider problem (Heckathorn 1989; Coleman 1990; Panchanathan and Boyd 2004) where third parties are unwilling to exert the social sanction on norm violators because they fear retribution or are otherwise unwilling to pay the cost of delivering the sanction.

Further problems arise with organizational norms when those norms reward behaviors that undermine the organizational mission, contradict organizational values, encourage shirking, or otherwise impede compliance with official rules or with participation in formal monitoring and sanctioning systems (Jones 1983). Organizations may introduce special offices, like that of the ombuds, to internally address issues of institutional justice and equity (Huther and Shah 2000). Such offices offer a partial solution to the fact that traditional hierarchies can themselves become sources of actions and influence that undermine the mission of the organization or violate rights of organizational members. Review boards, abuse hotlines, whistleblower rules, or other mechanisms can offer a partial solution to enabling distributed members of organizations to help enforce organizational missions or values. However, these types of mechanisms themselves can be abused and can be implemented for personalistic reasons or vendettas (Gould 2000). This work, research on norms, and one organizational culture all suggest that important improvements can be made in the capacity of organizations to facilitate effective social monitoring and sanctioning.

Informal Groups, Identity, Status, and Biases

All organizations face challenges in which the informal social network and identities of agents of the organization can give rise to factions or other subgroups that act in ways that undermine the mission and values of the organization. A major challenge for every organization is to create effective social control systems that reflect the mission and values of the organization. Organizational control systems that allow agents to sanction in knowledge of personalistic ties and identities open the organization to corruption of the very system of social control (Prendergast and Topel 1996). Agents in organizations can also be influenced by conscious as well as unconscious biases in their evaluation of others, according to research in status characteristics (Berger et al. 1972) and cognitive psychology (Greenwald and Krieger 2006). People evaluate the work quality of low-status actors lower; they defer expertise to higher-status actors, and in numerous ways allow status expectations to shape their assessments even when status differentials should have no bearing on content of a judgment (Berger et al. 1972). Status differences are especially problematic when they are correlated with organizationally defined sanctioning role. To the extent that personal identities and group memberships are connected to the evaluation of work performance, these biases threaten the legitimacy of the organizational control system.

Reputation, Evaluation, and Distortions of Judgment

People will work for fame and notoriety, and many organizations try to leverage the desire for esteem and respect in order to motivate members. In online communities, reputation systems are seen as important, though fickle, tools for motivating contribution and exerting social control (cites). In Wikipedia, authors award barnstars to recognize noteworthy contributors, Ebay sellers work hard to curry positive reviews, redditors grant karma to those whose contributions they value, and many other systems all participants make comments, reshare, or assign positive or negative sanctions. Through all of these processes, actors accumulate the results of their past behaviors, and in so doing develop reputations, and these reputations have the potential to influence their future behavior and to influence how others behave toward them.

While reputation systems have the potential to spur contributions, encourage prosocial interaction, and otherwise aid organizational missions, they can also distort and undermine evaluation systems in organizations. The Matthew effect (Merton 1968; others) describes the tendency for reputation systems to distort assessment of current merit: Those which are already well known, or perceived to be high status, attract more than their fair share of accolades and endorsements while those which are less known go wanting. Reputation systems can lead to a "winner takes all" society where the rewards for performance are disproportionate to objective differences in performance (Frank and Cook 1996). When people are given a judgment task, they tend to grant too much quality to the work of the famously good and not enough merit to the work of the obscure. When agents in organizations evaluate work of peers, their judgment is distorted by reputational information, and thus a system that blinds evaluators to the identity of the contributor will provide a more accurate assessment of that work.

Reviewers can clearly be influenced by the reputations of those they review, but they may also be influenced by the awareness that their review may contribute to their own reputation or assessment in the eyes of others. The story of the emperor's new clothes underlines a second problem that reputation systems create for evaluation processes in organizations: Public reviews of quality become subject to concerns about how exercises of judgment reflect upon the judge. Reviewers can be reluctant to express support for unpopular work or unpopular contributors, or under other circumstances, reviewers will use their assessments as statements to draw attention to themselves for a variety of reasons, distorting the assessment that they would make if they were purely trying to accurately measure the quality of the work in front of them.

Double-blinded peer review in the scientific academy demonstrates how an institution has adopted procedures to overcome reputational concerns in the assessment of quality. Even with those efforts, we can still point to examples of authors whose work makes them largely identifiable, and in others where reviewers may tip their hand, revealing their own identities. However, even if the implementation is imperfect, the effort to double blind reviews in the scientific process provides support to the claim that the reputational issues raised above present an important threat to the validity of assessment in organizations. Issues of reputation can distort evaluation processes in organizations, and without a system for double blinding the review process, such reputational concerns will always be present. This threat should be addressed at a larger scale and across the full range of organizational contexts where the accurate measurement of member contribution is important for the success of the organization.

We can use the term "organizational fidelity" to refer to the capacity of an organization to actually implement its mission, to encourage exemplification of its values, and to maximize the productive contributions of members. The concepts discussed above identified challenges to social control and evaluation in organizations. Taken together, these concepts suggest ways that organizations vary systematically in the degree to which they either facilitate or impede organizational fidelity. While it may seem utopian and unrealistic, digital institutions have the potential to combine features that previously were impractical, costly, or unrealistic. We can now build organizations that are not limited by the shortcomings that Michel saw as inescapable, and thus the next section provides an ideal typical model for high-fidelity organization that could actually be implemented.

Design Elements for Distributed Organizational Control

Computer-mediated work provides new opportunities for organizations to solve the long-standing problems of social control that lead to corruption in organizations. These problems stem from asymmetries in the flow of and access to information about organizational contributions, and in asymmetric constraints on capacities for exerting control through monitoring and sanctioning. This section outlines the necessary conditions that when combined would allow all organizational agents to constitute the systems of organizational control that will allow the membership of an organization to reward compliance with organizationally defined values and to prevent agents from exploiting asymmetries of information for their own advantage. The purpose of this section is to describe the full list of attributes that need to be designed into an organization that can overcome oligarchy through distributed organizational control. This section also helps to articulate an ideal type that can be used as a standard to compare to extant organizations that will reveal the sources of corruption in those organizational designs.

Mission A clear mission statement is needed to allow all agents to judge their contributions and the contributions of others according to how effectively those actions advance that mission. The mission helps to define how the values of the organization are expressed, and thus how the quality of contribution is created through agents' actions and how it is measured in the evaluation process.

Values Values are represented in dimensions of agents' actions on behalf of the organization. Organizations need value statements that can be clearly translated into actions, and the qualities of those actions need to be measurable and comparable in terms of those values. Organizational values need to be clear and concrete enough that when an agent commits an action on the behalf of the organization that this action can be at least interpreted as good or bad, and as better or worse than comparable actions.

Contributions All organizations require agents to take actions on the behalf of the organization. These contributions are of different types and will vary in terms of quality and impact. An effective system of distributed social control will require a set of definitions of the types of contributions that will be collected, evaluated, and sanctioned in the system. The effectiveness of the organizational control system depends on the capacity for the most important contributions to be recorded digitally and for their dimensions of quality to be accurately recorded in the system and thus be made available for distributed evaluation.

Criteria of Evaluation Criteria of evaluation are the definitions that apply the organizational values to the measurement of quality for the defined types of contributions. These criteria need to be expressed with enough clarity that third parties could employ those criteria upon a sample of actions and achieve a high degree of intercoder reliability. Reliably measurable attributes of content is a standard expectation for content analysis, and to the extent that actions are rendered as textual records of agent actions, measurement of the quality of individual actions in an organization will need to, at least, rise to the level of basic social scientific methodology.

Digital Work Efficient distributed organizational control requires a digital environment. Work performed in a digital environment creates digital records that can be stored and selectively displayed while linking to stable identities, but while also controlling the flow of identity information in ways that are not feasible in face-to-face interaction. Not only are costs of monitoring and sanctioning reduced, but the system can be used to both maintain and obscure the key links between identity, evaluation, and sanctioning discussed in the Section “Theoretical Models for Controlling Corruption in Organizations.”

Completeness of Information The greater the percentage of each agents contributions are accurately recorded in the work system the more effective the system will be at both motivating contribution and directing those actions toward the mission of the organization. The complete system would be both motivating (because all actions would “count”) and would be mission enhancing because contributors would be aware that the quality of their contributions would be measured.

Double-Blinded Evaluation Digital environments make it possible for the content of a contribution to be presented in a context that obscures the identity of the contributor and the evaluator while maintaining a connection to their identities. This allows evaluation to be based on the quality of the contribution rather than the reputation of the agent or any other biased rational. Furthermore, the identity of the evaluator is blinded during the act of evaluation to remove consideration of positive or negative second-order sanctions. Inescapably identifiable actions, like decisions by agents in leadership, are subject to single-blinded reviews, so that the reviewers can honestly assess the value of the contribution without fear of retribution. Inescapably identifiable contributions are a predictable result of leadership roles or highly distinctive tasks, and thus, given the fact that these will be subject to review that is not blinded, they should motivate extra compliance with known criteria of evaluation since the reviews of the evaluations will themselves be blind peer reviewed.

Evaluation Is Itself Reviewed Evaluation acts are contributions that are themselves subjected to double-blinded evaluation. This gives evaluators incentive to maximize the accuracy of their evaluations.

Performance-Based Compensation To the extent that dimensions of agents' contributions are accurately recorded in the system, their compensation should depend on the quality of their contributions. Unless the values of the organization are enforced through the compensation system, they are like laws without the sword, mere words (an analogy to Hobbes on the necessity of force as the defining dimension of the state).

Dependence and Distribution of Dependence Compliance with organizational rules and degree of contribution toward organizational mission will be higher when agents are more dependent on organizational compensation for the current and future well-being. The distribution of dependence of agents within the organization should reflect the importance of the compliance of those agents with the values of the organization and capacity to contribute to the success of the mission.

Uniform, Universal Constraints on Agent Control over Visibility of Contributions All agents have equal and limited capacity to influence the visibility of their own contributions and the visibility of others.

Uniform Evaluative Rights and Obligation All agents are equal participants in the organizational control system. Each agent has equal access to and obligation to participate in the evaluative role.

Uniform, Universally Subject to Evaluation All agents of the organization have their contributions recorded by the digital work systems and made available for evaluation.

Open Code and Transparency of Rules The only way to ensure that digital infrastructure is not itself corrupted is to open the code to knowledgeable review and correction. The only way to ensure that the rules of the organization are not being manipulated to favor some at the expense of others is to open those rules to review and correction. Openness, review and correction in the machinery of the digital work system is necessary to prevent organizational agents from distorting the function of the distributed organizational control system.

Summary of Design Elements for Organizational Fidelity and Effective Organizational Control

The attributes listed in the Section "Design Elements for Distributed Organizational Control" can be used as a preliminary ideal type. By holding up these standards, we can compare the ideal to extant organizations and see dimensions where current organizations have features that deviate from distributed organizational control and therefore where those organizations contribute to corruption and oligarchy. It is immediately apparent that traditional features of organizations exist in stark contrast to these design elements (Table 8.1).

In general, we see that despite the fact that high-level contributors in organizations act primarily as agents rather than as principals, privileges associated with rank make them less subject to organizational control and more likely to become corrupt. Organizational fidelity is actively undermined by the concentration of privilege related to capacities for social control. The more that privileges of social control are concentrated according to rank in hierarchy, the more oligarchic the organization will tend to be and the more likely that corruption will undermine organizational fidelity. The more universally distributed the items are, the more democratic the organization will tend to be and the higher fidelity we expect for implementing the organizational mission.

Implementations for Organizational Fidelity: Partial Examples from the Contemporary Social Media Ecosystem

Existing social media systems and some organizations currently implement procedures and rules that enact some aspects of organizational designs discussed above. These systems are far from perfect, and they retain many attributes typical of institutions that readily encourage corruption and oligarchy. In this discussion, I will try to focus on helpful lessons to draw from existing systems and identify paths to strengthen the potential for effective distributed social control.

Universally Distributed Capacity for Monitoring and Sanctioning, Semipublic Sanctioning, Reputation Systems

Both Reddit and Wikipedia represent large-scale experiments in universalizing the capacity for monitoring and sanctioning across a population of participants. In Reddit, every post and every comment is subject to evaluation and can be sanctioned. Every member of the community can exert an equal (though weak) sanction, delivered either to a comment or initial link post of any other participant. Quantitatively, sanctions are minimally expressive: Each login identity can grant a single upvote or downvote per item. Additionally, participants can reply to posts or comments, and in so doing, administer a qualitative sanction, which, to the degree that comment influences the reading and voting behavior of others, can result in larger-scale changes in the vote count or "karma" of the login identity to whom they have replied. However, upvoting and downvoting behaviors are not subject to review or evaluation, and no justifications or explanations of voting decisions are integrated into the social control procedure. Therefore this social control measure is not itself subject to social control, lending to potential arbitrary and antisocial implementations.

Table 8.1 Comparison between distributed social control and traditional organizations

Dimension	Distributed control	Traditional organizations
Mission	Clear, formal, public; directly linked to values	Varies
Values	Explicit, clear, concretely tied to desired contributions of organizational actors	Varies, seldom clear and complete enough to allow systematic, valid review
Contributions	Codified into measurable, important types that are digitally recorded	Widely varied by level in hierarchy, clarity of contributions decreases with status in hierarchy
Criteria of evaluation	Clear and complete enough to allow high intercoder reliability	Hazy and incomplete, especially at higher levels in hierarchy
Digital work	Extensive use of digital participation systems for all crucial types of contribution	Few if any of the key decisions or contributions are recorded digitally in any monitorable form
Completeness of information	Full records of all digital work is made available to evaluation system	Little if any of digitally recorded work is made available for systematic review
Evaluation	Double-blinded evaluation for all feasible actions, single-blinded review for inescapably identifiable actions	Organizationally determined by role and position in hierarchy, top down review of lower-status agents
Review of evaluation	Automatically and universally included as feature of evaluation process	Reviews are not systematically monitored, and if review is made, it is made by hierarchically privileged actors
Compensation	Compensation from organization depends upon quality measured from evaluation process	Compensation dependent on privilege in hierarchy as well as control over compensation of others
Dependence	Maximize dependence of agents on quality of contribution toward both long- and short-term organizational mission, distribute dependence uniformly through organization	Hierarchically defined privilege in freedom from dependence and capacity to control dependence of subordinates
Visibility of contributions	Uniform and universal obligation for contributions to be visible to digital recording system	Hierarchical control over access to own contributions as well as access to others
Evaluative rights/obligation	Uniform and universal right to and obligation to participate in evaluation system	Hierarchically defined privilege to engage in evaluation and to hide contributions from evaluation
Subject to evaluation	Uniform and universal subject to evaluation of contributions to organization	Organizationally determined by role and position in hierarchy, top down review of lower-status agents
Transparency of rules	Transparent and public rules, open source code for all digital institution systems	Access to rules a privilege of rank, capacity to change rules also a privilege of rank. Limited transparency

Another general source of bias in implementation comes from the fact that sanctioning of particular posts or comments results from the organic process of reading and commenting on whatever is interesting to each participant, and thus the vote count for any contribution can be greatly influenced by arbitrary circumstances. Another shortcoming relates to the potential for valuable contributions to receive too little attention, and trivial contributions to receive too much. These will distort the vote total measures in ways that diminish their potential to reflect community goals or standards. There are no automated processes for delivering contributions in need of sanctions to participants, and there is no obligation to sanction particular contributions, and there is no systematic process for making sure that the review process is providing the feedback that would enhance the organizational mission.

Wikipedia faces similar challenges for evaluating contributions of members. Because all edits are recorded, and the login identity accompanies the timestamp on the edit, the wiki software allows actors to sanction particular edits. However, this edit is typically made when reviewers also have access to knowledge about the identity of the contributor (Edits by IP addresses are normally treated with great skepticism as are edits by recently created login identities). The fact of automatic recording of editor identity is what allows Wikipedia to eventually identify contributors, like sock puppets who use multiple proxy accounts to circumvent organizational rules on appropriate editing practices (Owens 2013). However, the presence of identifiers with each edit makes all wiki systems susceptible to corruption in the forms of the Matthew effect, vendettas, patronage, and other types of bias based on identity of the contributor. However, the fact that the default condition in Wikipedia is for all editors to have equal access to contribution information and to all have potential to deliver some sanction takes a major and important step toward a distributed digital organization. Implementing a supplemental review system based on double-blinded peer review could allow additional progress in this direction.

Double-Blinded Peer Review Systems

Digital systems can automate procedures for double-blinded peer review. To implement review in online work settings requires identifying sets of equivalent tasks (in this case, homeworks) which can be collected during a period of time and then later redistributed to a pool of potential reviewers through the collection interface. CrowdGrader (de Alfaro and Shavlovsky 2013) is just such a tool designed for instructors to use in courses where the students can act as graders of the assignments of their peers. After submitting an assignment, students have a rubric, grading instructions, and are presented with a series of anonymized assignments to review. They score and range the assignments and leave qualitative feedback. After the period of review ends, students can view the scores and comments from the variable number of (typically 5 or 6) reviewers who judged their contribution. Students have the capacity to respond qualitatively to reviewer comments and grades as well as assign a negative or positive score to that review. I have used this tool in my courses, and it seems to have some beneficial effects on student's perceptions of the legitimacy of their marks as well as expanding students' awareness of the range of variation in quality of work. In addition to organizing the presentation of assignments for review, the system uses recommender system algorithms to judge the accuracy of reviews and penalizes students scores if their assessments are inconsistent with consensus, and it also assigns greater weights to the reviews of students whose own work is highly scored. Students, as a part of their assignment grade, are obligated to perform reviews as part of their assignment grade.

A system like CrowdGrader is not perfectly applicable to work environments, but it is an extremely valuable illustration of how computer-mediated systems can lower the transaction costs involved in a distributed, double-blinded review system. The general approach could be extended and developed and could be applied to other types of digital contributions ranging from writing prose, writing computer code, contributing to online discussion, editing resources like Wikipedia, and any other work-related task that can be divided up into relatively discrete and comparable chunks of work. The conflict of interest raised by within group comparisons, as well as difficulties in fully anonymizing work, can be solved by extending the population of reviewers beyond a particular organization. Reviewers, for some tasks at least, could be recruited through Amazon's mechanical turk, or other crowdsourcing resources, could be partly automated or could be identified from comparable external organizations. For instance, decision-making contributions in an online meeting from one part of an organization could be reviewed by members of a different part of that organization who lack the contextual knowledge to infer identity of participants.

Partial Illustrations in High Tech Firms

Some high tech firms make efforts to enforce a relatively flat organizational structure and work to minimize layers of management between engineers and top executives. Regardless of where this inclination comes from, this organizational structure can help reduce the tendency toward oligarchy to the extent that it enumerates items from a list based on theory and ideal type. Reportedly, companies like Google, 37signals, GitHub, Facebook, and Valve make efforts to keep their organizations relatively flat (Fried 2011). In particular, organizational structure described by the New Employee Handbook at Valve (2012) emphasizes the importance of a flat organizational system as well as other principles of worker autonomy and equality (see also Woffard 2012).

The New Employee Handbook describes several principles that are consistent with generating higher organizational fidelity but the list is partial, and some key issues such as the influence of reputation on assessment and the lack of double-blinding peer review are not addressed. The first principle is a fundamental commitment to role equality and the seemingly radical commitment to a flattened work environment "Welcome to Flatland" (Valve 2012, p. 4). In practice, this commitment implements some important features of a distributed organizational control because by flattening the work groups Valve eliminates many of the hierarchical asymmetries between manager and "managed" in terms of monitoring, sanctioning, and availability as the subject of monitoring and sanctioning. By working as peers in more freely flowing workgroups, particular employees do not have the same organizationally enabled capacity to exert favors or introduce biases and thus take important steps. Freedom of choice for selection of work is another foundational principle that makes the job much more attractive to creative workers, but also reduces leverage of holders of hierarchical positions, at least to the extent that multiple promising projects exist for a particular employee that flexibility reduces the coercive power of those in leadership positions.

Valve implements an extensive system of peer review, participation in which seems to be universally mandated. All employees are obligated to review some number of their peers, and some of their peers must evaluate them (Valve 2012, p. 25). This universal participation in evaluation takes an important step toward distributed organizational control. However, this process is not blinded, so it will likely be subject to the same reputational distortions and constraints that conventional review processes are. Also, to the extent that targets of review are self selected by reviewers, such a process may lead to greater "winner take all" effects than would a more regimented system of review. There seems to be a risk that those with more easily visible work or with bigger reputations would end up with inflated reviews while more obscure contributors are likely to have their work undervalued. The Valve Employee Handbook is noticeably missing provisions for assessing quality of contributions without knowledge of reputation, and thus issues of coalitions, the "Matthew effect," and other reputation driven flaws in organizational fidelity are likely to arise.

