

ScholarWorks@GSU

Polycentric Information Commons: A Theory Development and Empirical Investigation

Authors	Mindel, Vitali
Citation	Mindel, Vitali. "Polycentric Information Commons: A Theory Development and Empirical Investigation." 2018. Dissertation, Georgia State University https://doi.org/10.57709/13369847
DOI	https://doi.org/10.57709/13369847
Download date	2026-05-13 16:57:50
Link to Item	https://hdl.handle.net/20.500.14694/2604

Polycentric Information Commons: A Theory Development and
Empirical Investigation

By

VITALI MINDEL

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF

DOCTOR OF PHILOSOPHY

IN THE

ROBINSON COLLEGE OF BUSINESS

OF

GEORGIA STATE UNIVERSITY

GEORGIA STATE UNIVERSITY
ROBINSON COLLEGE OF BUSINESS
2018

Copyright by
Vitali Mindel
2018

ACCEPTANCE

This dissertation was prepared under the direction of the *Vitali Mindel* Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Business Administration in the J. Mack Robinson College of Business of Georgia State University.

Richard Phillips, Dean

DISSERTATION COMMITTEE

Chair: Dr. Lars Mathiassen (Center for Process Innovation & Computer Information Systems Department, Georgia State University)

Chair: Dr. Arun Rai (Center for Process Innovation & Computer Information Systems Department, Georgia State University)

Dr. Likoebe Maruping (Center for Process Innovation & Computer Information Systems Department, Georgia State University)

Dr. Brian Butler (College of Information Studies, University of Maryland)

ABSTRACT

Polycentric Information Commons: A Theory Development and Empirical Investigation

BY

VITALI MINDEL

11/26/2018

Committee Chair: *Lars Mathiassen & Arun Rai*

Major Academic Unit: *Center for Process Innovation & Computer Information Systems*

Decentralized systems online—such as open source software (OSS) development, online communities, wikis, and social media—often experience decline in participation which threatens their long-term sustainability. Building on a rich body of research on the sustainability of physical resource systems, this dissertation presents a novel theoretical framing that addresses the sustainability issues arising in decentralized systems online and which are amplified because of their open nature. The first essay develops the theory of polycentric information commons (PIC) which conceptualizes decentralized systems online as “information commons”. The theory defines information commons, the stakeholders that participate in them, the sustainability indicators of information commons and the collective-action threats putting pressure on their long-term sustainability. Drawing on Ostrom’s factors associated with stable common pool resource systems, PIC theory specifies four polycentric governance practices that can help information commons reduce the magnitude and impact of collective-action threats while improving the information commons’ sustainability. The second essay further develops PIC theory by applying it in an empirical context of “digital activism”. Specifically, it examines the role of polycentric governance in reducing the threats to the legitimacy of digital activism—a type of information commons with an overarching objective of instigating societal change. As such, it illustrates the applicability of PIC theory in the study of digital activism. The third essay focuses on the threat of “information pollution” and its impact on open collaboration, a type of information commons dedicated to the creation of value through open participation online. It uncovers the way polycentric governance mechanism help reduce the duration of pollution events. This essay contributes to PIC theory by expanding it to the realm of operational governance in open collaboration.

Acknowledgments

As corny as it may seem, I would like first and foremost to acknowledge the American Dream. I came to this country in 2007 and had to start from the bottom. For almost two years, I worked as a housekeeping aid in a hospital in Springfield, Illinois, cleaning surgery rooms for minimum wage on the evening shift while taking classes at the local community college during the day. I then moved to Atlanta, where I worked as a daytime groundskeeper in an apartment complex for three years, cleaning the pool and hallways. After taking a year off of my education to avoid paying out-of-state tuition, I began taking evening classes and completed my bachelor's degree at Georgia State University. To this day, when I see housekeepers in office buildings pushing carts equipped with spray bottles, paper towels, mop buckets, and other cleaning supplies, I get emotional. I pushed one of those carts for almost five years.