The Employee Handbook mentions many perks for Valve employees: free laundry service, espresso drinks, food, massage, and company trips (Valve 2012, p. 19). All of these benefits are collective goods that members enjoy as employees, but they are primarily valuable in terms of reducing transaction costs, saving time, and allowing employees to focus more of their time and energy on their work. Time is more valuable than money to the intrinsically motivated, and an employer who provides these perks makes the employees more strongly dependent on their employer not simply for the job but for the well-being that they enjoy from not having to spend focus on everyday hassles of life.

Taking Valve as an example, a truly double-blind system of review may be difficult to implement. Some aspects of creative work will be uniquely identifiable, and some aspects of publicly identifiable accomplishments are important to reward in the organizational compensation system. However, all workers will also make contributions to smaller dimensions of work, to decision making, to meetings, and to other less distinctive contributions. To the extent that these important but less distinctive contributions can be performed through digital means, they could represent a distinct data stream for generating an independent source of performance evaluation. One can imagine a firm like Valve implementing a split system of evaluation that maintains their current peer stack assessment system (with its known flaws related to reputational effects) while introducing a second stream based on a double-blinded measurement system of digital contributions. Combining these data streams to predict appropriate compensation would help identify where deviations between the measures occur. Deviations would arise both because of differences in the types of work contributions that are measurable by both and also to instances where reputational effects are distorting organizational fidelity.

Future Implementations for Improving the Fidelity of Organizations

The potential fidelity of an organization can be improved through distributed social control. Many design features of contemporary organizations undermine fidelity by exaggerating problems of agency, magnifying reputational and personalistic biases, and excluding most organizational members from exerting formal monitoring and evaluation processes. We can see that some recently developed organizations in the technology field have adopted flatter organizations, some aspects of peer review, and embraced principles of more open collaboration. These have shown some success in quality of work, creative productivity, and, to some extent, in the perceptions of legitimacy of the organizational systems. However, all of the example organizations omit key principles outlined in the ideal typical model of distributed organizational control.

Limited and partial implementations that we can observe in extant organizations can only provide limited insights into the operation of distributed organizational control. In the future, the best insights will require experimental tests of alternative designs of full systems within units of larger firms, in the context of social control systems in online games, in distributed collaborative projects, or in classroom environments.

Another productive direction for development would be to design a digitally mediated interface for meetings that would be "better than being there" and would include design features derived from the model of distributed social control, such as double-blinded review, equally distributed responsibility for evaluation and sanctioning, etc. Meetings are widely understood to be necessary, but also incredibly inefficient at advancing the organizational mission. We also know that meetings are very constrained by limitations of status, hierarchy, and reputation effects. A better interface for organizational decision making that made meetings better would be a tool that could be developed and implemented, and in so doing, it would spread examples of types of organizational principles that decrease the tendency toward oligarchy and increase organizational fidelity.

Another productive direction for implementation would be to introduce distributed organizational control as a secondary data stream for an existing review and evaluation system and to attach to it a separate stream of compensation in an organization. For instance, in an academic course, tools like CrowdGrader could be used to assign a portion of the points from assignments created, managed, and evaluated by the students themselves. Half of their grade could depend on a traditional set of assignments and the other on the set created by the students. Similar innovations could be explored in voluntary associations where members of organizations could experience both traditional and distributed systems of organizational control.

Implementation of the principles of distributed organizational control will require testing in small settings and in settings where the institutional hierarchy is not threatened by the loss of privilege that elites in the system enjoy. Certainly, the many deans of your typical university would not agree to be eliminated by a policy of distributed organizational control even if their roles of oversight and management could be organized with higher fidelity by a distributed network of faculty and administrators. Such a change would only happen once greater organizational efficiency became a necessity and if that faculty already had experiential insight into distributed organizational forms. This experience could come in their courses, in research groups, or within their departments but only if researchers develop accessible tools for those principles in circumstances that help them solve their existing problems.

Implications of Implementing Distributed Organization Control

What are the expected implications of implementing a system of distributed social control where all agents are subject to universal system of monitoring and sanctioning that captures accurate records of all of their organizational contributions, including their contributions to organizational decision making? A system where their organizational rewards are dependent on their evaluation in that system? A system where the evaluations they receive, as well as those that they make, are not biased by knowledge of identities or attributes of others. They evaluate as fairly and accurately as they can because their own evaluations are subject to blinded peer review according to the values of the organization and established, shared criteria of evaluation?

First, we expect a higher level of social order. There would be less corruption, less malfeasance, simply because of the presence of effective monitoring and sanctioning, combined with the knowledge that any such acts would generate sanctions delivered by the distributed evaluations of their peers. Second, we should expect that a higher percentage of agents' organizational actions would be productive work consistent with the mission of the organization. Third, we expect that opportunities to exert corrupt influence based on personalistic ties, favoritism, or other biases would be reduced. Fourth, agents who know that their work is being evaluated according to organizational criteria, that they themselves apply to the work of others, and whose work they are obligated to review will be likely to see resulting evaluations of their own work as legitimate. Finally, agents working in such an environment and implementing their part of the distributed system of monitoring and sanctioning would likely come to see such a system as legitimate, more so than conventional systems to which they are a part of. This experiential dimension could increase the demand for the practices and principles of distributed social control in other organizational systems. This effect is much like the existing demand among high tech workers who, having experienced effective peer collaboration systems, have developed a strong preference for that structure and who may be deeply skeptical of executives who impose hierarchy and demand compliance without the behavioral foundation of equality that those workers experienced in their work groups.

Digital institutions do not automatically create equality, justice, or efficiency, but depending on how they are structured, digital systems of interaction and social control offer new opportunities to address long-standing organizational problems (Kollock and Smith 1996; Glaser and Ebersbach 2004; Suh 2008). We have new opportunities to design our institutions in ways that have never been possible before and to implement them at unprecedented scales. While small-scale, localized solutions to social dilemmas have been documented (Keohane et al. 1993), digital institutions that implement distributed organizational control can address the corruption problem of organizational oligarchy that has long been recognized, but not fully addressed (Michels 1915; MacLennan 2005; Mason and Misener 2006).

Conclusion: We Need High-Fidelity Digital Organizations as Alternatives to Our Institutional Antiques

This chapter began by listing technological advances in science and engineering. If you watch enough TED talks, it would seem as if our social problems will be solved through technical means (we need yet another more efficient solar panel or a laser that zaps mosquitoes). However, I would contend that our most-limiting shortcomings are imposed by our social systems and that what we really need to work on is the development of digital institutions that create systems of distributed organizational control.

The greatest challenges faced by human societies will require unprecedented levels of cooperation and contribution. We need effective, high-fidelity organizations to overcome major new challenges. Succeeding on those challenges will not be possible while relying on institutional antiques and organizations with low-organizational fidelity. The problems inherent to current organizations are cast in stark relief when we seriously consider theoretical models of agency, social control, reputations, and biased expectations. Similarly, digital systems for interaction are already being developed that allow for structures of interaction that are not possible in face-to-face systems, and these digital institutions can enact double-blinded peer review, efficient systems of monitoring and sanctioning, as well as distributed systems that allow all members to participate in the exercise of social control. We need social scientists and computer scientists to team up to design and promote new digital institutions that overcome the many pitfalls of earlier hierarchical designs.

We need to keep our lofty ambitions, but we need to start small and think about building systems that can be used as tools within existing organizations. We can see glimpses of organizational innovations (Wikipedia, Reddit, Google, Valve, CrowdGrader), but we need to be thoughtful and integrative to build systems that more completely address the problems that reduce the fidelity of organizations. There are promising opportunities to experiment with distributed social control in online communities, online gaming systems, academic courses, and with specialized tools like computer-mediated systems that make key aspects of organizations (like meetings) work more effectively than they do in person. Cultivating experience and success in these limited areas will provide a baseline for integrating the same types of practices into larger and more socially influential organizations.

We need a global distributed effort, a space race for high-fidelity organizational design. The organizations that can effectively solve the problem of oligarchy will provide much more fair and equal work environments for their members, they will more accurately reward the productivity of all members, and they will minimize opportunities for those in leadership roles to turn organization into oligarchy. High-fidelity organizations should also enjoy substantial comparative advantages, such that eventually they simply outcompete the institutional antiques. But we will not actually see that better world unless we develop new institutions that let us work together better.

References

Anand, V., Ashforth, B. E., & Joshi, M. (2004). Business as usual: The acceptance and perpetuation of corruption in organizations. *The Academy of Management Executive, 18*(2), 39–53.

Arrow, K. J. (1984). *The economics of agency*. Stanford University Institute for Mathematical Studies in the Social Sciences, Stanford University, Technical Report, DTIC.

Berger, J., Cohen, B. P., & Zelditch, M. Jr. (1972). Status characteristics and social interaction. *American Sociological Review, 37,* 241–255.

Carzo, R. Jr., & Yanouzas, J. N. (1969). Effects of flat and tall organization structure. *Administrative Science Quarterly, 14*(2), 178–191.

Coleman, J. (1990). *Foundations of social theory*. Harvard: Harvard University Press.

Coleman, E. A., & Steed, B. C. (2009). Monitoring and sanctioning in the commons: An application to forestry. *Ecological Economics, 68*(7), 2106–2113.

de Alfaro, L., & Shavlovsky, M. (2013). CrowdGrader: Crowdsourcing the evaluation of homework assignments. *arXiv preprint arXiv:1308.5273*.

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *Academy of Management Review, 14*(1), 57–74.

Forbes Magazine. Dear CEO: They're just not that into you. *Forbes*. http://www.forbes.com/2011/02/02/flat-organization-structure-leadership-ceonetwork-ceo.html. Accessed 6 Dec 2013.

Frank, R., & Cook P. (1996). *The winner-take-all society: Why the few at the top get so much more than the rest of us*. New York: Penguin Books.

Fried, J. (2011). Why I run a flat company. *Inc. Magazine,* Originally printed in April 2011. http://www.inc.com/magazine/20110401/jason-fried-why-i-run-a-flat-company.html/1.

Glaser, Anja, & Ebersbach, M. (2004). Towards emancipatory use of a medium: The wiki. *International Journal of Information Ethics, 2,* 1–9.

Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Stanford University Press.

Gould, R. V. (2000). Revenge as sanction and solidarity display: An analysis of vendettas in nineteenth-century Corsica. *American Sociological Review, 65*(5), 682–704.

Greenwald, A. G., & Krieger, L. H. (2006). Implicit bias: Scientific foundations. *California Law Review, 94*(4), 945–967.

Hechter, M. (1988). *Principles of group solidarity*. Berkeley: University of California Press.

Hechter, M. (1990). The attainment of solidarity in intentional communities. *Rationality and Society, 2*(2), 142–155.

Heckathorn, D. D. (1989). Collective action and the second-order free-rider problem. *Rationality and Society, 1*(1), 78–100.

Hirschi, T. (1986). On the compatibility of rational choice and social control theories of crime. In D. B. Cornish & R. V. Clarke (Eds.), *The reasoning criminal: Rational choice perspectives on offending* (pp. 105–118). New York: Springer-Verlag.

Hirschi, T., & Gottfredson, M. (1986). The distinction between crime and criminality. In T. F. Hartnagel & R. A. Silverman (Eds.), *Critique and explanation: Essays in honor of Gwynne Nettler* (pp. 44–69). New Brunswick: Transaction.

Hirschi, T., & Gottfredson, M. (1987). Causes of white collar-crime. *Criminology, 25*(4), 949–974.

Hundley, K., & Taggart, K. (2013). Above the law: America's worst charities. *CNN U.S. Edition,* 13 June 2013. http://www.cnn.com/2013/06/13/us/worst-charities/.

Huther, J., & Shah, A. (2000). *Anti-corruption policies and programs: a framework for evaluation* (Vol. 2501). World Bank-free PDF.

Jones, G. R. (1983). Transaction costs, property rights, and organizational culture: An exchange perspective. *Administrative Science Quarterly, 28,* 454–467.

Keohane, R, Ostrom, E., & McGinnis, M. (1993). Linking local and global commons: Monitoring, sanctioning, and theories of self-organization in common pool resources and international regimes. *Proceedings of a Conference on Linking Local and Global Commons* (pp. 1–15).

Kiser, E. (1999). Comparing varieties of agency theory in economics, political science, and sociology: An illustration from state policy implementation. *Sociological Theory, 17*(2), 146–170.

Kiser, E., & Cai, Y. (2003). War and bureaucratization in Qin China: Exploring an anomalous case. *American Sociological Review, 68*(4), 511–539.

Kiser, E., & Hechter, M. (1991). The role of general theory in comparative-historical sociology. *American Journal of Sociology, 97*(1), 1–30.

Kiser, E., & Sacks, A. (2011). African Patrimonialism in historical perspective assessing decentralized and privatized tax administration. *The ANNALS of the American Academy of Political and Social Science, 636*(1), 129–149.

Kiser, E., & Schneider, J. (1994). Bureaucracy and efficiency: An analysis of taxation in early modern Prussia. *American Sociological Review, 59,* 187–204.

Kollock, P., & Smith M. (1996). Managing the virtual commons. In S. C. Herring (Ed.), *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives* (pp. 109–128). Amsterdam: J. Benjamins

Kunz, A. H., & Pfaff, D. (2002). Agency theory, performance evaluation, and the hypothetical construct of intrinsic motivation. *Accounting, Organizations and Society, 27*(3), 275–295.

MacLennan, C. (2005). Corruption in corporate America: Enron–before and after. In D. Haller & C. Shore (Eds.), *Corruption: Anthropological perspectives* (pp. 156–170). London: Pluto

Mason, DS, Thibault, L., & Misener, L. (2006). An agency theory perspective on corruption in sport: the case of the International Olympic Committee. *Journal of Sport Management, 20*(1), 52–73.

Measure to Venture. Why I run a flat company (Jason Fried). *Inc.com*. http://www.inc.com/magazine/20110401/jason-fried-why-i-run-a-flat-company.html/2. Accessed 6 Dec 2013.

Merton, R. (1968). *Social theory and social structure*. New York: Free Press.

Michels, R. (1915). *Political parties: A sociological study of the oligarchical tendencies of modern democracy*. New York: Free Press (Translated into English by Eden Paul and Cedar Paul; from the 1911 German source).

Owens, S. (2013). The battle to destroy Wikipedia's biggest sockpuppet army. *The Daily Dot,* 8 October.

Paltrow, S. J. (2013). *Special Report—The Pentagon's doctored ledgers conceal epic waste*. *Reuters,* 18 November 2013. http://uk.reuters.com/article/2013/11/18/uk-usa-pentagon-waste-specialreport-idUKBRE9AH0LP20131118.

Panchanathan, K., & Boyd, R. (2004) Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature, 432*(7016), 499–502.

Pearce, C. L., Manz, C. C., & Sims, H. P. Jr. (2008). The roles of vertical and shared leadership in the enactment of executive corruption: Implications for research and practice. *The Leadership Quarterly, 19*(3), 353–359.

Prendergast, C., & Topel, R. H. (1996). Favoritism in organizations. *Journal of Political Economy, 104,* 958–78.

Prophet Motive. (1999). *The Guardian.com,* 2 March 1999. http://www.theguardian.com/theguardian/1999/mar/03/features11.g210.

Ross, S. A.(1973). The economic theory of agency: The principal's problem. *The American Economic Review, 63*(2), 134–139.

Shapiro, S. P. (2005). Agency theory. *Annual Review of Sociology, 31,* 263–284.

Shleifer, A., & Vishny, R. W. (1993). Corruption. *The Quarterly Journal of Economics, 108*(3), 599–617.

Suh, B., et al. (2008). Lifting the veil: Improving accountability and social transparency in Wikipedia with wikidashboard. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (6 April, pp. 1037–1040).

Tenbrunsel, A. E., & Messick, D. M. (1999). Sanctioning systems, decision frames, and cooperation. *Administrative Science Quarterly, 44*(4), 684–707.

Theobald, R. (1990). *Corruption, development, and underdevelopment*. Durham: Duke University Press.

Valve Corporation. (2012). *Valve: Handbook for new employees* (1st ed.). Bellevue: Valve.

Voss, K., & Sherman, R. (2000). Breaking the iron law of oligarchy: Union revitalization in the american labor movement1. *American Journal of Sociology, 106*(2), 303–349.

Washburn, J. (2008). *University, Inc.: The corporate corruption of higher education*. New York: Basic Books.

Weber, M. (1978). *Economy and society: An outline of interpretative sociology*. Berkeley: University of California Press.

Woffard, J. (2012). Is valve too good to be true? http://www.jeffwofford.com/?p=943.

Yamagishi, T. (1992). Group size and the provision of a sanctioning system in a social dilemma.

Zyglidopoulos, S. C., Fleming, P. J., & Rothenberg, S. (2009). Rationalization, overcompensation and the escalation of corruption in organizations. *Journal of Business Ethics, 84*(1), 65–73.

Chapter 9
Cultural Differences in Social Media: Trust and Authority

Mei Kobayashi

Culture is always a collective phenomenon, because it is at least partly shared with people who live or lived within the same social environment . . .

(Hofstede 1991).

However, culture can only manifest itself through the individual. . . . In search of (cultural) effects on online trust formation . . . analysis at the individual level is most appropriate, as online activities are individually oriented . . .

(Hitosugi 2011).

Introduction

Over millennia, humans developed the ability to communicate, nurture trust, and identify experts and authorities who could provide helpful information. Sociologists have studied these phenomena within and between different cultures (Hofstede 2003; Kroeber and Kluckhohn 1952). Most communication was face-to-face until the development of writing. Reliable courier services made possible relationships between people who had never met, but had a solid, concrete reason for communication, e.g., leaders, scholars, businessmen, pen pals.

The advent and proliferation of the internet has led to the birth of *virtual* relationships, in which two parties may "meet" in cyberspace (Haythornthwaite 2005; Preece and Maloney-Krichmar 2005). Each party may have very limited information about the party at the other end—much less information than in a face-to-face meeting. Often, information exchanged may be limited to text messages, documents or downloaded media. Establishment and continuation of relationships in cyberspace pose new sets of challenges in the absence of facial and body cues, or direct, real-time communication (Cummings et al. 2002; Jiang et al. 2009b).

Despite these challenges, the internet is seen as a means for developing new relationships or maintaining existing ones. The Pew Research Center's 2013 Spring

[1] www.pewinternet.org/Static-Pages/Trend-Data-%28Adults%29/Online-Activites-Total.aspx Access 14 Aug 2013

M. Kobayashi (✉)
IBM Research, 5-6-52 Toyosu, Koto-ku, Tokyo 135–8511, Japan
e-mail: mei@jp.ibm.com

E. Bertino, S. A. Matei (eds.), *Roles, Trust, and Reputation in Social Media Knowledge Markets,* Computational Social Sciences, DOI 10.1007/978-3-319-05467-4_9,

Tracking survey of residents of the United States found 85 % of male adults and 84 % of female adults use the internet[1]. The greatest per capita use was by 18–29 year olds (98 %) followed by 30–49 year olds (92 %), 50–64 year olds (83 %), and 65 years and older (56 %). A 2006 Pew study on online dating found: "*Some 31 % of American adults say they know someone who has used a dating website, and 15 % of American adults—about 30 million people—say they know someone who has been in a long-term relationship or married someone he or she met online*" (Madden and Lenhart 2006). An August 2013 Pew study found 72 % adults who go online use social networking sites, with 18 % of them using Twitter, an increase from 16 % a year ago (Brenner and Smith 2013). Facebook dominates the social networking domain, but Twitter is steadily gaining ground. Surprisingly, the top activity on the internet (91 % of users) was "*using a search engine to find information on the internet*", followed by "*send or read e-mail*" (88 %)[2]. Rounding out the top seven activities were: "*look for info on a hobby or interest*" (84 %), "*search for a map or driving directions*" (84 %), "*look for info on a product or service thinking of buying*" (78 %), and "*get the news*" (78 %).

The growth in cyberspace activity has led to the need for a quick and reliable means for assessing the trustworthiness and level of expertise (or authority) of people and documents/information in cyberspace (Tidwell and Walther 2002). Moreover, "*There is a need for a high-level, abstract way of specifying and managing trust, which can be easily integrated into applications and used on any platform*" (Grandison and Sloman 2001). Below we discuss some scenarios in which measurement of trust and authority significantly improve the outcome:

e-Commerce is impacted by on-line perceptions. Studies have shown that a good on-line reputation can improve sales prices in on-line auctions and sales, even when the seller is undeserving of a highly positive reputation (Brown and Morgan 2006; Guertler and Grund 2006; Klewes and Wreschniok 2010; Livingston 2005).

Targeted Marketing by search engine, portal, and social network sites exploits visitor data to identify users who are likely to buy a product or service. Since many ad contracts are pay-per-purchase, increasing the rate of purchases per display is critical for revenue. Further improvements to internet site monetization may require a deeper analysis of the visitors and their cultural background, e.g., degree of cultural assimilation for immigrants (Chau et al. 2002).

Cybersecurity is a major application area for ratings of trust and reputation. Accurate algorithms for evaluating trustworthiness and authority (with very low false positives and false negatives) will be critical for their public acceptance and successful adoption. Since studies have shown the existence of demographic (age, gender, race) and cultural differences in internet use and practices—even within the same country[3] (Kobayashi 2012; Recabarren et al. 2008), a one-size-fits-all approach for identifying suspicious behavior that triggers alerts and leads to unwarranted close monitoring of an individual or party will not work. In addition to wasting resources,

[2] www.pewinternet.org/Trend-Data-%28Adults%29/Online-Activites-Total.aspx Access 14 Aug 2013

[3] www.pewinternet.org

it may lead to serious consequences, e.g., claims of breaches in personal privacy and racial profiling.

Crowdsourcing Data and social networks are being considered by some governments for efficient dissemination of critical information following natural or man-made disasters[4,5,6] (Yeomans 2012). Some prototypes and products have been deployed, e.g., *CrisisTracker*™ (Rogstadius et al. 2013), *NewsBrief*™[7] (Best et al. 2005), *Sahana*[8], *TweetTracker*™(Kumar et al. 2011), *Twitcident*™ (Abel et al. 2012), *Ushahidi*™[9], and *VirtualAgility*™[10] (Yin et al. 2012). Crowdsourcing is also being considered for gathering accurate, real-time traffic data to reduce congestion and emissions. Identification of trustworthy, influential people with many links to trustworthy and efficient disseminators of information will be the key to success. These people must be trusted by governments as well as ordinary citizens. To be of practical value, the yardstick for measuring trust must account for diverse cultural values and beliefs.

Online Knowledge Markets for question answering can be traced to crowd sourcing sites, such as *Quora*™[11], *stackoverflow*™[12], and the now defunct *Aardvark*™. In these marketplaces, people with good answers to questions will be rewarded with an increase in esteem and reputation (Hicks 2011). More recently, sites that require payment in real or virtual[13] currencies have emerged.

The primary goal of this position paper is to show that development of accurate, reliable, and automated systems for evaluating trust and authority of people and information (including documents) in cyberspace requires an understanding of cultural differences of parties involved in the creation, dissemination, and consumption of information. The remainder of this paper is organized as follows. The next section surveys works on cultural differences in cyberspace behavior, specifically: use of the internet (including blogs, messaging services and social network services), assessment online document quality, personal expertise, and online trust. The paper concludes with a discussion on open data, with a focus on Wikipedia.

Cultural Influences on Cyberspace Behaviors and Perceptions

Cultural influences on cyberspace perception and behavior is a topic that is receiving attention by scholars, businesses, and governments (Simon 2001). It has become

---

[4] FEMA, US: www.fema.gov/social-media,www.fema.gov/blog-newsroom-videos-photos
[5] Council of Australian Govts, nat'l. strategy for disaster resilience: www.coag.gov.au/node/81
[6] University of Colorado at Boulder, National Hazards Center: www.colorado.edu/hazards/
[7] EMM NewsBrief: press.jrc.it
[8] Sahana software: sahanafoundation.org
[9] Ushahidi: www.ushahidi.com
[10] VirtualAgility WorkCenter: www.virtualagility.com
[11] www.quora.com
[12] stackoverflow.com
[13] Experts Exchange: www.experts-exchange.com/

common practice for people to run an internet search to find more information about someone they just met or a business they recently heard. The practice of "*checking up*" on someone or something seems to have become socially acceptable and regarded as being "*akin to asking . . . friends about this fellow (or business)—offhand, sociable and benign*" (Cohen 2002). As a consequence, on-line reputation management (personal, professional, corporate, and institutional) has become a lucrative new profession (Bilton 2011).

However, the impression about a person or business based on information in cyberspace may not necessarily be accurate, complete, or unbiased. People from some cultures may be more apt to post personal information, while others may have very strong views regarding personal privacy (De Angeli 2009). Whether first impressions in cyberspace are accurate in the long run, and how the accuracy compares with first impressions in the real world are important questions. (Vazire and Gosling 2004) examined whether interpersonal perceptions in the real-world are mirrored in cyberspace, and they identify some correlations. Their work extends that by (Kenny 1994) who studied perceptions in interactions between two people in the real-world. In an earlier study, (Gosling et al. 2002) found simple physical data (e.g., tidiness, physical objects related to hobbies, and décor in offices and bedrooms) correlates well with real-world personality traits.

Understanding how cultural differences in the real-world extend into the cyber domain is an important first step towards the design of systems for international cyberspace users. Below is a brief review of works on cultural differences in: internet use; evaluation of document quality, identification of experts on a subject; and views of trust. It is intended to serve as a starting point for discussion and is by no means comprehensive.

Internet and Social Network Sites (SNSs)

Popular sites and services on the internet vary widely according to geographic regions. Quite arguably, *Google*™ has become synonymous with *internet search* in the United States. According to Alexa™[14],in September of 2013, the top 10 visited sites by internet users in the United States were: *Google*, *Facebook*™, *YouTube*™, *Yahoo!*™, *Amazon.com*™, *LinkedIn*™, *Wikipedia*™, *Twitter*™, *eBay*™, *Craigslist.org*™. However, in the People's Republic of China the most popular site is the search engine *Baidu*™, followed by the portal site *qq.com*™ (not *Yahoo!*) and the consumer to-consumer (C2C) marketplace *taobao.com*™ (not *eBay*). In fact, all of the top 10 most visited sites were Chinese. South Korea's top 10 sites had only three non-Asian sites: *Google* (ranked third), *Facebook* (ranked fourth), and *YouTube* (ranked sixth).

Site ranking is just one of countless measures of user behavior on the Internet. (Graff et al. 2004; Li and Kirkup 2007) examined gender and cultural differences in

[14] www.alexa.com

internet use in the People's Republic of China and the UK. The Pew Research Center initiated the Internet & American Life Project starting 2000 to gather data and publish reports on "*issues, attitudes and trends shaping America... (with special focus on) the social impact of the internet*".[15] Results from this Pew project show that within the United States, internet associated practices differ according to demographics, such as: gender, age, ethnicity, and level of education.

Studies of cultural differences in user behavior on the internet are increasing, for example: (Recabarren et al. 2008) studied relationships between culture, subcultures and the internet; (Jackson et al. 2008) conducted studies on use of information technologies by Chinese and American students. (Lee et al. 2002; Young 2004) compared Korean and Japanese internet use; (Vitartas et al 2004) studied cultural influences on internet and media use among Thai and Australian students; (Akman and Mishra 2010) examined the influence of gender, age and income differences on internet usage in Turkey; (Ibrahim and Ibrahim 2006) studied cultural influences on internet usage by Malaysian students; and (Canton 2012) examined cultural differences in mobile phone use.

A major phenomenon in the evolution of cyberspace was the sudden emergence of Weblogs (or blogs) on personal websites. As blogs grew in popularity, studies showed differences in motivations of readers and writers of blogs in different countries. Within the United States, blogging was initially regarded as a platform for self-promotion by college-educated, white males for career advancement or advancement of some political agenda. As inexpensive or free, user-friendly blogging tools became widely available, a more diverse community of bloggers ensued, with a broader range of agendas. They wrote about anything from hobbies, clubs, and activities involving children and families to cooking or pets.

Motivations for blogging and blogging communities in other parts of the world are markedly different (Kobayashi 2012). In South Korea and Japan, typical bloggers contribute to several blogs, each on a different hobby or interest. Bloggers often use different pseudonyms for different blogs in the (mistaken) hope of protecting their anonymity. The proliferation of sophisticated mobile phones and their use in Seoul and Tokyo during long commutes on public transportation expedited the trend towards reading and writing blogs from personal smart phones. Microblogs, including *Twitter*™, became popular, given the small screen size of personal phones. The Japanese language is well-suited for microblogging since each word typically consists of a single or pair of characters; A Japanese microblog with no special abbreviations has more information than a typical English tweet.

In the United States, blogging fell out of favor as people began to find it too consuming of time and energy. The debut of microblogging may have hastened the decline. In contrast to blogging, microblogging can be done anywhere (e.g., while waiting for a friend, child, or an order at a café), and it does not require grammatical knowledge or strong writing skills. Friends can respond to short messages in (near) real-time. People seem to enjoy the real-time interaction that was absent in

[15] www.pewinternet.org/

conventional blogging. Social interactions and experience sharing in real-time via cyberspace became popular. Soon thereafter, social networking became the next big thing.

As recently as 2008, *mixi*™[16] was the most popular SNS in Japan, and the fifth most visited site by Japanese, while *Facebook*™ was primarily a United States-based phenomenon and was the fifth most visited site by Americans. The initial success of the SNSs is related to cultural preferences for the site design—*mixi* for Japanese and *Facebook* for Americans (Fogg and Iizawa 2008). (Hargittai 2007) examined why certain people joined a SNS, while others did not. The study only involved internet users in the United States and four SNSs: *Facebook*, *MySpace*™, *Xanga*™, and *Friendster*™. Among those who joined, gender, race, ethnicity, and parental educational background influenced the choice in SNS. In studies involving 423 *Facebook* users from 5 countries, (Vasalou et al. 2010) found that experience with the site and culture were the strongest factors influencing "*true commitment*" to use of the SNS. A survey paper, "*Overview of research on cross-cultural impact on social networking sites*", by (Vitkauskaite 2010), gives additional references up to 2010.

Jiang and de Bruijn (2013) considered three levels of cross-cultural social networking: "*individual* (e.g., self-presentation, privacy), *interaction* (e.g., networks, motives for use), and *consequential* (e.g., cross-cultural social capital)". They propose a strategy for researching cultural differences and their effects on social networking at these levels. Other studies on cultural differences in users of SNSs are: (Cho 2010), who examined differences in Korean, American, and Korean–American users of SNSs, and (Huang and Park 2013), who found cultural differences in profile photographs between Taiwanese and American *Facebook* users. Their findings supported earlier studies by (Kenny 1994; Gosling et al. 2002) that found an Asian preference for group-oriented, context-related photographs and Western preference for close-up photos with a clear view of the author's face.

Marcus and Krishnamurthy (2009) compared the human-computer interfaces of Japanese, Korean and American SNSs and found that cultural factors tend to influence the design and personalization of virtual spaces. A study by (De Angeli 2009) examined sixty *MSN*™ virtual spaces[17] of British and Chinese students. Less variation in design was found among the Chinese students than the British students. Chinese favored symmetrical layouts, designed new functions more frequently, and posted far less personal information (if any). The main conclusion was the layout and design of personal pages in SNSs are biased towards Western users. There is "*a big challenge and moral responsibility for the HCI community to find effective methods and techniques to address cultural differences of users in the design of technology. … Designing for a global population requires understanding differences and similarities between heterogeneous groups of people*".

[16] mixi.jp

[17] www.microsoft.com/presspass/newsroom/msn/factsheet/msnspaces.mspx Access 14 Sept 2013

In Asia, a two-year messaging service named *LINE*™[18] has attracted a strong following, particularly among young people (Osawa 2013; Pfanner 2013). In early September of 2013, it had 230 million users and had been installed on 71 % of *iPhones*™ in Japan. The service is free, and unlike many other providers, users can register using a pseudonym, and ads are not based on user data, mitigating privacy concerns. A significant portion of *LINE's* revenue comes from games, which can be downloaded for free. However, many users opt to pay for special features. Another revenue source and popular feature are cute stickers that can serve as emoticons. By September of 2013, the average day saw the purchase of over 1 billion stickers that were sent to friends and family. The owner, *NHN Corporation*™ of Korea, plans to expand services to European countries and the United States. Its recent debut in Spain, which involved the recruitment of high-profile celebrities, was extremely successful. It remains to be seen if *LINE* and its revenue-generating features will be as appealing to users outside of Asia.

Document Quality

A number of approaches have been used to evaluate the quality of documents in cyberspace. The most straightforward is to ask members of the question answering (Q&A) community to set up rules for posting and rating posted answers during a fixed period of time, as in *Yahoo! Answers*™[19]. (Adamic et al. 2008) found Q&A on the site covers a broad range of topics, however, much of the discussion lacked depth.

Another approach is link analysis. The HITS (Kleinberg 1999) and PageRank (Page et al. 1999) algorithms work well for this application. A modified version of HITS which uses ExpertiseRank was particularly good at finding good questions and good answers (Zhang et al. 2007). (Agichtein et al. 2008) conducted one of the first large-scale studies to identify high quality content in community-driven Q&A sites. They introduced a new framework and system to assess the quality of input in the question answering domain and conducted implementation studies that showed "*accuracy close to that of humans*"

In Korea, more people use Q&A sites than search engines to find information on the internet (Bonfils 2011; Nam et al. 2009; Sang-Hun 2007). Specifically, *Naver*™ Q&A service[20] dominates search in Korea, with 30 % of the population visiting the site every day. Reliable methods to evaluate the quality of answers that are language independent will have high practical value in Korea.

In other countries, most people use traditional search engines to find information in cyberspace, so evaluation of retrieved document quality is important (Zhou 2007). The most straightforward approach for evaluation is relevance judgment by users who

[18] Line: line.naver.jp/en/
[19] answers.yahoo.com, Access 7 Sept 2013
[20] www.naver.com, Access 7 Sept 2013

manually assess the relevance of retrieved documents with respect to their queries. Evaluation involves the use of metrics, such as *mean average precision* (MAP), i.e., the mean of the average precision scores for each query[21] (where *precision* is the fraction of retrieved documents that are relevant or useful) or *discounted cumulative gain* (DCG), a measure of the usefulness, or *gain*, of a document based on its position in the result list[22] (Manning et al. 2008). However, manual assessment is a cumbersome and unrealistic approach due to the number of users, the volume of data that is searched, and volume of results retrieved.

A more realistic approach for evaluating retrieval quality is to use online user-behavior *in situ*, rather than from a test collection (Radlinski and Craswell 2013). User behavior can be evaluated through direct observations, e.g., time elapsed until users click, position of clicks, downloads (Huang et al. 2011; Kelly and Teevan 2003; Wang et al. 2009). In another approach known as *interleaved evaluation*, users are shown "*combinations of results retrieved by different ranked retrieval algorithms, and observing which results users select from this combination*" (Radlinski and Craswell 2013).

Automated evaluation of e-document quality has become an important research area with high potential for growth as a service business (e.g., *ARiSA*™[23]). Studies have shown that consumer perceptions of products are strongly influenced by the quality of supporting documents (Smart et al. 1996). Until recently, the focus of e-document assessment was on software code since most text documents were not in electronic form. (Wingkvist et al. 2010) extended a system originally developed for evaluating software quality to evaluate supporting documentation.

Evaluation of e-documents requires some basic groundwork. (Juran 1998; Klein 2001) proposed some definitions for information quality; (Hargis et al. 2009) proposed characteristics that should be considered when examining documents for quality; (Wingkvist et al. 2010) proposed some Key Performance Indicators (KPIs); and several research groups have proposed frameworks for evaluating document quality (Arthur and Stevens 1992; Chidamber and Kemerer 1994; Naumann 2002; Stvilia et al 2007; Wang and Strong 1996). (Ge and Helfert 2007) and (Knight and Burn 2005) give overviews of the subject and review relevant works.

Klein (2002) conducted a survey to determine characteristics people associate with document quality and found five main categories: "*accuracy* (discrepancy, source/author, bias/intentionally false information), *completeness* (lack of depth, technical problems, missing desired information, incomplete when compared with other sites, lack of breadth), *relevance* (irrelevant hits when searching, bias, too broad, purpose of Website), *timeliness* (information is not current, technical problems, publication date is unknown), and *amount of data* (too much information, too little information, information unavailable)".

[21] en.wikipedia.org/wiki/Information_retrieval, Access 9 Sept 2013
[22] en.wikipedia.org/wiki/Discounted_cumulative_gain Access 8 Sept 2013
[23] www.arisa.se/index.php?lang = en Access 14 Sept 2013

Zhou and Croft (2005) proposed and implemented a document quality model for a search engine to improve the quality of retrieved documents. The HITS algorithm (Kleinberg 1998) is effective for finding authorities on a topic on the open internet. It uses link analysis (links between documents and links between people). Since authorities will author or cite high quality documents, the algorithm can be used to find them. (Naveed et al. 2011) examined the problem of retrieving tweets with quality information. Language-independent methods, such as HITS, can be used to assess document quality in cross-cultural, multilingual comparison studies.

Despite advances in document quality assessment, more work is needed to prevent the spread of incorrect information (Carr 2013), particularly in the age of microblogging. False and unfounded accusation of a student, Sunil Tripathi, as a suspect in the Boston Marathon bombings, led to a horrific nightmare for his family: "*Someone will tweet, then retweet, and completely unsubstantiated things can proliferate so rapidly and destructively*" (Shih 2013). Unlike traditional media, for which editors and owners can be held accountable, pinpointing the blame is difficult, if not impossible, when stories spread like wildfire through social media.

Authority: Finding Experts in Cyberspace

Automated identification and ranking of experts on a topic is a difficult problem that has been studied for over a decade (Maybury 2006). Early studies examined the problem of finding experts in an organization using e-mail and corporate intranet data (e.g., internal reports, internal projects, report chain data). Some organizations facilitate expert search by requiring employees to input their skill sets into their corporate intranet. Since compliancy may be low or incomplete (especially within large institutions with many new hires), alternative means are often needed. Analysis of email communications has been shown to yield good results since most people communicate via email with workplace colleagues (Balog and de Rijke 2006; Campbell et al. 2003). Many people find experts using information on the institution's intranet (Becerra-Fernandez 2000; Pohs et al 2001; Reichling and Wulf 2009; Vivacqua and Lieberman 2000). As mentioned above, the HITS algorithm (Kleinberg 1998) is effective for finding experts on a topic in cyberspace. An extensive survey of expert finding is Chap. 2 of a PhD thesis by Serdyukov (2009).

Trust and Reputation

Merrian-Webster Dictionary defines reputation as, "*overall quality or character as seen or judged by people in general, recognition by other people of some characteristic or ability, a place in public esteem or regard: good name*"[24]. A more modern

[24] www.merriam-webster.com/dictionary/reputation Access 11 Sept 2013

definition that is more applicable to on-line relationship is a definition proposed in a Ph.D. Thesis by Mui (2002) on Trust & Reputation: "*(trust is) a perception that an agent has of another agent's intents and norms*".

One of the earliest surveys on trust and the internet is (Grandison and Sloman 2001). Mui et al. (2002) review computational models of trust and reputation, and they propose that information from social and friendship networks is important in modeling trust. A recent book by Castelfranchi and Falcone (2010) begins with a review of several different definitions of trust, then critiques each, with provocative questions, such as, "*Is trust only about predictibility?*", "*Is trust only willingness for any kind of vulnerability?*", etc. The book concludes with a dialog on trust and technology. A survey by Artrz and Gil (2007) examines research on trust and the Web from four perspectives: policy-based trust, reputation-based trust, general models of trust, and trust in information resources. They note that human beings currently judge and evaluate the trustworthiness of information and people in the real-world as well as cyberspace. However, in the Semantic Web, agents and so-called "automated reasoners" will have to perform these tasks.

Abdul-Rahman and Hailes (2000) propose a simple model for trust in virtual relationships that accounts for social activities such as prior experiences, word-of-mouth, reputations of recommenders, and intransitivity of trust. They note that trust in cyberspace has important consequences for e-commerce. (Chow and Chan 2008) examine trust in social networks and information sharing. (Cutillo et al. 2009) discuss *Safebook*™[25], a privacy-preserving social network that exploits features of real-life trust. (Downward et al. 2001; Hitosugi 2009, 2011; Karvonen et al. 2000, 2001) examine trust across different cultures in online contexts.

Open Data: Wikipedia

A major hurdle to conducting studies on cultural differences in cyberspace behavior is the dearth of open data. Reliable comparison of cultural differences can only be made if the data comes from similar types of sources. Although limited datasets have become available in the US and Europe, few open datasets are available in other regions. Japan has been making progress following the 2011 disasters; Mistrust of the government by ordinary citizens (Onishi and Fackler 2011; Keshet 2013; Wetherell 2013) led to successful demands to open government data files. In July 2012, the Open Data Promotion Consortium was established (Tanaka 2013).