Now, on my way to be a college professor in one of the premier higher education institutions in the land, I feel extremely fortunate and grateful to all the experiences and people that helped and inspired me along the way. I want to thank Hanny, my mentor in the housekeeping department at Memorial Hospital in Springfield. For some reason, she took me under her wings, making sure I got the “easier” assignments so that I had extra time to do my college class work. I want to thank the crew at the hospital's Cath-Lab for being nice to me and being understanding of my slacking while I was buried in textbooks. I want to thank Dr. Ellen Watkins, my sociology professor at Lincoln Land Community College (LLCC), for not laughing at me when I told her that I wanted to become a college professor. She and other professors I encountered in LLCC—especially Dr. Mark Roehrs and Dr. John Vinzant—were all very encouraging. The quality of courses at LLCC was the best I have experienced in my long journey as a student, and though I did not receive my degree there, I still regard LLCC as my *alma mater*. I want to thank Cedrick, the maintenance

manager in the Atlanta apartment complex where I worked, for turning a blind eye when I was often doing homework instead of my job. All of these people noticed the drive in me and implicitly and explicitly supported my efforts. I thank you all for that.

I want to thank Dr. Jeannie Grussendorf, who was my favorite professor at Georgia State University and whose teaching style inspires me now when designing my own classes. She was the one who told me about the internship opportunity with the World Affairs Council of Atlanta (WACA), and who connected me to the right people and wrote my recommendation letter. I would like to thank Dr. Wayne Lord, the president of WACA at the time, for hiring me as his assistant and introducing me to Dr. Lars Mathiassen, the academic director of Georgia State University's Executive Doctorate Program, which shared office space with WACA. Without this introduction, I would not be writing these acknowledgments.

This brings me to Lars, my advisor and mentor. No words of gratitude can express how I feel about this man. I know for sure that I owe EVERYTHING to him. Before meeting Lars, I applied for dozens of jobs and graduate programs, thinking that graduating at the top of my class would help get my foot in the door. I was wrong; almost everyone shut the door on me. For some reason that is not entirely clear to me to this day, Lars took a chance on me, giving me the opportunity I needed to advance from cleaning floors to having a fulfilling life. I had no master's degree, no background in the discipline, and no industry experience, and still this man saw something in me. Working with Lars was incredible, and I will cherish this for the rest of my career. Thank you for taking the chance on me, Lars. Hope you don't regret it. Over the five years we spent together, I learned so much from you—the most important thing being that successful academic pursuit is not only about doing research, but also about cultivating friendships and relationships. I am truly inspired by the way in which you treat everyone you encounter with respect and genuine friendliness.

I know that Lars would not have admitted me to the program without the blessing of his close colleague, Dr. Arun Rai. Having worked mostly with Lars, I was at first intimidated by Arun. Not because of his status in the field (which I was unaware of) but because he seemed to be fiercely brilliant, intense, and unabashedly confident. I kept my distance for a while, but over time I was very happy to discover the true Arun Rai—brilliant, intense, and confident, yes, but also very much an outstanding human being who deeply cares about others. One should see the sparkle in his eyes when he talks about his students and other young aspiring academics! I learned a lot from working with Arun, including the importance of seeking constant inspiration and cultivating your passion for the academic pursuit. It is a little funny to me now to acknowledge how clueless I was when I started the program, with no idea who Arun Rai or Lars Mathiassen were; it was only after a while into the PhD program that I was pleasantly surprised to learn that I was being mentored by some of the most esteemed scholars in the IS field. It was a little like realizing that, after years jamming with Paul and John, I was actually making music with the Beatles.