On a more positive note, on August 17, 2013, NHK news in Tokyo announced that twelve railway companies (including JR East and Tokyo Metro subway) will provide real-time information on trains in service as part of an open data project involving the Tokyo Metropolitan Government, the Transportation Ministry, and

[25] answers.yahoo.com, Access 7 Sept 2013

the University of Tokyo[26]. Information will include locations of trains and buses and levels of traffic congestion. The Japanese government's goal and hope is the open data project is will lead to the development of new business opportunities, including innovative new services for smart phone users, visually impaired persons, and the general public during emergencies and disasters. The aggregation of data from many transportation sources is important since most people need to know how to effectively use combinations of trains and buses, rather than just one line or mode of transportation.

In the United States, online communities have served as repositories of valuable open data sets. (Hansen 2007) reviews the development of these communities and the challenges they faced during their growth. Many communities created archives to help new members catch up with discussions that took place prior to their participation. Recently, a number of cities in the United States and Canada have made some of their data available to the public[27] (e.g., Chicago, San Francisco, Edmonton, Ottawa, Toronto, Vancouver[28]) for a variety of reasons, including government transparency, political accountability, and encouragement of technical innovation and civic participation[29]. The open data movement is also picking up in Europe[30]. The United Kingdom held the Open-data Cities Conference in April 2012[31] and published a 52-page white paper[32] two months later to opine on the benefits of open data (transparency, equality in the data market, enhancement of trust in public data, personalization of public services) and to elaborate on its policies.

Currently, one of the largest, international, well-documented, archived data set is the *Wikipedia*[33] dataset. It provides an excellent opportunity to study how authors, editors and readers behave in collaborative work in cyberspace (e.g., authoring and editing). Collaboration involves a number of important attributes, such as trust, respect for the scholarly authority of others, as well as distrust and disrespect (Brandes et al. 2009; de Alfaro et al. 2011; Wilkinson and Huberman 2007).

Several studies found cultural differences in the behavior of members of the *Wikipedia* communities (e.g., Callahan and Herring 2011; Hara et al. 2010; Pfeil et al. 2006). However, these studies have only examined a fraction of all of the communities (intra- and international), which can have very different characteristics even within the same country and language, e.g., communities associated with computer science vs. mathematics vs. literature (Halavais and Lackaff 2008, Iba et al. 2010). We believe

---

[26] NHK World, Tokyo rail companies unite in open data project, 17 Aug 2013: www3.nhk.or.jp/nhkworld/english/news/20130817_19.html (English) Access 18 Aug 2013 www3.nhk.or.jp/news/html/20130817/k10013841691000.html (Japanese) Access 19 Aug 2013

[27] www.data.gov/opendatasites

[28] data.cityofchicago.org, opensf.wordpress.com, data.edmonton.ca, data.ottawa.ca data.toronto.ca, data.vancouver.ca

[29] http://opencityapps.org/,http://citycampsf.govfresh.com/

[30] www.govdata.de/,www.data.gouv.fr/,www.dati.gov.it/, data.gov.uk/

[31] http://opendatacitiesconference.com/

[32] Open Data White Paper: Unleashing the Potential, #opendata @uktransparency, @cabinetofficeuk, Her Majesty's Government: data.gov.uk/sites/default/files/Open_data_White_Paper.pdf

[33] http://en.wikipedia.org/wiki/Main_Page

that a more comprehensive study of differences in *Wikipedia* authoring, editing, and evaluation by the readership is an important first step towards understanding and appreciating different practices in diverse cyberspace communities. The results will be important for the design of systems that rely on quantitative measures of trust and authority of people and documents in cyberspace.

Acknowledgements The author would like to thank organizers of the KredibleNet project for the opportunity to participate in a friendly and fruitful workshop with experts from diverse backgrounds.

References

Abdul-Rahman, A., & Hailes, S. (2000). Supporting trust in virtual communities. In Proceedings of the Hawaii International Conference on Systems Science 6(6):6007.

Abel, F., Hauff, C., Houben, G., Stronkman, R., & Tao, K. (2012). Semantics + Filtering + Search = Twitcident. Exploring information in social web streams. In Proceedings of the Hypertext Social Media, Wincosin, pp. 285–294.

Adamic, L., Zhang, J., Backshy, E., & Ackerman, M. (2008). Knowledge sharing and Yahoo Answers: Everyone knows something. In Proceedings of the WWW, Beijing, pp. 665–674.

Agichtein, D., et al. (2008). Finding high quality content in social media. Proceedings of the WSDM. www.mathcs.emory.edu/ eugene/papers/wsdm2008quality.pdf. Accessed 7 Sept 2013.

Akman, I., & Mishra, A. (2010). Gender, age and income differences in internet usage among employees in organizations. *Journal of Computers in Human Behavior, 26*(3):482–490.

Arthur, J., & Stevens, K. (1992). Document quality indicators: A framework for assessing document adequacy. *Journal of Software Maintenance, 4*:129–142.

Artrz, D., & Gil, Y. (2007). A survey of trust in computer science and the Semantic Web. *Web Semantics: Science, Services & Agents on WWW, 5*(2):58–71.

Balog, K., & de Rijke, M. (2006). Finding experts & their details in e-mail corpora. Proceedings of the WWW, Edinburgh, pp. 1035–1036.

Becerra-Fernandez, I. (2000). Facilitating the online search of experts at NASA using expert seeker people-finder. Proceedings of the PAKM. http://ceur-ws.org/Vol-34/becerra-fernandez.pdf. Accessed 11 Sept 2013.

Best, C., van der Goot, E., Blackler, K., Garcia, T., & Hornby, D. (2005). European media monitor—System description. European Commission, Brussels, Rep No. EUR 22173 EN.

Bilton, N. (2011). Erasing the digital past. *NY Times*, 3 April.

Bonfils, M. (2011). Search marketing guide to Naver, Korea's most popular search engine. Search Engine Watch May 11. http://searchenginewatch.com/article/2070244/Search-Marketing-Guide-to-Naver-Koreas-Most-Popular-Search-Engine. Accessed 7 Sept 2013.

Brandes, U., Kenis, P., Lerner, J., & van Raaij, D. (2009). Network analysis of collaboration structure in Wikipedia. Proceedings of the WWW, Madrid, pp. 731–740.

Brenner, J., & Smith, A. (2013). 72 % of online adults are social networking site users. Pew Report Aug 5. www.pewinternet.org/Reports/2013/social-networking-sites.aspx. Accessed 8 Aug 2013.

Callahan, E., & Herring, S. (2011). Cultural bias in Wikipedia content on famous persons. *Journal of the American Society for Information Science and Technology, 62*(10):1899–1915.

Campbell, C, Maglio, P, Cozzi, A, & Dom, B. (2003). Expertise identification using email communications. Proceedings of the ACM CIKM, Louisiana, pp. 528–531.

Canton, N. (2012). Cell phone culture: how cultural differences affect mobile phone use. CNN on-line: edition.cnn.com/2012/09/27/tech/mobile-culture-usage. Accessed 14 Aug 2013.

Carr, D. (2013). The Pressure to be the TV news leader tarnishes a big brand. *NY Times*, 22 April.

Castelfranchi, C., & Falcone, R. (2010). Trust theory: A socio-cognitive & computational model. New York: Wiley.

Chau, P., et al. (2002). Cultural differences in the online behavior of consumers. *Communications of the ACM, 45*(10): 138–143.

Chidamber, S., & Kemerer, C. (1994). A metrics suite for object-oriented design. *IEEE Transactions on Software Engineers, 20*(6):476–493.

Cho, S. E. (2010). Cross-Cultural Comparison of Korean & American Social Network Sites. PhD Thesis, Communication, Info & Library Studies, Rutgers University.

Chow, W., & Chan, L. (2008). Social network social trust, and shared goals in organizational knowledge sharing. *Information and Managementm, 45*:458–465.

Cohen, R. (2002). Is Googling OK? The Ethicist, *NY Times*, 15 Dec.

Cummings, J., Butler, B., & Kraut R. (2002). The quality of online social relationships. *Communications of the ACM, 45*(7):103–108.

Cutillo, L. A., Molva, R., & Strufe, T. (2009). Safebook: A privacy-preserving online social network leveraging on real-life trust. *IEEE Communications Magazine, 47*(12): 94–101.

de Alfaro, L., Kulshreshtha, A., Pye, I., & Adler, B. (2011). Reputation systems for open collaboration. *Communations of the ACM, 54*(8): 81–87.

de Angeli, A. (2009). Cultural variations in virtual spaces design. *AI & Society, 24*(3):213–223.

Downward, P., Pawlowski, T., & Rasciutte, S. (2001). Trust, trustworthiness, relational goods & social capital: A cross-country economic analysis. North American Associatiom Sports Economists Paper 11-10.

Fogg, B., & Iizawa, D. (2008). Online persuasion in Facebook and Mixi: A cross-cultural comparison. In Proceedings of the PERSUASIVE 2008, LNCS 5033, Springer, Berlin, pp. 35–46.

Ge, M., & Helfert, M. (2007). A review of information quality research—Develop a research agenda. Proceedings of the International Conference on Information Quality. http://mitiq.mit.edu/ICIQ/PDF/A%20REVIEW%20OF%20INFORMATION%20QUALITY%20RESEARCH.pdf. Accessed 15 Sept 2013.

Gosling, S., Ko, S., Mannarelli, T., & Morris, M. (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal on Personality and Social Psychology, 82*:379–398.

Graff, M., Davies, J., & McNorton, M. (2004). Cognitive style & cross cultural differences in Internet use & computer attitudes. http://www.eurodl.org/materials/contrib/2004/Graff_Davies McNorton.html. Accessed 14 Aug 2013.

Grandison, T., & Sloman, M. (2001). A survey of trust in internet applications. IEEE Communication Surveys & Tutorials. http://www.doc.ic.ac.uk/~mss/Papers/Trust_Survey.pdf. Accessed 13 Aug 2013.

Guertler, O., & Grund, C. (2006). The effect of reputation on selling prices in auctions, Paper 114, Mannheim University. http://www.gesy.uni-mannheim.de/dipa/114.pdf. Accessed 11 Sept 2013.

Halavais, A., & Lackaff, D. (2008). An analysis of topical coverage of Wikipedia. *Journal of Computer-Mediated Comm, 13*(2): 429–440.

Hansen, D. (2007). Knowledge sharing, maintenance, and use in online support communities. PhD Thesis, Department of Information, University of Michigan.

Hara, N., Shachar, P., & Hew, K. (2010). Cross-cultural analysis of the Wikipedia community. *Journal of American Society on Information Sciences, 61*(10):2097–2108.

Hargittai, E. (2007). Whose space? Differences among users & non-users of social network sites. *Journal of Computer-Mediated Communication, 31*(1). http://webuse.org/pdf/Hargittai WhoseSpaceJCMC07.pdf. Accessed 11 Sept 2013.

Haythornthwaite, C. (ed.). (2005). Special issue on computer-mediated collaborative practices. *Journal of Computer-Mediated Communnication, 10*(4). http://jcmc.indiana.edu/vol10/issue4/. Accessed 14 Aug 2012.

Hicks, J. (2011). The rise of the knowledge market. *Forbes*, 27 June.

Hitosugi, C. (2009). Effects of culture on online initial trust: Individual level analysis. PhD Thesis, University of Hawaii, Manoa.

Hitosugi, C. (2011). Trust on the web: the power of subjective norm across cultures. Proceedings of the Cross-Culture Resource Conference. http://claireh.weebly.com/uploads/5/2/9/9/5299882/hitosugi_sninonlinetrust_crosscultureconf_dec2011.pdf. Accessed 15 Sept 2013.

Hofstede, G. (1991). *Cultures and organizations—Software of the mind*. London: McGraw-Hill.

Huang, C.-M., & Park, D. (2013). Cultural influences on Facebook photographs. *International Journal on Psychology, 48*(3):334–343.

Huang, J., White, R., & Dumais, S. (2011). No clicks, no problem: Using cursor movements to understand and improve search. Proceedings of the ACM CHI Human Factors in Computer System, Vancouver, pp. 1225–1234.

Iba, T., Nemoto, K., Peters, B., & Gloor, P. (2010). Analyzing the creative editing behavior of Wikipedia editors. *Procedia—Social & Behavioral Sciences, 2*(4): 6441–6456.

Ibrahim, A., & Ibrahim, D. Z. D. (2006). Significance of cross cultural background on internet usage among university students. *Unitar E-Journal, 2*(2):39–49.

Jackson, L., et al. (2008). Culture, gender and information technology use: A comparison of Chinese and US children. *Computers in Human Behavior, 24*(6): 2817–2829.

Jiang, Y., & de Bruijn, O. (2013). A three-level approach to the study of multi-cultural social networking. Proc ACM HCI, LNCS 8029, Vol. 26, Springer, Berlin, pp. 365–374.

Jiang Y, de Bruijn O, & De Angeli A (2009b) The perception of cultural differences in online self-presentation. Proc INTERACT 2009, LNCS 5726, Springer, Berlin, pp. 672–685.

Juran, J. (1998). *Juran's quality control handbook*. New York: McGrall-Hill.

Karvonen, K., & Cardholm, L., Karvonen, S. (2000). Cultures of trust: A cross-cultural study on the formation of trust in an electronic environment. Proceedings of the Nord-Sec. http://www.tml.tkk.fi/Research/TeSSA/Papers/Karvonen/Karvonen_Cardholm_Nordsec_final.pdf. Accessed 15 Sept 2013.

Karvonen, K., Cardholm, L., & Karvonen, S. (2001). Designing trust for a universal audience: A multicultural study on the formation of trust in the internet in Nordic countries. Proceedings of the Universal Access in HCI. www.tml.tkk.fi/Research/TeSSA/Papers/Karvonen/designing-trust.pdf. Accessed 15 Sept 2013.

Kelly, D., & Teevan, J. (2003). Implicit feedback for inferring user preference: A bibliography. *SIGIR Forum, 37*(2):18–28.

Kenny D (1994) *Interpersonal perception: A social relations analysis*. New York: Guilford.

Keshet, I. (2013). In Japan, a portrait of mistrust. *NY Times*, 27 March.

Klein, B. (2001). User perceptions of information data quality: Internet and traditional text sources. *Journal of Computer Information Systems, 41*(4):5–15.

Klein, B. (2002). When do users detect information quality problems on the World Wide Web? Proceedings of the Americas Conference on Information Systems, Dallas, pp 1101–1103.

Kleinberg, J. (1998). Authoritative sources in a hyperlinked environment. Proceedings of ACM SODA. www.cs.cornell.edu/home/kleinber/auth.pdf. Accessed 11 Sept 2013.

Klewes, J., & Wreschniok, R. (Eds). (2010). *Reputation capital: Building and maintaining trust in the 21st century*. Berlin: Springer.

Knight, S., & Burn, J. (2005). Developing a framework for assessing information quality on the World Wide Web. *Informing Science Journal, 8*:159–172.

Kobayashi, M. (2012). Blogging around the globe. In Abraham A (ed.) *Computational social networks* (pp. 55–86). Berlin: Springer.

Kroeber, A., & Kluckhohn, C. (1952). *Culture: A critical review of concepts and definitions*. Cambridge: The Museum.

Kumar, S., Barbier, G., Abbasi, M., & Liu, H. (2011). TweetTracker: an analysis tool for humanitarian and disaster relief. AAAI Conference on Weblogs & Social Media, Barcelona, pp. 661–662.

Lee, Y., et al. (2002). A cross-cultural study on the value structure of mobile internet usage: Comparison between Korea and Japan. *Journal of Electronic Commerce Research, 3*(4):227–239.

Li, M., & Kirkup, G. (2007). Gender and cultural differences in Internet use: A study of China and the UK. *Computers & Education, 48*(2):301–317.

Livingston, J. (2005). How valuable is a good reputation? A sample selection model of internet auctions. *Review of Economics & Statistics, 87*(5):453–465.

Madden, M., & Lenhart, A. (2006). Online dating. Pew Internet & American Life Project Report Mar 5. www.pewinternet.org/Reports/2006/Online-Dating.aspx. Accessed 8 Aug 2013.

Manning, C., Raghavan, P., & Schuetze, H. (2008). *Introduction to information retrieval*. Cambridge: Cambridge University Press.

Marcus, A., & Krishnamurthi, N. (2009). Cross-cultural analysis of social network services in Japan, Korea, and the USA. Proceedings IDGD (Part of HCI), LNCS 5623, Springer, Berlin, pp 59–68.

Maybury, M. (2006). Expert finding systems. Mitre Tech Rep 06–1115. www.mitre.org/work/tech_papers/tech_papers_06/06_1115/06_1115.pdf. Accessed 9 Sept 2013.

Mui, M. (2002). Computational models of trust and reputation: agents, evolutionary games, and social networks. PhD Thesis, EECS, MIT.

Mui, M., Mohtashemi, M., & Halberstadt, A. (2002). A computational model of trust and reputation. Proceedings of IEEE Hawaii International Conference System Sciences, Hawaii, p. 188.

Nam, K., Ackerman, M., & Adamic, L. (2009). Questions in, knowledge iN? A study of Naver's question answering community. Proceedings of ACM Human Factors in Computer Systems, Boston, pp. 779–788.

Naumann. (2002). *Quality-driven question answering for integrated information systems*. Berlin: Springer.

Naveed, N., Gottron, T., Kunegis, J., & Alhadi, A. (2011). Searching microblogs: Coping with sparsity and document quality. Proceedings of ACM CIKM, Glasgow, pp. 183–188.

Onishi, N., & Fackler, M. (2011). In Fukushima nuclear plant crisis, crippling mistrust. *NY Times*, 13 June.

Osawa, J. (2013). Japan messaging app line expands aggressively outside of Japan. *Wall Street Journal*, 27 Aug.

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). *The PageRank citation ranking: Bringing order to the Web*. Stanford: Stanford University Press (http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf. Accessed 10 Sept 2013).

Pfanner, E. (2013). A Japanese social app contacts new shores. *NY Times*, 6 Sept.

Pfeil, U., Zaphiris, P., & Ang, C. (2006) Cultural differences in collaborative authoring in Wikipedia. *Journal of Computer Mediated Communications, 12*(1):88–113.

Pohs, W., Pinder, G., Dougherty, C., & White, M. (2001). The lotus knowledge discovery system: Tools and experiences. *IBM Systems Journal, 40*(4):956–966.

Preece, J., & Maloney-Krichmar, D. (Eds.) (2005) Special issue on online communities. *Journal of Computer-Mediated Communications, 10*(4). http://jcmc.indiana.edu/vol10/issue4/. Accessed 4 Sept 2013.

Radlinski, F., & Craswell, N. (2013). Optimized interleaving for online retrieval evaluation. Proceedings of WSDM. http://research.microsoft.com/pubs/179433/Radlinski_Optimized_WSDM2013.pdf.pdf. Accessed 9 Sept 2013.

Recabarren, M., Nussbaum, M., & Leiva, C. (2008). Cultural divide and the Internet. *Computers in Human Behavior, 24*: 2917–2926.

Reichling, T., & Wulf, V. (2009). Expert recommender systems in practice: evaluating semi-automatic profile generation. Proceedings of ACM CHI, Boston, pp. 59–68.

Rogstadius, J., et al. (2013). CrisisTracker: Crowdsourced social media curation for disaster awareness. *IBM Journal of Research & Development, 57*(5):4:1–4:13.

Sang-Hun, C. (2007). To outdo Google, Naver taps into Korea's collective wisdom. *NY Times*, 4 July.

Shih, G. (2013). Boston Marathon bombings: how Twitter and reddit got it wrong. The Independent. www.independent.co.uk/news/world/americas/boston-marathon-bombings-how-twitter-and-reddit-got-it-wrong-8581167.html. Accessed 17 Sept 2013.

Smart, K., Madrigal, J., & Seawright, K. (1996). The effect of documentation on customer perception of product quality. *IEEE Transactions on Professional Communications, 39*(3):157–162.

Stvilia, B., Gasser, L., Twidale, M., & Smith, L. (2007). A framework for information quality assessment. *JASIST, 58*(12):1720–1733.

Tidwell, L., & Walther, J. (2002). Computer-mediated communications effects on disclosure, impressions, and interpersonal evaluations. *Human Communication Research, 28*(3):317–348.

Vasalou, A., Joinson, A., & Courvoisier, D. (2010). Cultural differences, experience with social networks & the nature of true commitment in Facebook. *International Journal of Human–Computer Studies, 68*:719–728.

Vazire, S., & Gosling, S. (2004). e-perceptions: Personality impressions based on personal websites. *Journal of Personality & Social Psychology, 87*:123–132.

Vitartas, P., Sangkamanee, S., & Scott, D. (2004). Cultural influences on Internet and media use: A comparison of Thai & Australian student usage. *Australasian Journal of Business & Society Inquiry, 2*(2). pandora.nla.gov.au/pan/37232/200410270000/www.scu.edu.au/schools/socialsciences/ajbsi/archive.html#v2n2. Accessed 14 Aug 2013.