I could not have asked for a better place to do my PhD than the Center for Process Innovation (CEPRIN). Everyone I met there—from the many visiting scholars to the PhD students and the professors—are incredible people. I want to thank Dr. Likoebe Maruping, who joined CEPRIN during my third year, for his guidance and for providing an example of excellence as a scholar, on top of being one of the nicest people one may encounter. I want to thank my fellow PhD students. Chaitanya, I appreciate your friendship and the many times you disrupted my work by starting hour-long conversations that had little to do with research. I needed that. Liwei and Jessica, thank you for helping me with various matters, ranging from registration and class selection to sharing your materials with me. Zhitao, I appreciate your wisdom and recognize that you are the smarter PhD student in our cohort. Yanran, I appreciate our many conversations, and you are truly a sweet

person. Peiwei and Khaleed, although I did not get to know you as well as some other students, I enjoyed our interactions; take care of Lars. All of you, my fellow students, have worked hard and inspired me to do better. I wish you all success in your pursuits! Finally, I want to acknowledge Vanessa Browne. Over time, we developed a close friendship, and I will miss dearly our conversations and laughs. With all respect to the others, you are my favorite person in Robinson College of Business. I always knew you had my back.

Table of Content

Acknowledgments.....	5
Chapter 1: Dissertation Overview.....	12
1.1 Introduction.....	12
1.2 Essay 1	17
1.3 Essay 2	18
1.4 Essay 3	20
1.5 References.....	23
Chapter 2: The Sustainability of Polycentric Information Commons.....	26
2.1 Sustainability of Decentralized Systems Online	27
2.2 Information Commons	32
2.3 Derived Stakeholder Value	38
2.3.1 Derived Producer Value.....	39
2.3.2 Derived Provider Value	40
2.3.3 Derived Appropriator Value	41
2.4 Sustainability.....	41
2.4.1 Provision	42
2.4.2 Appropriation.....	43
2.4.3 Revitalization	43
2.4.4 Equitability.....	44
2.5 Collective-Action Threats.....	45
2.5.1 Free-riding.....	47
2.5.2 Congestion	48
2.5.3 Pollution.....	49
2.5.4 Violation	50
2.5.5 Rebellion.....	51
2.6 Polycentric Governance Practices.....	51
2.6.1 Boundary Regulation	55
2.6.2 Incremental Adaptation.....	56
2.6.3 Shared Accountability.....	57
2.6.4 Provider Recognition	57
2.7 Theorizing Relationships	58

2.8 Discussion	65
2.9 References	72
Appendix	82
Reviewed IS Research	82
Chapter 3: Harnessing Digital Activism for Institutional Change: A Collective-Action Perspective	99
3.1 Decentralized Systems Online and Societal Change	100
3.2 Digital Activism	100
3.2.1 Digital Online Activism	102
3.2.2 Digital Offline Activism	103
3.3 Theoretical Framing	105
3.3.1 Digital Activism as Polycentric Information Commons	108
3.3.2 Digital Activism for Institutional Change	113
3.4 Research Method	116
3.5 Case Study Analysis	117
3.5.1 Institutional Field	119
3.5.2 Instigating Institutional Change (September 2012–March 2013)	120
3.5.3 Emerging Collective Action (April 2013–July 2013)	120
3.5.4 Enacting Polycentric Governance (May 2013–September 2013)	127
3.5.5 Cascading Collective Action (May 2013–March 2014)	127
3.5.6 Avoiding Provider Rebellion (November 2014–December 2014)	131
3.5.7 Institutional Change	136
3.6 Discussion and Theory Development	136
3.7 Conclusion	147
3.8 References	147
Chapter 4: Polycentric Resolution of Information Pollution in Open Collaboration Systems	158
4.1 Information Pollution in Open Collaboration Systems Online	159
4.2 Addressing Information Pollution in Open Collaboration	162
4.3 Information Pollution in Polycentric Information Commons	167
4.3.1 Operational Level of Governance	169
4.3.2 Polycentric Principles	171
4.4 Study Design and Hypotheses	174
4.4.1 Dependent Variable	175
4.4.2 Independent Variables	176

4.4.3 Control Variables	181
4.5 Dataset Construction	181
4.6 Analysis Results	183
4.6.1 Descriptive Statistics	183
4.6.2 Model Specification and Analysis Procedures	185
4.6.3 Results	186
4.6.4. Robustness and Additional Analyses	188
4.7 Discussion	192
4.7.1 Resolving Information Pollution in Open Collaboration Systems	193
4.7.2 Contribution to PIC Theory	195
4.7.3 Practical Implications	199
4.8 References	201
Appendix A: Dataset Construction	208
Appendix B: Robustness Analysis	212

CHAPTER 1: DISSERTATION OVERVIEW

1.1 Introduction

Decentralized systems online—such as open source software (OSS) development, online communities, wikis, and social media—are highly accessible, dispersed and propelled by voluntary participation that varies in its intensity and stability (von Krogh and Spaeth 2007). The openness and decentralization of these systems has been shown to spur incredible outcomes (Benkler and Nissenbaum 2006; Tapscott and Williams 2008). Software created by volunteers and openly distributed online propels private and government organizations around the world and is used by private individuals daily. Online communities and social media platforms connect people and improve their well-being in countless ways, ranging from providing entertainment and professional functionality to offering emotional support and information. Wikis provide content on all sorts of topics, and are increasingly widely used as a source of information for lay people as well as a wide range of professionals, including academics, healthcare professionals, journalists, lawyers, and judges (Brokowski and Sheehan 2009; Brown 2011; Lim and Simon 2011; Miller and Murray 2010; Peoples 2009). The impact that decentralized systems online have on our everyday lives cannot be overstated; they impact how we spend our leisure time, connect, learn, absorb news and cultural trends, shop and find other transaction opportunities, and much more.

In addition to affecting the everyday lives of anyone with an internet connection, the openness and connectivity-potential of decentralized systems online has become a major catalyst for new business models. Some of the most valued companies today started not so long ago as simple platforms on which individuals from all walks of life were invited to freely participate. At first, in the mid-90s, websites such as eBay and Craigslist invited individuals and small businesses to

advertise all sorts of goods and services with little restrictions. At the same time, online dating was born and Match.com invited individuals to advertise themselves to others to promote a different type of transaction. In the early 2000s, the internet's openness and connectivity gave rise to social media websites such as Myspace and Facebook, where people can freely share personal information, photos, and videos with anyone they choose to include in their virtual social circle. Social media's value proposition is similar to that of transaction-oriented websites such as eBay and Craigslist and even Match; all are based on the notion of virtuous network effects (although the revenue model has shifted from charging brokerage and platform usage fees to selling eyeballs to advertisers). Half a decade later, the multisided platform business model progressed with the advent of smartphone technologies that incorporate global positioning systems (GPS) to allow people to coordinate transactions in real time. Multisided platform services driving the emerging sharing economy is exemplified by Uber and Lyft, which—with few restrictions—allow people with cars to offer rides for a fee to people needing a ride in real time. Similarly, Airbnb, also with little restrictions, allows individuals with extra living space to offer it for a fee to those looking for a place to stay. The idea of using decentralized online technologies to invite people to freely transact with each other or simply exchange information proved to be business gold; as of Summer 2018, the companies mentioned above are estimated to be worth collectively more than half a trillion dollars.¹

The increase in the importance of decentralized systems online is well reflected in IS research, which shifted its focus in the past decade from “traditional” business systems such as enterprise resource planning (ERP) to the various “new” emerging systems online. In 2005, only three research papers on decentralized systems online were published in the top two IS journals—

¹ Yahoo Finance

Management Information Systems Quarterly (MISQ) and *Information Systems Research (ISR)*; 10 years later, in 2014, the same two journals published 26 research papers on the subject, reflecting a steady trend (see Table 1.1).