Vitkauskaite, E. (2010). Overview of research on cross-cultural impact on social networking sites. *Economics & Management, 15*: 844–848.

Vivacqua, A., & Lieberman, H. (2000). Agents to assist in finding hep. Proceedings of ACM CHI, Boston, pp. 65–72.

Wang, R., & Strong, D. (1996). Beyond accuracy: What data quality means to consumers. *Journal of Management Information System, 12*(4), 5–33.

Wang, K., Walker, T., & Zheng, Z. (2009). PSkip: Estimating relevance raking quality from web search clickthrough data. Proceedings of ACM KDD, Paris, pp. 1355–1364.

Wetherell, J. (2013). Is this Japan's year for open data? Feb 1. http://techpresident.com/news/wegov/23454/japans-year-open-data. Accessed 14 Aug 2013.

Wilkinson, D., & Huberman, B. (2007). Cooperation & quality in Wikipedia. Proceedings of WikiSym, Montreal, pp. 157–163.

Wingkvist, A., Ericsson, M., Loewe, W., & Lincke, R. (2010). A metrics-based approach to technical document quality. Proceedings of International Conference on Quality of Information & Communications Technology, Porto, pp. 476–481.

Yeomans, M. (2012). Social media's crucial role in disaster relief efforts. *The Guardian*, 6 Nov.

Yin, J., Lampert, A., Cameron, M., Robinson, B., & Power, R. (2012). Using social media to enhance emergency situation awareness. *IEEE Intelligent Systems, 27*(6):52–59.

Young, B. (2004). Diffusion and usage patterns of the Internet in Korea and Japan. *Development and Society, 33*(2):229–250.

Zhang, J., Ackerman, M., & Adamic, L. (2007). Expertise networks in online communities: structure and algorithms. Proceedings of WWW, Banff, pp. 221–230.

Zhou, Y. (2007) Retrieval performance prediction and document quality. PhD Thesis, CS, U Mass Amhearst.

Zhou, Y., & Croft, B. (2005). Document quality models for web ad hoc retrieval. Proceedings of ACM CIKM, Bremen, pp. 331–332.

Chapter 10
Convincing Evidence

Andrew Gelman and Keith O'Rourke

The rules of evidence as presented in statistics textbooks are not the same as the informal criteria that statisticians and practitioners use in deciding what methods to use.

According to the official rules, statistical decisions should be based on careful design of data collection, reliable and valid measurement, and something approximating unbiased or calibrated estimation. The first allows both some choice of the assumptions and an opportunity to increase their credibility, the second tries to avoid avoidable noise and error, and the third tries to restrict to methods that are seemingly fair. This may be fine for evaluating psychological experiments, medical treatments, or economic policies, but we as statisticians do not generally follow these rules when considering improvements in our teaching (Gelman and Loken 2012) nor while deciding what statistical methods to use.

Did Fisher decide to use maximum likelihood, because he evaluated its performance, and the method had a high likelihood? Did Neyman decide to accept a hypothesis testing framework for statistics because it was not rejected at a 5 % level? Did Jeffreys use probability calculations to determine that there were high posterior odds of Bayesian inference being correct? Did Tukey perform a multiple comparison analysis to evaluate the effectiveness of his multiple comparison procedure? Did Rubin use matching and regression to analyze the efficacy of the potential-outcome framework for causal inference? Did Efron perform a bootstrap of existing statistical analyses to demonstrate the empirical effectiveness of resampling? Do the authors of the textbook on experimental design use their principles to decide what to put in their books? No, no, no, no, no, no, and no. We do know some psychometricians who fit item response models to evaluate their exam questions, and this is one of the very few examples we can think of where statistics researchers are using statistical principles to make professional decisions. Gigerenzer et al. have decided to use percentages versus natural frequencies for better understanding of analysis while the analysis are done by the medical students or faculty.

A. Gelman (✉)
Columbia University, New York, NY USA
e-mail: gelman@stat.columbia.edu

K. O'Rourke
O'Rourke Consulting, Ottawa, ON, Canada

E. Bertino, S. A. Matei (eds.), *Roles, Trust, and Reputation in Social Media Knowledge Markets,* Computational Social Sciences, DOI 10.1007/978-3-319-05467-4_10,

Then, how, do we gain our knowledge as in how to analyze the data? This is a question that arises over and over as we encounter new sources of data that are larger and more structured than ever before. How do we decide to believe in the effectiveness of a statistical method? Following Gelman (2013), here are a few potential sources of evidence:

1. Mathematical theory (e.g., coherence of inference or asymptotic convergence);
2. Computer simulations (e.g., demonstrating approximate coverage of interval estimates under some range of deviations from an assumed model);
3. Solutions to toy problems (e.g., the comparison of Rubin (1981) of a partial pooling estimate for a test-preparation program in eight schools to the no pooling or complete pooling estimates);
4. Improved performance on benchmark problems (e.g., getting better predictions for the Boston Housing Data (Harrison and Rubinfeld 1978), an example that is loved by textbook writers in statistics and computer science);
5. Crossvalidation and external validation of predictions (see Vehtari and Ojanen 2012), as can be done in various examples ranging from education to business to election forecasting;
6. Success as recognized in a field of application (e.g., a statistical method that is used and respected by biologists, economists, or political scientists);
7. Success in the marketplace of software or textbooks (under the theory that if people are willing to pay for something, it is likely to have something to offer);
8. Face validity: whether the method seems reasonable. This can be a minimum requirement for considering a new method.

As noted by Gelman (2013), "None of these is enough on its own. Theory and simulations are only as good as their assumptions; results from toy problems and benchmarks don't necessarily generalize to applications of interest; cross-validation and external validation can work for some sorts of predictions but not others; and subject-matter experts and paying customers can be fooled. The very imperfections of each of these sorts of evidence and how they apply to different user populations and settings gives a clue as to why it makes sense to care about all of them. We can't know for sure so it makes sense to have many ways of knowing." Informal heuristic reasoning is important even in pure mathematics (Polya 1941).

There is also the concern that a statistical method will be used differently in the field than in the lab, so to speak—or, to give this problem a pharmaceutical spin, that a new method, approved for some particular class of problems, will be used "off-label" in some other setting. Rubin (1984) discusses concerns for recommending methods of analysis for repeated use by those (and often ourselves) with limited statistical expertise, limited resources, and limited time. To further complicate this, there has long been experimental evidence that optimal methods of information processing do not always lead to optimal human performance, and this varies by level of skill, incentives, and time pressure (Driver and Streufert 1969).

We may also wish to consider as how to choose between methods in a given application for ourselves, to recommend to colleagues of similar or different levels of technical skill, and to communities of users who are not full-time statisticians or quantitative analysts, i.e., how should we go about approving statistical methods for use in various applications by various users, making reasoned, critical choices. To do this we lean on background material, again following the model of the choice of medical treatments for use by various professionals or end users. Our primary objective is to maximize the rate of learning about the empirical application while minimizing the rate and magnitude of mistakes. These goals require a sort of metaevidence that is not captured by any single sort of inquiry. More generally, we have argued that stories, to the extent that they are anomalous and immutable, are central to building understanding in social science (Gelman and Basbøll 2013).

How do we build trust in statistical methods and statistical results? There are lots of examples but out of familiarity we will start with my (Gelman's) own career. My most cited publications are my books and my methods papers, but I think that much of my credibility as a statistical researcher comes from my applied work. It somehow matters, I think, when judging my statistical work, that I've done (and continue to do) real research in social and environmental science.

Why is this? It's not just that my applied work gives me good examples for my textbooks. It's also that the applied work motivated the new methods. Most of the successful theory and methods that my collaborators and I have developed in the context of trying to solve active applied problems. The methods have faced real challenges and likely have been appropriately tempered in some generally relevant ways. At the very least, any of our new methods are computationally feasible, have face validity, and solve at least one applied problem. In developing and applying these methods, we weren't trying to shave a half a point off the predictive error in the Boston housing data; rather, we were attacking new problems that we couldn't solve in any reasonable way using existing approaches.

That's fine, but in that case who cares if the applied work is any good? To put it in another way, suppose my new and useful methods had been developed in the context of crappy research projects where nobody gave a damn about the results? The methods wouldn't be any worse, right?

The mathematics does not care whether the numbers are real or fake. I have an answer to this one: If nobody cared about our results we would have little motivation to improve. Here's an example. A few years ago I posted some maps based on multilevel regression and post stratification of preelection polls to show how different groups of white people voted in 2008. The blogger and political activist Kos noticed that some of the numbers in my maps didn't make sense. Kos wasn't very polite in pointing out my mistakes, but he was right. So I want back and improved the model with the collaboration of my student Yair Ghitza (Ghitza and Gelman 2013). It took a few months, but at the end we had better maps—and also a better method (which was later published in the *American Journal of Political Science*). This all happened only because I and others cared about the results. Kos and other outsiders performed severe testing of our conclusions, which ruined the face validity of our claims, and then we went through a process of trying to get the representation (model) less wrong for this particular empirical problem. If all we were doing was trying to minimize mean squared predictive error, I doubt the improvements would've done anything at all. Indeed, it turns out that it is difficult to identify improvements in hierarchical models via cross-validated mean squared error, even with large sample sizes (see Wang and Gelman 2013, a paper that was developed as our attempt to understand our challenges in comparing different models for this and similar small-area estimation problems).

This is not to say that excellent and innovative statistical theory can't be developed in the absence of applications or, for that matter, in the context of shallow applications. For example, my popular paper on prior distributions for group-level variance parameters (Gelman 2006) came through my repeated study of the 8-schools problem of Rubin (1981), a dead horse if there ever was one. Hierarchical modeling is still unsettled, and it was possible to make a contribution in this field using an example without any current applied interest. In many cases, though, seriousness of the application, the focus required to get details right (the representation less wrong), was what made it all work.

As suggested earlier, one possible way to better cover all the issues and challenges for deciding on which statistical method is most credible is to borrow from the regulatory review and approval perspective that is brought to bear on deciding which medical treatments should be approved, for what purpose, by whom, in which circumscribed situations. The actual reason for statisticians and nonstatisticians choosing statistical methods in practice may be less interesting, the reason being largely more to do with perceived authority, sociology (peer group), psychology, and economics (skill level, access to software, and budget limitations of client) than any putative evidence. A regulatory perspective would attempt to set this aside, at least until the benefits and harms of methods have been assessed along with their relative value/importance and the real uncertainties about these are clarified. We are not literally suggesting that statistical methods be subject to regulatory approval but rather that this perspective can help us make sense for the mix of information available about the effectiveness of different research methods.

More recently, it has been argued that the truth comes from "big data" (see Hardy 2013, for a contrary view). We agree with the saying "data trumps analysis" but in practice it can be easier to work with small datasets that we understand rather than with large datasets with unknown selection biases. For example, in our analysis of home radon levels (Lin et al. 1999), we used national and state-level random sample surveys of about 80,000 homes (which sounds like a lot but is not that much when you consider that radon concentrations vary spatially, and there are over 3000 counties in the USA), ignoring millions of measurements that were collected by individual homeowners, buyers, and sellers, outside of any survey context. We suspect that if we had ready access to the large set of self-selected data, it would be possible to perform some analysis to calibrate with respect to the more carefully gathered measurements, and get the best of both worlds. In that particular example, however, we doubt this will ever happen, because there is not such a sense of urgency about the problem; our impression is that everyone who is worried about radon has already had their house measured. For other problems such as medical treatments (or, in the business world, social advertising), we suspect that much can be learned (indeed, is already being learned) by combining experimental measurements with larger available observational data (see, e.g., Kaizar 2011; Chen et al. 2013).

This article is appearing in a volume and its goal is to consider "a future agenda for theoretical or methodological research on authorship, functional roles, reputation, and credibility on social media." We do not have a clear sense of what this agenda should be, but we think it is important to recognize the disconnect between our official and unofficial modes of reasoning in statistics, and the many different sources of practical evidence we use to make professional decisions. A fruitful direction of future research could be the formalization of some of our informal rules, much in the way that Rosenbaum (2010) formalized and critiqued the well-known rules of Hill (1965) in epidemiology.

References

Chen, A., Owen, A. B., & Shi, M. (2013). Data enriched linear regression. http://arxiv.org/abs/1304.1837. Accessed 4 June 2014.

Driver, M. J., & Streufert, S. (1969). Integrative complexity: An approach to individuals and groups as information-processing systems. *Administrative Science Quarterly, 14,* 272–285.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis, 1,* 515–533.

Gelman, A. (2013). How do we choose our default methods? For the Committee of Presidents of Statistical Societies 50th anniversary volume.

Gelman, A., & Basbøll, T. (2013). When do stories work? Evidence and illustration in the social sciences. Technical report, Department of Statistics, Columbia University.

Gelman, A., & Loken, E. (2012). Statisticians: When we teach, we don't practice what we preach. *Chance, 25*(1), 47–48.

Ghitza, Y., & Gelman, A. (2013). Deep interactions with MRP: Election turnout and voting patterns among small electoral subgroups. *American Journal of Political Science, 57,* 762–776.

Hardy, Q. (2013). Why big data is not truth. *New York Times*. http://bits.blogs.nytimes.com/2013/06/01/why-big-data-is-not-truth/.

Harrison, D., & Rubinfeld, D. L. (1978). Hedonic prices and the demand for clean air. *Journal of Environmental Economics and Management*, *5,* 81–102.

Hill, A. B. (1965). The environment and disease: Association or causation? *Proceedings of the Royal Society of Medicine, 58,* 295–300.

Kaizar, E. E. (2011). Estimating treatment effect via simple cross design synthesis. *Statistics in Medicine, 30,* 2986–3009.

Lin, C. Y., Gelman, A., Price, P. N., & Krantz, D. H. (1999). Analysis of local decisions using hierarchical modeling, applied to home radon measurement and remediation (with discussion). *Statistical Science, 14,* 305–337.

Polya, G. (1941). Heuristic reasoning and the theory of probability. *American Mathematical Monthly, 48,* 450–465.

Rosenbaum, P. R. (2010). *Observational studies* (2nd ed.). New York: Springer.

Rubin, D. B. (1981). Estimation in parallel randomized experiments. *Journal of Educational Statistics, 6,* 377–401.

Rubin, D. B. (1984). Bayesianly justifiable and relevant frequency calculations for the applied statistician. *Annals of Statistics, 12,* 1151–1172.

Vehtari, A., & Ojanen, J. (2012). A survey of Bayesian predictive methods for model assessment, selection and comparison. *Statistics Surveys, 6,* 142–228.

Wang, W., & Gelman, A. (2013). Cross-validation for hierarchical models. Technical report, Department of Statistics, Columbia University.

Part V
Research Opportunities and Gaps in Trust, Credibility, and Authorship Research

Chapter 11
The Trajectory of Current and Future Knowledge Market Research: Insights from the First KredibleNet Workshop

Sorin Adam Matei, Brian Britt, Elisa Bertino and Jeremy Foote

Introduction

A. Background Increasingly, research among scholars and practitioners alike has become a collaborative enterprise (Schleyer et al. 2012). The fundamental logic, supported throughout the academic literature (see, for instance, Katz and Martin 1997), is that a group can produce a superior product than that which an individual can generate. Consequently, collaboration leads to better studies and more sophisticated products. These benefits are pushed to their limits when a collaborative group comprises a diverse assortment of experts who can uniquely contribute to any given endeavor. More ambitious visions, such as Weinberger's (2011), see the networked version of knowledge as the main avenue for future scholarship.

It is only natural, then, that numerous scholarship forms and software packages have been developed to support networked research collaboration. One need only search for collaborative software online to gain a sense of the wide range of efforts to support group work. Such resources, which stand alongside other communication technologies ranging from e-mail to video conferencing programs, were developed to support collaboration across scholarly and geographic domains, allowing groups of scholars to freely interact and develop insightful projects regardless of their particular disciplines or locales. Yet none of these collaborative packages have succeeded at attracting and retaining a sizeable population of researchers who consistently make use of the software. By and large, these systems, which were intended to form the coreof users' research activities, have moved into niche roles or been abandoned—e.g., FLOSSmole (2013), DataSift (2014), and Mendeley (2014).

The authors thank the participants to the first KredibleNet workshop for the valuable input offered to the final roundtable that constituted the starting point for this chapter.

S. A. Matei (✉) · B. Britt · E. Bertino · J. Foote
Purdue University, West Lafayette, IN, USA
e-mail: smatei@purdue.edu

E. Bertino
e-mail: bertino@purdue.edu

E. Bertino, S. A. Matei (eds.), *Roles, Trust, and Reputation in Social Media Knowledge Markets,* Computational Social Sciences, DOI 10.1007/978-3-319-05467-4_11,

A few more general collaborative tools have maintained relatively high usage, from lesser-known software packages like Basecamp (37signals 2013) and Alfresco (Alfresco Software 2013) to ubiquitous offerings such as Google Drive (Google 2013) and Microsoft SharePoint (Microsoft 2013a; see also Berman et al. 2012). Yet, with some exceptions focused on very narrow scientific collaboration, such as NanoHub (Klimeck et al. 2008), these are hardly tailored to researchers' needs, allowing them to transmit and develop documents with one another but compelling them to look elsewhere to collect and manage data, conduct analyses, and share their results within a larger scholarly community. Furthermore, social sciences, even those dealing with large datasets generated by social media, are woefully underrepresented in this arena.

It has become increasingly obvious that researchers, especially those with an interest in social behavior at large scale, would benefit from a computing infrastructure that actually facilitates academic research beyond the benefits of a mere file-sharing system. The most ambitious goal of the KredibleNet project is to develop such an infrastructure that would allow researchers interested in social behavior online to be part of a community and to more freely pursue their research tasks together within a single environment rather than dividing the workload among team members and between countless other software packages. The present chapter, which assesses current research practices and future trajectory for scholarship in order to identify the most critical needs for the new infrastructure, is a first step toward this objective.

B. Approach It would be virtually impossible to predict, with any degree of certainty, every single need for all researchers in the present and future. As such, the authors sought to get a broad picture of potential requirements for a collaborative infrastructure, rather than to identify individual details.

To this end, the authors used the first KredibleNet workshop on knowledge markets, functional roles, authority, and credibility hosted at Purdue University as an opportunity to develop this broad picture. While it was important to develop a set of features and characteristics for a system architecture that would be suited to researchers' needs, this could not be done without first thoroughly exploring current research paradigms and anticipated research directions. Considering that knowledge market research is a particularly broad scholarly domain spanning the boundaries of numerous academic departments, it was deemed a useful subject matter to use as a foundation for exploring the spectrum of present and future scholarly work. Therefore, the first round-table session focused on the most prominent current and future research topics and the most critical research challenges in four broadly defined research areas of interest to the workshop participants: reputation, trust, and authority; social influence and social structures; large data and data mining; and text mining and natural language processing.

It should be noted that this approach was not intended to develop a set of requirements to which a new infrastructure could be tailored, as that infrastructure would consequently overlook all other research needs that happened to escape the perspective of the workshop participants. Rather, the idea was to construct a research agenda that would be both sophisticated and flexible enough to handle the particular issues that the four workshop groups highlighted; in so doing, we could expect it to also be flexible enough to handle any other needs that emerge around and in between our groups on the scholarly spectrum. Knowledge markets in particular represent an especially useful topic area on which to base this kind of initial, foundational work, as scholars studying this area span a wide range of home disciplines, goals, and theoretical/methodological backgrounds and approaches. As such, it is easier to say that research on knowledge markets is reasonably representative of research across fields, as knowledge market research is inherently interdisciplinary, more so than most other academic and practical domains.

Thus, the results of the first KredibleNet workshop round-table sessions offer an especially wide picture of the research spectrum. Even without substantial depth in any particular area, the results provide a sufficient picture of current and future research to provide a litmus test for the flexibility and, therefore, utility of any proposed collaborative system.

The results of the groups' discussions are reviewed below, starting with the most prominent cutting-edge theories and methodologies in use today, transitioning to projections of the most promising research topics in the near future, and ultimately using those to identify the key theoretical, methodological, and organizational shortcomings that will need to be addressed in order for the research to move forward.

State-of-the-Art Theoretical Frameworks

During the first round-table session of the KredibleNet workshop, two broad areas emerged as critical for current research on knowledge markets: (a) the development of community content and individual reputations within such communities, and (b) competing theories about societal change and preferential attachment to explain why individuals who behave in similar ways and share similar viewpoints tend to connect through social, friendship, or work ties.