Year	MISQ	ISR	MISQ+ISR
2014	13	13	26
2013	5	17	22
2012	6	6	12
2011	4	4	8
2010	3	2	5
2009	3	0	3
2008	0	8	8
2007	3	2	5
2006	3	4	7
2005	2	1	3
Total	42	57	99

Although existing research has made some important theoretical contributions to the study of decentralized systems online (Bateman et al. 2011; Butler et al. 2014; Howison et al. 2014; Kane et al. 2012; Levina et al. 2014), most studies in this area are light on theory. To the extent that researchers use theories, those theories are typically longstanding frameworks developed in other social science fields. Borrowing established “outside” theories for framing research is a common practice in the IS field. Still, the lack of new theories is problematic for two main reasons: (i) it limits the cross-fertilization of insights, which contributes to (ii) theoretical stagnation in an increasingly mature field. The shortage of new theories on decentralized systems online is noticeable, and several prominent researchers have called for their development (Johnson et al. 2014; Majchrzak 2009; Singh et al. 2011; Von Krogh et al. 2012). New theories are needed to

consolidate the many empirical insights generated shortly after the emergence of decentralized systems online and to serve as a stepping stone for future research in the field.

In response to the identified need, this dissertation advances the theoretical understanding of governance of decentralized systems online to help address the challenges that arise due to their high degree of openness to participation (Table 1.2 summarizes the dissertation essays). The dissertation presents a novel theoretical framework for examining the systems' sustainability in the face of degenerative threats. In addition, it exemplifies the theory's versatility and applicability by providing evidence from two distinct research contexts and research methods. The dissertation then leverages the distinctive aspects of these contexts, coupled with different modes of inquiry, to further advance theory regarding decentralized information systems online. The first empirical research study—an explanatory case study—aims to contribute to the understanding of how decentralized systems online may be harnessed for promoting institutional change. The second empirical research—a quantitative study—aims to contribute to our understanding of how open collaboration (OC) systems online resolve arising information pollution problems. The developed theory, coupled with the two empirical investigations that build upon it, contribute to a better understanding of governance of decentralized systems online.

Table 1.2. Summary of Context, Theory, and Methods

	Essay 1	Essay 2	Essay 3	
Context	<i>Technological Context</i>	Online communities, social media, multisided platforms, online reviews, wikis, crowdsourcing	Smartphone-based crowdsourcing and social media	Open collaboration (OC) systems online
	<i>Area of research</i>	Theory development on the governance of decentralized systems online	Digital activism for promoting institutional change	OC systems governance against information pollution
	<i>Target audience</i>	Researchers of decentralized systems online; firms using decentralized systems online as part of their business model	Researchers of technology and social activism; organizers of social activism	Researchers of OC systems; participants in OC systems
Theory	<i>Informing theoretical perspectives</i>	Tragedy of the commons; common pool resources (CPR) governance	Polycentric information commons; legitimacy and institutional change	Polycentric information commons
	<i>Data source</i>	Coding papers published in <i>MISQ</i> and <i>ISR</i> between 2005 and 2014	Interviews with key informants; electronic communications, and official and media reports	Wikipedia data release processed with a Python script developed by coauthor Aleksi Aaltonen
Method	<i>Sample</i>	72 articles	18 interviews, 95 emails, 43 Facebook communication streams, 439 Facebook posts, 621 media stories, 7 official reports	4,325 pollution events and the editing activities in Wikipedia between the posting of a “cleanup” tag and its removal
	<i>Analysis approach</i>	Manual coding	Episodes and encounters analysis of qualitative data	Hierarchical linear regression modeling

1.2 Essay 1

To add to the body of theories on decentralized information systems online, Essay 1 focuses on the theoretical question of how these systems may be governed to achieve sustainability in the face of threats arising from within because of their openness to participation.² At present, the literature examining the question of sustainability of decentralized systems online is limited and focused mostly on online communities. This stream of research finds that an increase in the community's size reduces the efficiency of communication, and that participants are generally sensitive to barriers that raise the opportunity cost of engagement; it also identifies an inherent tension between the need to attract new members while also avoiding alienating old-timers that might be dissatisfied with the changes that new participants bring (Butler 2001; Butler and Wang 2012; Butler et al. 2014). The theory development essay seeks to further advance our understanding about the sustainability of decentralized information systems online by identifying the threats impacting them and the conditions under which these threats may be addressed.

Drawing on the well-established paradigms of *tragedy of the commons* (Hardin 1968) and *polycentric* governance of common pool resource (CPR) systems (Ostrom 1990), in Essay 1, we³ analyze a sample of 72 research articles published in *MISQ* and *ISR* on decentralized systems online. We appropriate the Hardin and Ostrom terminologies to code the literature and create a “common language” repository of insights that we subsequently use, in conjunction with Ostrom's findings on CPR governance in the physical world, to develop a novel theoretical framework. Conceptualizing decentralized systems online as “information commons” because of their high

² Mindel, V., Mathiassen, L. and Rai, A., 2018. “The Sustainability of Polycentric Information Commons,” *MIS Quarterly* (42:2), pp 607-631.

³ The paper was coauthored with Lars Mathiassen and Arun Rai

degree of openness to participation, the theory of polycentric information commons (PIC), theorizes on the categories of actors deriving value from participating in information commons; the factors associated with their sustainability; the collective-action threats that arise from within and threaten their sustainability; and the polycentric governance practices that may help them reduce the prevalence of threats and strengthen them in the long run.

Understanding the sustainability conditions for the various web-based decentralized information systems is important for theory for two reasons: (i) these systems represent a novel mode of information production and governance that is distinctly different from past hierarchical models (Benkler 2006; Benkler and Nissenbaum 2006), and (ii) they impact societal outcomes, ranging from affecting markets (Dellarocas 2005; Xu and Zhang 2013) and economic displacement (Chan and Ghose 2014) to outcomes in politics (Gibson and McAllister 2011; Wattal et al. 2010), health (Barak et al. 2008; Chan and Ghose 2014) and education (Agazio and Buckley 2009; Burke et al. 2009). Hence, it is important to understand theoretically what differentiates successful systems from the numerous systems that become unsustainable. Moreover, understanding practically how decentralized information systems may be governed to improve their sustainability odds becomes increasingly important for companies relying on these systems as part of their business model; those companies include YouTube, Airbnb, Uber, eBay, Waze, Facebook, Yelp, TripAdvisor, Twitter, and Craigslist, all of which rely on self-selecting volunteers to provide the bulk of their content.

1.3 Essay 2

Decentralized systems online share the characteristic of open access to participants, but otherwise differ in their overarching goals. Increasingly, researchers from various disciplines examine how

decentralized systems online are being used to promote social change (Agarwal et al. 2014; Aouragh and Alexander 2011; Bennett and Segerberg 2011; Edwards et al. 2013). The literature demonstrates that the use of online communities, social media, crowdsourcing, and crowdfunding online can be useful for fundraising and promoting awareness for various causes, but they are not very impactful beyond these narrow objectives (Bimber et al. 2005; DeLuca et al. 2012). Digital activism online also suffers from the negative perception that it is simply an easy way for people to feel as if they are doing something without actually engaging with the issues beyond cyberspace (Butler 2011). Even when decentralized systems online enable mobilization offline in the real world, the resulting impact is typically disappointing (Friedersdorf 2015; White 2016). This is typically because the congregation of people sharing similar views feeds the echo-chamber effect of continuous reinforcement of agitation among activists (Garrett 2009), which alienates non-activists, while the lack of strong leadership makes it hard for social movements born online to articulate a clear and unified agenda (White 2016, 2017). Thus, despite their rising profile in promoting social change, we still do not fully understand how decentralized systems online can be effectively harnessed to achieve lasting institutional change.