A. Content and Reputation Multiple groups were particularly concerned with reputations as they develop online, as well as the potential relationship between the perceived credibility of content and contributors within a knowledge market and its activity level relevant to other similar markets. As the reputation, trust, and authority workshop group noted, it is possible that, for instance, a popular article on Wikipedia may be especially well vetted. Presumably, the more editors contribute toward a given article, the more refined that article becomes. This echoes the famous dictum of large programming groups: "Given enough eyeballs, all bugs are shallow" (Raymond 1999, p. 30).

On the other hand, an article may instead be popular because it is controversial, so it may hold a constant state of flux as more, newer editors strive to add their own perspective to the article at hand. Thus, rather than the article being refined to approach an error-free state, its editors may instead play an endless game of tug-of-war over how to bias the article in a way that privileges their own respective viewpoints (Matei and Dobrescu 2011). As such, while there may be a relationship between activity and the quality of content produced, that relationship might not be strictly linear.

Likewise, authority figures within a community are often identified solely based on their amount of activity, with the most active community members deemed the leaders. Certainly, an individual with virtually no activity who has contributed almost nothing to the community cannot be considered a leader, but those with the most activity might be considered a type of functional leaders if a number of other conditions are met. The top authorities within a community generally do, indeed, contribute a lot, but to become a leader one also needs a broad involvement in the conversation. They should also be found among those frequent contributors who add fuel to the metaphorical fire in arguments rather than resolving the issues at hand. In short, the roles of leader and fire-stoker are easily confused, yet they need to be carefully considered and their complementarities identified (see Welser et al. 2007; Welser et al. 2011).

One such type of complementary approach is the "Professional of the Month" award launched by Microsoft to honor especially helpful contributors to its Technet technical support forums (Microsoft 2013b). The award attempted to counteract the tendency to determine authority based solely on contribution quantity. Rather than attempt to identify community leaders solely based on which individuals posted the most responses to user questions, as many other community award systems do, the Professional of the Month title was awarded based on whose contributions tended to be the very last in a given discussion. In other words, individuals who stepped in and definitively ended arguments over how to solve users' technical problems were deemed especially productive. Certainly, these individuals were often among those who posted the most messages in a given month, but rarely did they hold the very highest post counts in the work community for the month in which they won the award, as they tended to post only when necessary to help resolve a dispute rather than engaging in aimless back-and-forth arguments themselves.

Of course, that still leaves the so-called Matthew effect as the dark side of reputation. According to Merton (1968), those individuals who already carry strong reputations will tend to garner undue credit for further contributions, thus strengthening their reputations ever further and widening the gap between themselves and the relative unknowns—even if those lesser-known individuals are, in fact, making just as much of a difference through their efforts. This effect is closely related to the phenomenon of selective affiliation within many social networks wherein the most central nodes tend to further increase their centrality over time relative to others in the same network.

Online authorship and reputation are also related to the proper adherence to cultural norms. Unfortunately, websites tend to offer minimal cultural context for their users, making it very difficult to tailor one's behavior to any particular community. Such dynamics are usually quite clear when visiting another country and seeing how people behave, but online, visitors cannot necessarily observe the same social cues. Cultural mistakes online may degrade one's reputation, complicate personal relationships, or, on a national security level, even lead to needless investigations and privacy invasions. On the whole, Internet users could benefit from a greater understanding of the global landscape, which is an area in which scholarly research could greatly contribute.

Reputation's pervasiveness even extends to personal preferences and consumer behavior. Salganik et al. (2006) considered the underlying attitudes of individuals when they choose to buy, download, and listen to music, and questioned whether they did so because they actually liked the music in question, or because they saw others listening to the same things. In their study, Salganik et al. provided false feedback to consumers from others who were ostensibly "like you," and they found that more people downloaded and "liked" the songs with artificially inflated ratings. The music that was inherently superior but had not been socially validated was cast aside, both in terms of the number of times each song was downloaded and the average ratings that they received from research subjects. Thus, both purchasing behavior and personal opinions were decisively swayed by reputation, even beyond the inherent quality of the products in question.

B. Social Association and Change Societal change and social behavior was the second most prominent theme of research highlighted by the KredibleNet research community. For instance, the social influence and social structures workshop group was especially concerned with the causes of behavioral changes, particularly the roots of diffusion and of major societal movements. Why, they questioned, were people compelled to forward equal sign images supporting same-sex marriages when they saw them on Facebook? Regardless of their respective beliefs, why did some people choose to forward the images, personally spurring the movement forward, while others remained spectators?

A number of theories were considered to address this issue. One was that of the tipping point dynamic, which appears when masses of slacktivists who minimally contribute toward a given social movement reach a maximum diffusion moment in a social system (see, for instance, Bakshy et al. 2012; Tufekci and Wilson 2012; Valente 1995). However, this question might be better addressed by the emerging debate over the meaning of Christakis and Fowler's (2007) social influence findings. According to Christakis and Fowler, people influence their network "alters" to engage in similar actions, so certain individual characteristics such as obesity tend to spread through social networks. By and large, social influence theories have grown to hold a much stronger presence in the literature than the earlier message transmission models that are quite rudimentary by comparison (see, for instance, Cacioppo et al. 2009; Onnela and Reed-Tsochas 2010; Sobkowicz et al. 2012).

Despite the attractiveness of the social influence explanation, however, other scholars have cast doubt on Christakis and Fowler's work (see, for instance, McPherson et al. 2001; Rader and Wash 2008), contending instead that individuals who already share similar attributes are by design more likely to forge interpersonal connections with one another. Their affinity is determined more by essential characteristics than by mere communication. They contended that research on dynamic networks, along with data structures specifically developed to explore network changes over time, suggest that it is not the attributes of connected individuals that move toward one another over time, but instead that people who already share certain attributes tend to cluster together. Homophily or other types of social affiliation, such as complementarity, indeed, may trump mediated influence. Furthermore, we have historically relied primarily on demographics as a model of similarity, but in fact, these may not be the best measures of social affiliation. Collaborative work relationships, cohort effects, or cultural-social identification can be just as important factors in shaping social association.

In this context, it is worth remarking that homophily is by far the preferred concept for imagining social affiliation. However, association by complementarity and symbiotic relationships may also be important drivers of association and interaction. There are likely certain benefits that individuals may derive from interacting with others who are not exactly like them, but who are just similar enough to keep the interaction from merely becoming aggravating. There may be little point in associating with someone who is exactly like oneself; an alter with just enough differences may offer much more, such that connected individuals who are not exactly alike can learn and gain a great deal from one another. The question, then, is how individuals need to align across dimensions for their relationship to be complementary in this manner, and what dimensions are particularly salient for complementarity effects. In short, on what dimensions should connected individuals differ, and on which dimensions should they vary—or does even that vary across connected pairs and across social contexts?

The long-standing dichotomy of strong ties and weak ties may play its own role. Granovetter (1973), among others, argued that weak ties may be particularly important for diversifying one's own personal network and gaining access to a greater variety of populations and skill sets, as alters connected by weak ties generally have different features or talents. This may directly relate to the aforementioned complementarity, as pairs connected by strong ties may have to be similar across more relevant dimensions, or at least a different set of dimensions, than those which share weak ties.

Online social networks operationalize the strong versus weak tie dichotomy in ways which may or may not be accurate representations of the interpersonal phenomenon. Facebook social engineers treat the sheer amount of interaction between two parties as a measure of strength. This is likely the most straightforward measure of tie strength, but it would be reasonable to suspect that other tie characteristics may be just as important. As an example, it is quite possible that the types of interactions that two individuals share are more important than the mere quantity. If an individual maintains constant yet superficial communication with one alter, and has sporadic but deeply meaningful interactions with another alter, who is to say which tie is stronger? Yet again, if the strength of a tie is not defined by actual interactions, how can the external factors, uncaptured by the social medium, be defined? Quantity alone may be an insufficient operationalization, yet it has the lure of practicality. At the same time, although it is easy to measure, it may have limited utility for researchers hoping to gain a better understanding of tie strength, complementarity effects, social influence, and homophily (see also Gilbert and Karahalios 2009).

C. Summary Overall, two key themes developed as major topic areas for current knowledge market research: community content and individual reputation development, and competing theories about observed social association within such communities. For both, the key question was one of community change, whether in terms of the development of a strong content base and an individual's stature within that community or in terms of how societal change moves through a community based, at least in part, upon the ways in which individual members are connected with one another within the system.

Most Promising Research Topics and Approaches

Naturally, in developing new social practices and tools to support trust, authority, and credibility research in knowledge markets, it is insufficient to focus on the scholarship being conducted today. The KredibleNet community conversations were also directed toward projecting future research in order to obtain some potential theoretical domains and methodological approaches that a useful program must be able to support. During the workshop sessions, participants highlighted four primary areas toward which scholars are currently moving: advertising, textual commands and natural language processing, the integration of multiple data sources, and community design and feedback systems.

A. Evaluating Potential Research Topics Before the KredibleNet workshop participants could identify the most significant research challenges looming on the horizon, they first had to consider the topics that would likely hold the most currency across disciplines. With that said, the evaluation of promising future research topics developed into one of the most contentious questions of the entire workshop.

In some cases, the notion of such an evaluation itself caused considerable strife. The participants struggled over the very definition of the term "promising" or disputed whether it was more important to emphasize commercial value or societal impact, which was complicated by considerations about whether that societal impact would be tangible enough to be directly observed (such as campaigns designed to save lives) or if it might remain relatively imperceptible.

Some workshop participants were especially concerned about how to quantify the benefits of research with indirect outputs, as is the case for much of the social media literature. Aside from a few isolated examples, such as tracking the spread of diseases through social networks (Newman 2002) and aggregating data for the sake of coordinating disaster relief efforts (Gao et al. 2011; Huang et al. 2010; Jansen et al. 2009), such research is often directed toward developing a body of knowledge upon which other scholars and practitioners might build. In most cases, the tangible results of such studies are more indirect, with others using and building upon the initial social media research for their own end. The indirect nature of the output makes it difficult to resolve debates over which topics are more or less promising than others, yet such studies' usefulness to others who build upon the initial work in fact makes these foci especially critical areas for further development, as they can have wide-ranging effects that extend well beyond the confines of the initial work.

B. Advertising and Public Campaigns In terms of commercial value, the participants concluded that advertising revenue and public persuasion were the only worthwhile research topics that could be generalized across various knowledge markets. As one participant said, such work effectively aims to "take a manipulative message and deliver the highest level of manipulation possible to the largest number of people at the lowest possible cost." Regardless of one's personal attitudes about advertising research, any number of companies and businesspeople across regions and industries advertise online, whether or not their work is inherently connected with the Internet. Consequently, one could argue that advertising keeps businesses alive and running, allowing the general public access to greater commercial comforts. In short, even as detractors point out the psychological manipulation inherent in advertising, its advocates would herald it as a critical subject for society as a whole.

However, as one member of the social influence and social structures group noted, advertising research may not be as important in all online domains. Certainly, companies like Google and Facebook have long investigated various methods of incorporating others' advertisements into their own interfaces, but in terms of e-commerce, there are only a few major players who rarely even bother advertising. A consumer may go years without seeing an advertisement for Amazon.com, for instance, yet it remains one of the most well-known, dominant marketplaces on the Internet (Wohlsen 2013). At the same time, one can add that the Amazon.com Affiliate Program, which allows ordinary users to post links of Amazon products in exchange for a commission, is a form of viral advertising.

In the online realm, much smaller companies tend to disseminate the bulk of advertisements rather than the few corporations that already control the vast majority of the market share. Among these smaller companies are the "schemers," as some workshop participants described them, whose practices consist primarily of pestering Web consumers with pop-up windows, animated banners that obscure content, and any number of other unwanted distractions and intrusions. The major players in the online marketplace have no need to advertise as their respective target markets already know who they are.

With that said, this advertising dynamic may result from the structure of economic practices online, particularly as they allow a few companies to develop veritable monopolies simply by outspending others or by establishing themselves before competitors enter the marketplace. Even within the current system, advertising research may still help companies whose operation is not strictly online. For instance, advertisements for cars and movies have been common on many prominent websites, just as commercials and abbreviated trailers became a staple on television years ago. Thus, both the marketing professionals creating new advertising campaigns and the content-producing websites that funnel those advertisements to visitors stand to benefit from improvements to their current advertising approaches. So, too, might the consumers in the advertisers' respective target markets.

C. Textual Commands and Natural Language Processing In discussing text mining and natural language processing, participants focused their attention on the immediate future of text-to-cognition and textual command recognition research.

From a consumer standpoint, Siri, the personal assistant application developed for Apple computing devices, is likely the most prominent example of text-to-cognition technology. Such software does not require that the program actually understands human language, but merely that it be able to interpret language syntax well enough to parse out keywords and to connect phrases for the purpose of conducting online searches. In fact, workshop participants proclaimed that they were "not believers" in the natural language processing domain, with several of them contending that the effectiveness of more traditional search algorithms has made natural language processing unnecessary despite the academic attention it has drawn in recent years (see Jurafsky and Martin 2008).

Others claimed that natural language processing is, at best, still a pipe dream, many decades away from being even remotely feasible. After all, they said, even the most prominent natural language processing effort, the supercomputer Watson developed by IBM to compete on the game show *Jeopardy*!, performed very inconsistently. Considering Watson's notion that "Delete key is where the heart is," among others (Sohmer 2011), the idea that it actually understood the array of text clues and responses presented is almost laughable.

This view, however, was not consistent across all workshop discussions. One participant noted, as an example, the TAKMI (Text Analysis and Knowledge Mining) tool, which researchers at IBM hope to use to assess the sentiment of customer service callers. The project team's goal of determining customers' feelings toward IBM and the problem at hand requires significantly deeper text understanding than mere keyword identification, so they have found natural language processing to be absolutely vital. Given the number of possible "answers" for a sentiment analysis, the task may also be more feasible than Watson's mission on *Jeopardy*! The degree to which an individual feels positively or negatively about a topic may span a limited number of possible numeric responses on a finite scale, while the range of possible textual *Jeopardy*! responses from which Watson had to choose is effectively infinite.

TAKMI is not an isolated case as sentiment analysis has been attempted elsewhere, such as on social media sites like Twitter (Jansen et al. 2009; Thelwall et al. 2011). While it remains an under-researched topic, there are certainly other scholars and organizations besides IBM who see its potential.

The workshop participants further noted other possible uses for natural language processing, such as identifying terrorist plots based on phone and e-mail interactions—despite ethical and legal concerns about governmental efforts to monitor citizens' personal interactions (Ball and Ackerman 2013; Donohue 2008; Greenwald 2013; Kerr 2003). This, too, extends well beyond simple keyword searches. Nonetheless, it remains quite possible that natural language processing may be unnecessary for all but a few unique text-mining tasks. Other than applications to search, natural language processing remains an area for development and commercial engagement.

Regardless of commercial implications, many programmers still see natural language processing as one of the most exciting research areas to be found. Case in point: Participants highlighted the textual analysis of expertise as another area in which researchers can make significant contributions, and the reputation, trust, and authority group specifically highlighted natural language processing as a technique that might be particularly helpful for discerning credible sources of information.

As they explained, it can be difficult and time-consuming to determine which sources of information in a given community are authoritative or reliable, as well as to identify group leaders. Search engines like Google fall short in this regard, they claimed. After all, such searches can pinpoint prominent groups, but once you find their discussions and contributions, you have to delve into them yourself in order to discern group roles. An automated tool that understands the content of the conversation, not just the language syntax, could be very helpful in isolating trustworthy authority figures within such groups. Such a tool might not be far away, a conclusion of the workshop being that with recent developments in artificial neural networks as well as better understanding of brain structure, we might shift from machine learning to machine *understanding*.

On a methodological level, predictive modeling and data analytics applied to machine understanding might generate a particularly promising set of approaches within the hazard research domain. For instance, in the economic realm, researchers have begun using textual analyses of financial documents for risk predictions, whether of market trends, fraud, or mere costs (e.g., Kolyshkina and van Rooyen 2006; Nakatoh et al. 2013). Similar approaches have proven fruitful for assessing the risk of medical events (Poulin et al. 2013), with some researchers turning to data sources like Newswire and social media sites as well as filings by the Securities and Exchanges Commission.

Data analytics are also important for predicting and tracking natural disasters (see, for instance, Jansen et al. 2009; Starbird et al. 2010), a particularly difficult subject matter from a methodological standpoint since data on such events tend to be sparse and unreliable.

D. Integrating Multiple Data Sources As a potential counterpoint to the preceding discussion, some workshop participants contended that trustworthiness research might advance even more quickly if scholars identified and combined information from multiple types of evidence to form a synthesized picture of a given authority system. Presently, most studies view communities through only a single perspective, with just one kind of data analyzed. This inevitably results in an incomplete evaluation of the organization being assessed, regardless of the analytic tool used. Combining several sources of data would help to mitigate this problem.

As an example, pure textual information could be linked with data on social relationships, with the network structure supplementing and enriching conclusions about the presence of personal expertise and the way in which clout develops within a community structure (see, as an example, Ehrlich et al. 2007). Diversifying the data used in analyses could ultimately be far more beneficial than developing more sophisticated analyses to handle incomplete data.

This, naturally, also represents an example of linking multiple methodologies to enhance the analytic process. In some cases, however, researchers might choose to gather a wide range of data sources while still maintaining their focus on a single methodology. For instance, text-mining scholarship has a great deal of potential to tap into large-scale repositories of textual, audio, and video interactions. Many scholars are particularly interested in exploring such questions as the ways in which individuals within collaborative communities establish their own trustworthiness, how they evaluate that of others over time, and differences in the nature of credibility and trust as they occur within different types of groups, such as the distinction between online and offline teams. By combining text, audio, and video data sources, such researchers can more easily distinguish between and connect verbal and nonverbal communication processes, along with the sociological and psychological effects that accompany them. A possible approach would be that of enhancing the C-Span video archive search utility with "practice capital" metadata. Such metadata captures the degree to which a contributor to a debate is central or peripheral not as a function of how much he or she said but of centrality in a network of connections determined by turns of speech (Matei 2014).

However, it is often very difficult for researchers to synthesize data from different sources given the wide variety of ways in which data are presently collected, recorded, reported, and stored. Likewise, in general, different types of data have been collected on different communities, which makes it challenging for meta-analysts to synthesize multiple studies into more comprehensive findings. The development of standard formats for data sets and metrics, along with more universal measures and evaluation procedures, could resolve this problem, even if fully realizing such a scenario proves to be a utopian ideal.

E. Community Design and Feedback Systems The conflict between processing speed and cost is a problem facing researchers who rely on transcription systems, but one that will be resolved over time. The challenge of fostering and maintaining activity within a knowledge market while still keeping its content credible for consumers, on the other hand, is a major question that scholars can address today.

Shneiderman (Chap. 2; see also Sopan et al. 2012), in this volume, focused on design principles for online communities to build a four-component research framework on how practitioners can develop a thriving community that would also offer credible content for consumers. Broadly speaking, he suggested that researchers would be able to make a significant impact by devoting greater attention to (1) strategies to maximize the proportion of trusted contributors within a community, such as by raising barriers to entry or pinpointing troublemakers through network analysis; (2) reviews of the content itself that are designed to encourage high quality and filter out unhelpful contributions; (3) efforts to provide contributors with credible, secure resources to support their work; and (4) the use of such processes as audits and external oversight to ensure that communities and their leaders are themselves responsible and trustworthy. Shneiderman's essay provides substantially greater detail on these four scholarly areas of emphasis.

Along the same lines, another primary concern among workshop contributors was the push to make online institutions and quasi-institutional social media more transparent, democratic, and open to user feedback. At present, most online tools are ineffective at making the organization available for evaluation and monitoring, largely because it has not been a design or research priority. However, making the organization's actions more open to feedback could help to resolve myriad problems such as concerns over biases and rule enforcement practices on Wikipedia. If the organization was able to elicit feedback from users, in a manner that would strip the feedback of any connection to the their reputation, then that feedback system might mitigate instances in which one's reputations carries undue weight or in which established users are inclined to "pick on the newbies," as one workshop participant said.

Of course, such problems are not limited to online communities. Some reciprocal evaluations of leaders and ordinary users might be untrustworthy, especially in politically charged or partisan situations. The evaluation itself might become untrustworthy, if all of the respondents hold a personal stake in the result (see also Teng et al. 2010).

With that in mind, the question becomes how one might change the organizational models to elicit better, more credible feedback about leaders. After all, it is very difficult to trust any responses that may have been biased by defensive self-interest. As one member of the workshop said, offering as an example an administratively controlled university:

> How would you change some dials in a university to make it more democratic; make the monitoring more complete; make the sanctioning more multidirectional, less concentrated? You need to be able to tell [that] people did their job. You need to be able to have criteria for how they're doing it. You need to be evaluated by people that don't have their skin in the game.

In short, organizations need to have proper reward and sanctioning systems for leaders based on a trustworthy set of feedback mechanisms. Regardless of the particular knowledge market in question, whether it is a social media platform, a university, or an entirely different organization, the cautionary tales and instructive lessons about feedback credibility may be applied across contexts.