Drawing on PIC theory (Essay 1) and legitimacy in institutional change theory (Suchman 1995), Essay 2 examines how harnessing digital activism online and offline can effectively promote institutional change. Unlike most information commons, digital activism is less concerned about its long-run sustainability and more about achieving a result, whether it be raising money, signing people on a petition, passing a new legislation or reversing an existing one, pressuring an organization to change its practices, or even changing practices across institutions. In our research, we examine the case of Cabotagestudien (CS), a research initiative in Sweden in which researchers used smartphone crowdsourcing and social media to collect data. The generated data, to the

surprise of many, impacted large-scale institutional change, including noticeably shifting public opinion, shifting political alliances, passing new legislations, and changing industry practices. In Essay 2, we analyze CS through the prism of PIC and legitimacy theories. We find that for digital activism to achieve noticeable institutional change, it must establish its pragmatic, moral, and cognitive legitimacy, which is difficult when facing collective-action threats arising from within as well as attacks from opposing actors. By adopting more of a polycentric approach to managing the decentralized systems facilitating it, digital activism is more likely to overcome such treats and achieve the needed legitimacy to impact change.

Essay 2 contributes to the literature on digital activism by showing how it can be harnessed for institutional change. Drawing on the prisms of PIC and legitimacy theories, Essay 2 uses the empirical findings to formulate broad theoretical propositions on the relationship between digital activism and institutional change. In addition, Essay 2 contributes to PIC theory by illustrating its applicability for examining decentralized systems online and for qualitative research in general. Finally, Essay 2 contributes to practice by outlining the problems social movement organizers are likely to encounter in digital activism and how those problems can be lessened by incorporating polycentric practices in the governance and design of their initiatives.

1.4 Essay 3

OC systems online are a type of information commons in which providers of information build on each other's work to create value for appropriators in the form of a knowledge repository or software that is made freely available as an open-access alternative to professionally created retail products (Tapscott and Williams 2008). Much has been written about the OC model's novelty—specifically, how products created by mostly uncoordinated and unpaid volunteers often prove to have quality that is equivalent to products produced by paid specialists (Benkler 2006; Benkler

and Nissenbaum 2006; von Hippel and von Krogh 2003). This fact stands in stark contrast to traditional economic theories. No theory of production could have predicted that unorganized and uncompensated self-selecting volunteers can sustain complex enterprises requiring constant updating and maintenance to keep generating value over time (Tapscott and Williams 2008). OC systems emerged in the past two decades, and our limited time perspective prevents us from knowing who is right: those who hype OC systems as a superior way of organizing production, or those who doubt the capacity of OC systems to continuously provide value over the long run. The conflicting findings from empirical research on the quality of OC system outputs—some finding it to be on par with professionally created products (Brown 2011; Chesney 2006; Giles 2005), while others finding it to be subpar (Holman Rector 2008; Kupferberg and Protus 2011; Lavsa et al. 2011)—continue to remind us that the questions of how OC systems resolve quality issues is important for their long-term viability.

At present, research on the quality of OC systems is exploratory and almost entirely atheoretical, providing insights on certain factors associated with quality, but for the most part unable to explain how OC systems successfully manage to balance their high-level inclusiveness and openness to participation with inevitably arising quality problems. Using PIC theory (Essay 1) as our lens, Essay 3 examines how OC systems address quality issues through the prism of polycentric governance and information pollution resolution. PIC theory asserts that all information commons are more susceptible than professionally organized production to incidental and even deliberate pollution. When unlimited numbers of mostly anonymous people are free to upload content with little supervision, PIC theory asserts that some of it is bound to conflict with the overarching goal of the information commons. Examples of information pollution include fake news,

unsubstantiated rumors and conspiracy theories, manipulated and biased content, and all sorts of incomplete and erroneous information.