Along similar logistical lines, the reputation, trust, and authority group considered the nature of rules and norms on the Internet, comparing the reputation systems that tend to govern online behavior with the body of laws held over offline society. A few group members contended that having a more rigid set of Internet laws in place, rather like the Consumer Privacy Bill of Rights that President Barack Obama proposed in 2012, would make many online activities function much more smoothly. If so, the question then turns to the design and implementation of such laws, particularly since, from a legal standpoint, it would be difficult to claim that any particular police force or government holds any authority over the Internet itself. Furthermore, such laws would clash in the USA with the First Amendment, as it might curtail the freedom of expression in the name of social control.

F. Summary All told, multiple workshop participants placed considerable emphasis on the theories, methods, and tools that would best support the analysis and modeling of trustworthiness information. This came alongside concerns over standardization of data sets and evaluation procedures, as well as the debate over whether natural language processing was really a worthwhile approach to solving some of these analytical problems. However, aside from advertising research, some participants were concerned about how one could evaluate the relative importance of research topics that might not have a direct economic impact. Such concerns pervaded numerous other potentially promising areas of study, including design principles for online communities, the use of feedback systems within knowledge markets, and even proposed rules to govern online behavior.

Future Methodological Approaches for Improving Study of Trust, Authorship, and Credibility on Social Media

Even as scholars gain access to more efficient resources for their research, and as the analytic techniques that have been developed for their use grow ever more sophisticated, there still remain several critical methodological problems to be resolved. The KredibleNet workshop participants highlighted three of the most troublesome problems that pervade a great deal of current research: failures to demonstrate causality, the use of inappropriate data that does not properly connect with underlying theories—especially due to the increasingly heavy use of and over-reliance on big data (see Lazer et al. 2009)—and struggles against computational limitations.

A. Testing Causality Scholars have often observed that individuals who are connected with one another in a social network tend to share numerous similarities, but it is not clear whether this is because similar individuals are likely to form connections, or if those individuals who are already connected tend to become more similar to each other over time.

On a broader scale, the key question is how researchers can show causation with large data, which is often purely observational. To expand upon the earlier example, we might like to test whether individuals connected in a social network share attributes because of homophilic or other types of social influence forces, but this is a challenge with most large data sets. As scholars who are well versed in methodology would be quick to note, any claims of causality are suspect unless they arise from a formal experiment that randomized participants into distinct groups and applied treatment and control conditions. "Big data," such as large repositories of information on user activity within social networks, tend to instead come from more passive field observations for which no treatment and control conditions were ever clearly applied (see Williford et al. 2012). Although some scholars have developed sophisticated statistical analyses for big data sets to demonstrate intriguing connections between concepts, individuals, and groups (e.g., Gelman and Hill 2006; Ho et al. 2012; Xie and Xing 2013), in the absence of an experimental design, even these techniques cannot demonstrate causality.

With that said, some scholars hope that big data may at least be treated as quasi-experimental, or perhaps even as natural experiments, with some environmental force affecting a random subset of organizational members as opposed to an experimental condition randomly applied to subjects. In such a natural experiment, researchers could claim that a causal relationship existed between the naturally occurring treatment and any observed differences between the various groups—provided that the treatment was indeed random, that the environmental force was not related to the observed effects through some other mechanism, and certainly that the environmental force took effect prior to the observed differences between groups.

One workshop participant suggested that if the problem is simply that big data come from passive observations, we could rectify it simply by being more active in applying treatments. For instance, Bond et al. (2012) conducted a field experiment using Facebook in which political mobilization messages were sent to 61 million users. The researchers' subsequent observations of how different treatment conditions—whether or not someone received a message—prompted users, their friends, and their friends' friends to change their behavior constituted one of the largest experiments ever conducted. Granted, it can be difficult to get permission to conduct interventions within a community like Facebook, and doing so certainly prompts questions about the rights of unknowing study participants, but the opportunity to combine the methodological advantages of an experiment with the practical benefits of big data serves as an invaluable wellspring for research.

B. Data Appropriateness These concerns are all part of the larger problem of connecting theory with data. In many cases, the data that are available to scholars only indirectly or tangentially address the research questions to be answered. Even the best-designed surveys, deliberately applied to treatment groups, may not exactly address the construct of interest or otherwise have less than perfect validity. Such problems are exacerbated with secondary data sources, as the scholars who receive a given data set after the fact have little control over what kinds of information were initially collected, how the information was gathered, the ways in which it was organized, and so forth. When the researchers cannot deliberately target a particular research question in the data collection phase, it becomes much more difficult to convincingly show that any findings do, indeed, address the problem at heart.

Regardless of whether primary or secondary data sources are being used, it is also important to use data from multiple sources as opposed to a single frame of reference. Just as the text-mining and natural language-processing workshop group was concerned about using multiple sources of evidence to develop a more complete picture of any particular authority system, any scholar studying online communities would be well served by including offline activities in his or her analyses as well. For instance, it would be impossible to gain a complete sense of how the top Wikipedia editors operate and interact with one another without acknowledging their periodic gatherings in the "real world." Even a big data archive is only a single source of data, after all, with blind spots of its own. Incorporating a range of data sources would help to mitigate the tendency to overlook important attributes and behaviors that happened to escape a single data source with its necessarily limited scope.

Of course, one of the biggest concerns for many social and natural scientists today is the growing perception that you can do anything and solve any problem as long as you have big data—in other words, the idea that big data are a modern-day panacea for research that will somehow magically fix any obstacle facing scholars. The notion is understandable as one of the most common limitations cited in academic research articles is a dearth of data, yet this does little to resolve the wide variety of other challenges for researchers.

In fact, big data themselves can sometimes be limited. For instance, one workshop participant noted that many studies using big data are only able to focus on major "jumps" within a particular system, overlooking other effects that may be smaller but which are still significant. The natural solution, then, is to put more probes within big data sources in order to capture the finer nuances. This, too, is an imperfect approach. As any scientist knows, raw data are rife with noise, whether it comes from the natural sciences or social phenomena. As a researcher tries to capture smaller effects, the risk of confusing meaningful content with noise dramatically increases. Big data are no different (Williford et al. 2012).

Filtering data in order to separate genuine effects from noise therefore becomes a paramount concern, along with the balance between maintaining high data quality and striving for data completeness. An appropriate cyberinfrastructure could be especially useful if it incorporated data quality control mechanisms—especially the automated filtering of invalid or incorrectly recorded data—as early in the research process as possible.

Similarly, the workshop participants were particularly concerned with the ease with which information can be verified. In particular, they discussed the importance of having multiple, reliable, credible secondary sources for the knowledge production process. For instance, scholars of physics and related fields certainly rely on traditional academic journals for their research, but they also recognize arXiv (Cornell University Library 2013), a long-standing archive of work within the field, as a useful resource. This is certainly important for researchers within these particular academic domains, but just as importantly, the presence of multiple resources also allows scholars to better share their work with the general public. After all, physics pages on Wikipedia tend to be highly reliable, as it is especially easy for editors to check their contributions against multiple sources. Elsewhere on Wikipedia, article accuracy tends to suffer, as other fields lack the same array of resources, which simply makes it more difficult to check information. All told, the presence of large archives facilitates much more effective and efficient research, whether in terms of comprehensiveness or information verifiability.

With that said, Howison et al. (2011) highlighted substantial issues with the validity of big data after researchers have distilled them for the purposes of interpretation, regardless of how well intentioned their choices may have been. Similarly, boyd and Crawford (2012) voiced concerns about the privacy of unwitting research subjects and the need for serious discussions about the ways in which big data are being used and how to keep individuals protected. On both an academic and a political level, discussions about how to protect individuals should establish and review technical approaches and scholarly policies for anonymizing data while also maintaining their usefulness for researchers, among other topic areas.

The presence of massive data sets necessarily creates many other challenges for analysis. This is especially true in domains like network analysis, where statistical inference and modeling can sometimes prove to be imposing or even impossible tasks for networks much larger than a few thousand nodes in size. Such problems tend to be exacerbated within studies of network change, a cutting-edge topic that remains particularly difficult to manage on a large scale.

Some of the aforementioned challenges with unstandardized data are especially problematic for network researchers. For instance, how can scholars pair dynamic networks with static ones when studying network change? The answer is far from clear, as the conceptualizations of dynamic and static networks are fundamentally different. Such problems are compounded if changes within the various networks do not match one another in terms of time, creating a staggering effect for which scholars may be hard-pressed to account. Imperfect or inconsistent data collection and recording practices could compel scholars to manipulate the data in order to develop a clearer picture of the network evolution, but there is no clear protocol on how to do so, which largely leaves researchers to improvise their own solutions—yet another source of inconsistency that makes it tougher for future scholars to interpret and synthesize findings across multiple studies.

Furthermore, data-mining algorithms tend to have minimal connection with the key questions that social scientists would like to address. Since so few researchers have the ability to cross between the scientific, methodological, and programming domains in order to fully link their research questions with the data that they hope will provide answers, many scholars have to rely on secondary data sets for their analyses. This naturally causes problems with data interpretation, as the data set was not originally developed to resolve the scholars' research questions. In particular, several workshop participants noted a division between data-mining algorithms and network models, and they expressed the need for a way to link the two.

C. Computational Limitations In some cases, when the most significant challenges facing a given research project are mere computing speed or memory issues, skillful scientists and programmers are able to devise alternative approaches to perform a given analysis. For instance, Brandes (2001) developed a new algorithm to calculate betweenness centrality, one of the most prominent network attributes, in a manner that proved much less costly for large, sparse networks. It is likely that many similar problems with computational resources could be resolved in a similar manner, although fewer researchers may be compelled to pursue such solutions for lesser-known measures. Nonetheless, for significant research topics like the prediction of natural disasters, which revolves around sparse big data sets, innovative research approaches that make it feasible to analyze such data could prove to be a tremendous boon.

Mercifully, such computational issues are not shared across all research domains. Some of the workshop participants argued that computing power was not a significant challenge for examining a large corpus of text. With that said, there was some debate about whether the size of a given data set could at least play a noticeable role in how much computing time and processing power would be necessary for analysis, particularly based on the size of the training data set, but the consensus was that computational issues were, at the very least, not an obstacle that would keep analyses from being completed within a reasonable timetable.

D. Summary In sum, the KredibleNet workshop participants cited three major methodological problems that remain unresolved: the question of testing causality with big data, which related to the broader question of finding or generating data that are appropriate to resolve the research questions at hand, along with computational limitations on large-scale analyses. These shortcomings have been addressed in some isolated cases, such as in Bond et al.'s (2012) Facebook field experiment and Brandes' (2001) betweenness centrality algorithm. Nonetheless, these issues remain prevalent across a great deal of modern research, particularly as big data studies continue to rise in prominence as an increasingly attractive scholarly approach.

Organizational Shortcomings to Be Addressed

The workshop participants also identified three organizational issues that inhibit the emergence of a new paradigm of research on social media roles, authority, and credibility in knowledge markets: paucity of funding, the relative absence of researchers who hold skill sets that cross disciplinary boundaries, and the lack of data-sharing activities among scholars.

A. Funding The biggest obstacle for knowledge market research is very likely funding; some would argue that the same is true across all research domains. Yet, in our particular case, the issue is made more severe by the fact that a research agenda related to roles, authority, leadership, and credibility, although socially significant, does not have the kind of public recognition as the topics that are stereotypically associated with social media research, such as crowdsourcing, user motivation, or contribution equalization. Several workshop participants talked about the indirect connection between many promising research topics and tangible outputs. The obvious consequence is that when it is not clear how a company might profit from a particular research line, it is unlikely that the scholarship will draw any industry support. Major corporations are generally much more concerned with protecting their short-term bottom lines and satisfying stockholders than with engaging in exploratory studies, regardless of how fruitful the results might be for society at large or even to them in terms of long-term strategy.

This challenge, however, may be resolved by major funding agencies like the National Science Foundation and the National Institute of Health. Grants from these organizations may be directed toward projects which would have minimal economic impact for a particular organization but which would instead serve the broader needs of society at large, even if only to lay the theoretical and practical groundwork for further research by the primary investigators and others. Particularly important is to broaden the spectrum of research to include issues that are not immediately intuitive. Theories related to complementary association, the natural accretion of roles, social structures, and cultural norms within social media, and the evolution of credibility and trust mechanisms that rely on leadership models need to be particularly targeted. Too much social media research is conducted from simplistic, common sense assumptions, such as the abovementioned homophily axiom. A major cause of this oversimplification also comes from the fact that much social media research is conducted by scientists with a strong background in methodologies and computer science but no formal social scientific education. Funding agencies need to encourage cross-disciplinary and cross-university collaborative teams like the Social Media Research Foundation, developers of NodeXL, an open-source network analysis tool that has allowed countless researchers with minimal programming expertise to initiate data-mining studies and network analyses throughout a variety of disciplines.

B. Multidisciplinarity On a practical level, one of the most unique challenges for large data research is the fact that researchers tend to adopt distinct roles. Some scholars are data managers or computer scientists who focus on managing the analytic infrastructures and large databases. Others are adept at conducting sophisticated statistical analyses on the data themselves. Then there are the theoreticians who know what to do with the results. By and large, these are three wholly distinct groups, and very few people can effectively execute all three roles—after considerable discussion, the only example that the workshop participants could identify was *New York Times* blogger Nate Silver.

It is no surprise that so many scholars focus their efforts on one particular area rather than dividing their attention across a more diverse skill set, especially considering research indicating a (slight) improvement in work quality as a result of a narrow focus (Adamic et al. 2010). Consequently, few scholars are able to manage the entire data-mining process, so those academics who work on data-mining projects tend to organize themselves into interdisciplinary teams. Most importantly, the members of those teams often have minimal understanding of what their colleagues on the project are actually doing as their skill sets do not overlap.

This has exacerbated other problems like the aforementioned rift between data and theory, as the experts who understand the data structures struggle to communicate with those who conduct the analyses, who likewise find it difficult to share their knowledge with and grasp the work of those scholars who actually understand and hope to develop the underlying theories. As long as the data scientists, statisticians, and theoreticians fail to comprehend one another's work, it will remain very difficult for them to connect data with theory through a well-designed statistical analysis.

With that in mind, for social media data-mining research to move forward, education needs to improve, as we cannot simply maintain classes of scholars who are just data managers, just computer scientists, just statisticians, or just theoreticians. Multidisciplinarity is becoming increasingly important, so while it is useful to have experts within each unique domain, it is just as necessary to have academics who can cross between these otherwise distinct fields. Scholars who can do so are inherently better equipped to develop analyses that actually resolve the core theoretical questions in which they are interested.

This, however, is itself a major problem among researchers because it begs the question of how to foster multidisciplinarity. The workshop discussions highlighted that education will be important for this purpose, but they could not agree on the exact form that a multidisciplinary education would take. Some suggested changing current university curricula to open new paths for scholars in one domain or another to experience other disciplines in order to gain a degree of multidisciplinarity. Others argued for the development of an entirely new major area of study that would draw from multiple other fields. Still others claimed that such radical changes were unnecessary, but that current institutions simply need to understand the need for scholars to have some kind of background in statistics, a limited understanding of big data and programming, and at least sufficient database skills to convert data structures into a suitable form for analysis. The group was unable to reach any kind of conclusion about the best approach, so this may itself constitute a set of research questions about how the university system can best be organized to help scholars who want to cross over multiple areas.

C. Data Availability As previously noted, data appropriateness is a serious concern affecting some research domains. However, many analyses of social phenomena are constrained by the simpler question of data availability. Scholars who study credibility, for example, would like to have large datasets to model behavior and to connect those behaviors with others' evaluations of individual trustworthiness. By and large, such datasets do not exist or have limited availability for interested researchers—which makes sense, considering how difficult it can be to gather responses from a large-scale group on the trustworthiness of everyone within the system.

Cross-cultural research serves as a particularly instructive example as even those studies that claim to explore a variety of cultures are often very limited in terms of the spectrum that they actually compare (Zhang and Lowry 2008). There remains a dearth of cross-cultural and international studies that compare nearby cultures and countries in tandem with those spanning multiple continents. In other words, many studies have focused on the nuanced distinctions between geographically proximal groups, but fewer have contrasted countries and cultures with a broader geographic dispersion due to the inherent challenges and obstacles that inhibit data collection within certain communities throughout the world. Both nuanced and radical cultural distinctions are essential in order to obtain a complete picture of cross-cultural differences, but the absence of suitable data makes it difficult for researchers to move the field in that direction.

Nonetheless, while such research may be difficult, it is not necessarily impossible, depending on the context. For instance, a few workshop participants highlighted recent studies comparing distinct Wikipedias of different languages as well as differences between categories and types of pages within the same Wikipedia. To put it another way, these studies contrasted the work of highly distinct cultural groups (using different Wikipedias) as well as that of subcultures of editors and academicians from similar cultural backgrounds but different academic fields (Gallagher and Savage 2013; Hara et al. 2010; Pfeil et al. 2006), capturing both dramatic and fine-grained cultural differences.

The problem of availability also extends to proprietary data sets. Many companies have aggregated tremendous amounts of information about various populations, particularly in the online domain. Google, for example, maintains extensive records of user searches in order to better tailor search results to individual needs (McDermott 2013). Facebook, likewise, is notorious for maintaining records on virtually everything that has ever been written, uploaded, or done on their site (Rosenbush 2013). Despite the tremendous research potential of such sources, however, public access to such data is largely restricted, whether to one's own Google search history or to the activities of an individual's finite number of Facebook "friends," making comprehensive data-mining efforts exceedingly difficult at best.

This is not to say that all data repositories should necessarily be completely open for public use. After all, even discounting the right for corporations to strategically collect and use information, haphazardly releasing information would create myriad privacy problems on ethical and legal levels alike (boyd and Crawford 2012). At the very least, when data is released to academicians, those scholars must keep standard IRB processes and scientific values in mind. On a broader scale, researchers and organizations alike must adhere to all applicable policies for the use and protection of individuals' private information—whether or not such policies are standardized across research domains—including the organization's own terms of service and internal practices.

In addition to such ethical standards, technical restrictions may certainly come into play, particularly as they relate to researchers' ability to access, manipulate, and protect the data that they hope to use. Workshop participants noted that this issue is compounded by the present ineffectiveness of software tools, arguing that we need better-designed tools for information retrieval and text mining in order to conduct appropriate and effective social scientific research. This is especially important for scholars who lack the programming expertise to develop software that serves their own needs.

Aside from the simple approach of making existing data sets available for public or research use, crowdsourcing may be a useful approach for overcoming the problem of data availability. By soliciting contributions from a large group or the general public, researchers can rapidly generate very large amounts of data, from quantitative survey responses to anecdotal case studies. Certainly, crowdsourcing is useful in many other research processes as well, from developing folksonomies to proposing hypotheses to be analyzed, but its utility for generating data where none exist is of special importance for many scholars across domains.

Finally, data management and preservation are closely related to this question of data availability. The library scientists who participated in the KredibleNet workshop round-table sessions were especially interested in the ways that data sets are derived from raw information, how they are described, and how they can be discovered and reused by other scholars. Their core concern was ensuring that data which one researcher or research team collects are not simply lost to the rest of the academic community. It is therefore important to preserve data and sustain the systems in which they are based, develop useful descriptions and meta-data for those larger data sets, and help other scholars to cite secondary data sets so that their original creators are credited for their work and motivated to share it with the world at large.

D. Summary While funding remains the most critical organizational issue facing researchers today, the lack of multidisciplinarity among scholars and the absence of appropriate data remain significant problems in their own right. Importantly, the latter two challenges may at least be partially resolved through collaboration. Scholars without multidisciplinary skills can cross disciplinary boundaries by connecting with peers with different backgrounds and talents and mutually developing their research.

Similarly, many data sets are collected for one team of scholars to use in a small number of publications and are then discarded, yet mutual data-sharing practices would permit researchers across domains to have much easier access to appropriate data and to develop far more sophisticated analyses with more interesting findings. Facilitating collaborative interactions and data-sharing practices must therefore hold great importance for any collaborative research interface, including that which the authors intend to develop.

Requirements for a Collaborative Research Interface

The various challenges highlighted throughout this chapter are connected by the simple fact that scholars and practitioners need better collaborative tools and systems to facilitate joint research efforts. Many common problems could be resolved by the simple presence of software that would better support research projects, yet such a system has proven elusive, as countless programs have been developed for that very purpose and have been abandoned or neglected from the very beginning. Anecdotally, most scholars appear to have avoided the tools specifically designed for their research needs in favor of more generic, yet adaptable, file-sharing systems like Google Drive and Microsoft SharePoint along with their preferred set of data analysis programs.

Over the course of the first KredibleNet workshop round-table session, the authors identified two fundamental problems with neglected collaborative software packages that lessened their usefulness and inhibited their adoption. Those problems were (1) the limitation to a small portion of the research process, such as collaborative interactions or data analysis, and (2) the stagnation of tools that fails to account for researchers' rapidly evolving needs. In developing a new collaborative system for researchers, we aim to correct these problems via the convergence of technologies to support every phase of the research project within a single system, and by using an open-source approach to feature development in order to foster easy extensibility and to keep the infrastructure from becoming defunct.

A. Technology Convergence The problem goes well beyond merely providing a space for interactions between scholars. Rather, an intelligently designed interface also needs to facilitate the data analyses themselves, and it must facilitate data sharing, results reporting, and any other key tasks throughout a given research project. The methodological need is particularly salient now, as many scholars who explore sophisticated methodologies continue to find standard software packages available for purchase or download to be inadequate, and given that researchers with interdisciplinary capabilities remain highly uncommon, individuals with more focused skill sets tend to lack the ability to develop software that will satisfy their methodological needs themselves.

Consequently, a truly effective collaborative interface would allow researchers to easily share their ideas with peers; collect, distribute, manage, and report data; conduct analyses; and deliver results. We may say more simply that the ideal is a one-stop research domain. If the full scope of research tasks could indeed be satisfied within a single interface, it would permit much smoother collaborative efforts than a collection of many different software packages—rather like the current array of sharing, data management, analysis, and collaborative suites that have inundated the marketplace—as it would mitigate the problem of fragmentation among researchers into distinct camps depending on their favored tools.

To this end, we need to go beyond what is currently available. Virtually all other software packages designed for collaborative research have lacked major components of the process, whether the capability to interact with peers, manage data, conduct analyses, or share results. Even DataONE (2013) and FLOSSmole (2013), which are among the most sophisticated and successful collaborative platforms, miss at least one or two of these major steps (although they may still offer useful examples to be emulated, especially in terms of their respective data-sharing functionalities). Just as importantly, even those tools which target one aspect of the process tend to be limited themselves, with their functionality growing outdated over time.

In short, the platform we envision would synthesize each phase of the research process and the features necessary to proceed through each phase as a collaborative team. Just as there are many different competing programs to perform any number of other tasks, from filing taxes to painting digital artwork, there are many different collaborative and research tools for a reason: Each has a set of strengths and weaknesses, and since each program fulfills some functions better than others, users gravitate toward the one that best satisfies their needs. This can potentially foster fragmentation, yet we aim to stop the cycle of this fragmentation. The ultimate goal of the KredibleNet project is to design an interface that incorporates all of the features that researchers across disciplines need from day to day.

B. Open-Source Framework It is easy to anticipate certain tools that would be critical for such an interface, such as communication functions from basic chat features to sophisticated video conferencing frameworks, data management functions like filtering noise and serving preprocessed data, and the spectrum of basic methodological tools. However, more specific needs are more difficult to predict, especially in terms of methodology where scholars develop new approaches on a daily basis.

As such, this interface cannot rely on a static set of "one size fits all" tools, but should be easily and quickly extensible to accommodate researchers' needs as they develop; otherwise, it would quickly become defunct as researchers' needs exceeded the program's capabilities. This is particularly true for a collaborative tool whose value, according to Metcalfe's law, is largely dependent on the number of other people using the system (Hendler and Golbeck 2008). The loss of a few dissatisfied users who reach the program's limits would likewise reduce its usefulness for the rest of the user base, so such a loss must be deemed unacceptable.

With that said, no finite set of staff members could realistically keep up with the ever-changing needs of users. This is as true for a collaborative research interface as it is for any other evolving system (see, for example, Britt 2011). Given that more specific methodological needs are difficult to predict, this interface cannot rely on a static set of "one size fits all" tools but should be easily and quickly extensible to accommodate researchers' needs as they develop.

As for how this extensibility can be achieved, the answer is surprisingly simple. When the users themselves are able to develop tools—and do so easily, without a substantial learning curve—the system has the opportunity to keep pace with its users' demands. Roughly speaking, this is the philosophy that has worked for Wikipedia, allowing users to contribute knowledge to the system so that the user community at large will have that resource available to them. This approach has made Wikipedia the go-to source of basic information for curious individuals around the world, transforming a modest initial effort into a near-global monopoly (Garber 2011; see also Britt 2011).

The idea of allowing users to be involved in such an ongoing cocreation process also forms the foundation of R, an open-source software package for statistical analyses (R Foundation 2013). The more technical nature of R development makes it an especially illustrative example against which to compare the proposed infrastructure. Unlike R, however, the authors intend to develop a collaborative suite that transcends the analysis stage of the research process and which is more easily accessible for first-time users who want to quickly add a new tool to the system without the extra hassle of learning a new programming language. This demands both that the system be more comprehensive than R, covering the full scope of scholarly needs from initial idea generation to the dissemination of final results and reports, and that it not be restricted to any one programming language with which a contributor may or may not be familiar. Naturally, the system should also be as accessible as possible for users with limited or no programming background, similar to the aforementioned front-end features available through NodeXL. Only by making it feasible for as many users as possible to engage with the open-source development process can the system reach its full potential, serving the needs of the broadest possible audience of researchers.

C. Summary As highlighted during the KredibleNet workshop, a truly effective collaborative suite must combine each stage of the research process, from initial idea discussions to data collection and management, and from analysis to the reporting of final data and results, into one system. However, this breadth must also be balanced with sufficient depth, which means that users must have a sufficient range of tools at each stage of the process to satisfy their needs, such as a comprehensive array of methodologies supported by the software. The only realistic way to achieve satisfactory breadth and depth is to involve the users in the development process, much as R users have contributed to the open-source software package for statistical analysis.

On the largest scale, the efforts to devise these resources constitute a critical practical and scholarly challenge spanning countless academic fields and scientific domains, and to which the entire base of potential users must have the opportunity to contribute.

Conclusion

Clearly, this is an exciting period for researchers across scientific domains, with numerous opportunities for groundbreaking scholarship on authorship, roles, and social media credibility made possible by recent advances in collaborative and data analytic technologies. Yet, this potential still has not been adequately harnessed, as numerous research problems that demand collaborative attention remain unresolved. As the opportunities for revolutionary research grow richer, these shortcomings will inevitably become a source of great consternation if left unaddressed.

At the core of these issues is the lack of a suitable framework for collaboration. To this point, no collaborative research environment has been remotely effective at allowing scholars to seamlessly transition between engaging other researchers, uploading and analyzing data, and reporting findings for the world to observe. Theories are still to be developed to continue, rather than ignore, social and mass media theory. Even the strategies, theories, or platforms which purport to offer useful collaborative and analytic tools only tend to provide a few isolated objects, and minimal consideration is given to on-demand expandability in order to serve researchers needs as they arise.

In our increasingly global society with its growing emphasis on collaborative work, this is a critical gap in the research process. Many of the problems that form the core of current research, as well as those which hold the most promise for the future, appear to be well within our grasp given the proper perspectives and tools, but such a perspective has historically eluded the academic and public realm. Ultimately, this chapter highlights the need for such a carefully designed approach and framework that would more actively listen to the needs of its users in order to offer the resources and tools that would best allow them to address the most important research problems of today and of tomorrow in the fields of authorship, roles, and credibility.

Based upon present research needs and those projected for the near future, the authors propose the development of a new approach that problematizes the lack of theoretical coherence of current work, that advocates for strong multidisciplinary collaboration and for funding of nonintuitive research agendas and of collaborative platforms that would support scholars at all stages of the scholarly process. This demands the convergence of many perspectives, practices, or technologies into a single framework, mitigating the fragmentation of academics and practitioners into distinct camps based upon their particular preferences. In turn, the only practical way to achieve such comprehensive convergence is to develop the collaboration frameworks in such a manner that it is easily extensible, allowing researchers from a diverse range of backgrounds to quickly add additional tools to the existing repository of tools or to the future analytic platform so that it continues to serve their ever-evolving needs. This must be done without demanding that contributors develop an entirely new skill set, as the divisions between researchers and the relative lack of researchers who possess cross-disciplinary skills (such as social science theory paired with a programming background) is a key reason why collaboration, and a system which adequately supports it, is so vital in the first place.

The authors hope that this proposed approach will better support scholars and practitioners in their future research work, permitting stronger and more effective collaborative endeavors than ever before. Researchers around the world have tremendous potential to conduct groundbreaking research on roles, authority, credibility, and social structure on online knowledge markets and to share it with their peers and the public at large. Our goal is simply to offer a framework that will allow those researchers to harness that untapped potential and to finally conduct the revolutionary work of which they are capable.

References

37signals, LLC. (2013). Project management software, online collaboration: Basecamp. https://basecamp.com. Accessed 8 Nov 2013.

Adamic, L. A., Wei, X., Yang, J., Gerrish, S., Nam, K. K., & Clarkson, G. S. (2010). Individual focus and knowledge contribution. *First Monday, 15*(3). http://firstmonday.org/ojs/index.php/fm/article/view/2841/2475. Accessed 1 Jan 2014.

Alfresco Software, Inc. (2013). Document management | Open source collaboration software | Alfresco. http://www.alfresco.com. Accessed 8 Nov 2013.

Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. (2012). The role of social networks in information diffusion. In *Proceedings of the 21st international conference on the World Wide Web (WWW '12)* (pp. 519–528). New York: ACM.

Ball, J., & Ackerman, S. (2013). NSA loophole allows warrantless search for US citizens' emails and phone calls. *The Guardian*. http://www.theguardian.com/world/2013/aug/09/nsa-loophole-warrantless-searches-email-calls. Accessed 1 Jan 2014.

Berman, A. E., Barnett, W. K., & Mooney, S. D. (2012). Collaborative software for traditional and translational research. *Human Genomics, 6,* 21.

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature, 489,* 295–298.

boyd, d., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society, 15,* 662–679.

Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology, 25,* 163–177.

Britt, B. C. (2011). *System-level motivating factors for collaboration on Wikipedia: A longitudinal network analysis*. (Master's thesis). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. AAI1501175).

Cacioppo, J. T., Fowler, J. H., & Christakis, N. A. (2009). Alone in the crowd: The structure and spread of loneliness in a large social network. *Journal of Personality and Social Psychology, 97,* 977–991.

Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine, 357*(4), 370–379.

Cornell University Library. (2013). arXiv.org e-Print archive. http://arxiv.org. Accessed 8 Nov 2013.

DataONE. (2013). DataONE. http://www.dataone.org. Accessed 8 Nov 2013.

DataSift. (2014). http://datasift.com/. Accessed 4 Feb 2014.

Donohue, L. K. (2008). *The cost of counterterrorism: Power, politics, and liberty*. New York: Cambridge University Press.

Ehrlich, K., Lin, C.-Y., & Griffiths-Fisher, V. (2007). Searching for experts in the enterprise: Combining text and social network analysis. In *Proceedings of the 2007 international ACM conference on supporting group work* (pp. 117–126). New York: ACM.

FLOSSmole. (2013). FLOSSmole. http://flossmole.org. Accessed 8 Nov 2013.
Gallagher, S. E., & Savage, T. (2013). Cross-cultural analysis in online community research: A literature review. *Computers in Human Behavior, 29,* 1028–1038.
Gao, H., Barbier, G., & Goolsby, R. (2011). Harnessing the crowdsourcing power of social media for disaster relief. *IEEE Intelligent Systems, 26*(3), 10–14.
Garber, M. (2011). The contribution conundrum: Why did Wikipedia succeed while other encyclopedias failed? http://www.niemanlab.org/2011/10/the-contribution-conundrum-why-did-wikipedia-succeed-while-other-encyclopedias-failed. Accessed 1 Jan 2014.
Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models* (1st ed.). New York: Cambridge University Press.
Gilbert, E., & Karahalios, K. (2009). Predicting tie strength with social media. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '09)* (p. 211). New York City: ACM.
Google, Inc. (2013). Google Drive. http://drive.google.com. Accessed 8 Nov 2013.
Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology, 78,* 1360–1380.
Greenwald, G. (2013). NSA collecting phone records of millions of Verizon customers daily. *The Guardian*. http://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order. Accessed 1 Jan 2014.
Hara, N., Shachaf, P., & Hew, K. F. (2010). Cross-cultural analysis of the Wikipedia community. *Journal of the American Society for Information Science and Technology, 61,* 2097–2108.
Hendler, J., & Golbeck, J. (2008). Metcalfe's law, Web 2.0, and the semantic Web. *Journal of Web Semantics, 6*(1), 14–20.
Ho, Q., Yan, R., Raina, R., & Xing, E. P. (2012). Understanding the interaction between interests, conversations and friendships in Facebook. http://arxiv.org/pdf/1211.0028v1.pdf. Accessed 1 Jan 2014.
Howison, J., Wiggins, A., & Crowston, K. (2011). Validity issues in the use of social network analysis with digital trace data. *Journal of the Association for Information Systems, 12*(12), 2.
Huang, C. M., Chan, E., & Hyder, A. A. (2010). Web 2.0 and Internet social networking: A new tool for disaster management? Lessons from Taiwan. *BMC Medical Informatics and Decision Making, 10,* 57.
Jansen, B. J., Zhang, M. M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology, 60,* 2169–2188.
Jurafsky, D., & Martin, J. H. (2008). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (2nd ed.). Englewood Cliffs: Prentice-Hall.
Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy, 26*(1), 1–18.
Kerr, O. S. (2003). Internet surveillance law after the USA Patriot Act: The big brother that isn't. *Northwestern University Law Review, 97*(2), 607–673.
Klimeck, G., McLennan, M., Brophy, S. P., Adams, G. B., & Lundstrom, M. S. (2008). nanohub.org: Advancing education and research in nanotechnology. *Computing in Science Engineering, 10*(5), 17–23.
Kolyshkina, I., & van Rooyen, M. (2006). Text mining for insurance claim cost prediction. *Lecture Notes in Computer Science, 3755,* 192–202.
Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A. L., Brewer, D., et al. (2009). Life in the network: The coming age of computational social science. *Science, 323,* 721–723.
Matei, S. A. (2014). A social network analysis "practice capital" approach to enhance the C-Span Archive with meta-communication data to support public affairs debates and data journalism. In R. X. Browning (Ed.), *The C-SPAN archives: An interdisciplinary resource for discovery, learning, and engagement*. West Lafayette: Purdue University Press.
Matei, S. A., & Dobrescu, C. (2011). Wikipedia's "Neutral Point of View": Settling conflict through ambiguity. *The Information Society, 27,* 1–12.

McDermott, J. (2013). Google takes its tracking into the real world. http://digiday.com/platforms/google-tracking/. Accessed 1 Jan 2014.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology, 27,* 415–444.

Mendeley. (2014). http://www.mendeley.com/. Accessed 4 Feb 2014.

Merton, R. K. (1968). The Matthew effect in science. *Science, 159,* 56–63.

Microsoft Corporation. (2013a). Microsoft SharePoint—collaboration software—Office.com. http://office.microsoft.com/en-us/sharepoint. Accessed 8 Nov 2013.

Microsoft Corporation. (2013b). Technet forums. http://social.technet.microsoft.com/Forums/en-US/home. Accessed 8 Nov 2013.

Nakatoh, T., Amano, H., & Hirokawa, S. (Sept 2013). Prediction of growth rate of operating income using securities reports. Paper presented at the IIAI International Conference on Advanced Applied Informatics 2013 (IIAI AAI 2013), Matsue, Japan.

Newman, M. E. J. (2002). Spread of epidemic disease on networks. *Physical Review E, 66*(1), 016128-1–016128-11.

Obama, B. (2012). Consumer data privacy in a networked world: A framework for protecting privacy and promoting innovation in the global digital economy. http://www.whitehouse.gov/sites/default/files/privacy-final.pdf. Accessed 1 Jan 2014.

Onnela, J.-P., & Reed-Tsochas, F. (2010). Spontaneous emergence of social influence in online systems. *Proceedings of the National Academy of Sciences, 107,* 18375–18380.

Pfeil, U., Zaphiris, P., & Ang, C. S. (2006). Cultural differences in collaborative authoring of Wikipedia. *Journal of Computer-Mediated Communication, 12*(1), 88–113.

Poulin, C., Shiner, B., Thompson, P., Vepstas, Y., Young-Xu, Y., Goertzel, B. V., et al. (2013). Predicting the risk of suicide by analyzing the text of clinical notes. http://www.durkheimproject.org/results. Accessed 8 Nov 2013.

R Foundation. (2013). The R project for statistical computing. http://www.r-project.org. Accessed 8 Nov 2013.

Rader, E., & Wash, R. (2008). Influences on tag choices in del.icio.us. In *Proceedings of the 2008 ACM conference on computer supported cooperative work (CSCW '08)* (pp. 239–248). New York: ACM.

Raymond, E. S. (1999). *The cathedral and the bazaar*. Sebastopol: O'Reilly Media.

Rosenbush, S. (2013). Facebook tests software to track your cursor on screen. *Wall Street Journal*. http://blogs.wsj.com/cio/2013/10/30/facebook-considers-vast-increase-in-data-collection. Accessed 1 Jan 2014.

Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science, 311,* 854–856.

Schleyer, T., Butler, B. S., Song, M., & Spallek, H. (2012). Conceptualizing and advancing research networking systems. *ACM Transactions of Computer-Human Interaction, 19*(1), 2.

Shneiderman, B. (2014). Building trusted social media communities: A research roadmap for promoting credible content. In S. A. Matei & E. Bertino (Eds.), *Roles, trust, and reputation in social media knowledge markets: Theory and methods*. New York: Springer.

Sobkowicz, P., Kaschesky, M., & Bouchard, G. (2012). Opinion mining in social media: Modeling, simulating, and forecasting political opinions in the web. *Government Information Quarterly, 29*(4), 470–479.

Sohmer, S. (2011). Does Watson think a poor workman blames Yogi Berra? http://hypervocal.com/entertainment/2011/does-watson-think-a-poor-workman-blames-yogi-berra. Accessed 1 Jan 2014.

Sopan, A., Rey, P. J., Ahn, J., Plaisant, C., & Shneiderman, B. (2012). The dynamics of web-based community safety groups: Lessons learned from the nation of neighbors. http://hcil2.cs.umd.edu/trs/2012-02/2012-02.pdf. Accessed 1 Jan 2014.

Starbird, K., Palen, L., Hughes, A. L., & Vieweg, S. (2010). Chatter on the red. In *Proceedings of the 2010 ACM conference on computer supported collaborative work* (pp. 241–250). New York: ACM.

Teng, C.-Y., Lauterbach, D., & Adamic, L. A. (June 2010). *I rate you. You rate me. Should we do so publicly?* Paper presented at the 3rd workshop on online social networks (WOSN '10), Boston, MA.

Thelwall, M., Buckley, K., & Paltoglou, G. (2011). Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology, 62,* 406–418.

Tufekci, Z., & Wilson, C. (2012). Social media and the decision to participate in political protest: Observations from Tahrir Square. *Journal of Communication, 62*(2), 363–379.

Valente, T. W. (1995). *Network models of the diffusion of innovations (Quantitative methods in communication subseries).* Cresskill: Hampton.

Weinberger, D. (2011). *Too big to know: Rethinking knowledge now that the facts aren't the facts, experts are everywhere, and the smartest person in the room is the room*. New York: Basic Books.

Welser, H. T., Gleave, E., Fisher, D., & Smith, M. (2007). Visualizing the signatures of social roles in online discussion groups. *Journal of Social Structure, 8*(2), 1–32.

Welser, H. T., Cosley, D., Kossinets, G., Lin, A., Dokshin, F., Gay, G., & Smith, M. (2011). Finding social roles in Wikipedia. In *Proceedings of the 2011 iConference* (pp. 122–129). New York: ACM.

Williford, C., Henry, C., & Friedlander, A. (2012). *One culture: Computationally intensive research in the humanities and social sciences*. Washington, DC: Council on Library and Information Resources.

Wohlsen, M. (2013). The only thing Amazon has to fear is Amazon itself. http://www.wired.com/business/2013/05/amazon-dominance. Accessed 1 Jan 2014.

Xie, P., & Xing, E. P. (July 2013). *Integrating document clustering and topic modeling*. Paper presented at the 29th International Conference on Uncertainty in Artificial Intelligence (UAI 2013), Bellevue, WA.

Zhang, D., & Lowry, P. B. (2008). Issues, limitations, and opportunities in cross-cultural research on collaborative software in information systems. *Journal of Global Information Management, 16*(1), 61–92.

Index

E. Bertino, S. A. Matei (eds.), *Roles, Trust, and Reputation in Social Media Knowledge Markets,* Computational Social Sciences, DOI 10.1007/978-3-319-05467-4,