In Essay 3, we zero-in on how polycentric practices of *shared accountability*, *boundary regulation*, *incremental adaptation*, and *provider recognition* are associated with pollution event resolution in Wikipedia—one of the most successful and arguably the most polycentric OC system to date. Conceptualizing that the posting of tags calling for the cleanup of articles is the start of “pollution events,” we examine the relationship between the four polycentric practices mentioned above and the time until the tag is removed, which signifies the temporary pollution resolution. We find that shared accountability and incremental adaption is directly associated with the reduction in the time it takes to resolve pollution. Boundary regulation, on the other hand, at first increases pollution resolution time to a point; thereafter, it is associated with a decrease in pollution resolution time. This suggests that boundary regulation initially creates confusion about the article’s boundary, but after the confusion is resolved, it helps it to improve information quality. Finally, although Wikipedia does not have a provider recognition mechanism at the article editing level, we find evidence that persistent provider rejection—that is, reverts of revisions—is associated with a decrease in the pollution resolution time. Essay 3 contributes to both the literature on OC system quality management and to PIC theory by confirming the relationship between polycentric practices and pollution alleviation.

1.5 References

- Agarwal, S. D., Barthel, M. L., Rost, C., Borning, A., Bennett, W. L., and Johnson, C. N. 2014. Grassroots Organizing in the Digital Age: Considering values and Technology in Tea Party and Occupy Wall Street,” *Information, Communication & Society* (17:3), pp 326-341.
- Agazio, J. and Buckley, K.M., 2009. “An untapped resource: Using YouTube in nursing education,” *Nurse educator*, (34:1), pp23-28.
- Aouragh, M., and Alexander, A. 2011. “The Arab Spring| the Egyptian Experience: Sense and Nonsense of the Internet Revolution,” *International Journal of communication* (15:5), pp 1344-1358.
- Barak, A., Boniel-Nissim, M. and Suler, J., 2008. “Fostering empowerment in online support groups,” *Computers in Human Behavior*, (24:5), pp 1867-1883.
- Bateman, P. J., Gray, P. H., and Butler, B. S. 2011. “Research note-the impact of community commitment on participation in online communities,” *Information Systems Research* (22:4), pp 841-854.
- Benkler, Y. 2006. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*, Yale University Press.
- Benkler, Y., and Nissenbaum, H. 2006. “Commons-based peer production and virtue*,” *Journal of Political Philosophy* (14:4), pp 394-419.
- Bennett, W. L., and Segerberg, A. 2011. “Digital Media and the Personalization of Collective Action: Social Technology and the Organization of Protests Against the Global Economic Crisis,” *Information, Communication & Society* (14:6), pp 770-799.
- Bimber, B., Flanagin, A. J., and Stohl, C. 2005. “Reconceptualizing Collective Action in the Contemporary Media Environment,” *Communication Theory* (15:4), pp 365-388.
- Brokowski, L., and Sheehan, A. H. 2009. “Evaluation of pharmacist use and perception of Wikipedia as a drug information resource,” *Annals of Pharmacotherapy* (43:11), pp 1912-1913.
- Brown, A.R., 2011. “Wikipedia as a data source for political scientists: Accuracy and completeness of coverage,” *PS: Political Science & Politics* (44:2), pp 339-343.
- Brown, A.R., 2011. “Wikipedia as a data source for political scientists: Accuracy and completeness of coverage,” *PS: Political Science & Politics* (44:2), pp 339-343.
- Butler, B. S. 2001. “Membership size, communication activity, and sustainability: A resource-based model of online social structures,” *Information Systems Research* (12:4), pp 346-362.
- Butler, B. S., and Wang, X. 2012. “The cross-purposes of cross-posting: Boundary reshaping behavior in online discussion communities,” *Information Systems Research* (23:3-part-2), pp 993-1010.
- Butler, B. S., Bateman, P. J., Gray, P. H., and Diamant, E. I. 2014. “An Attraction-Selection-Attrition Theory of Online Community Size and Resilience,” *MIS Quarterly* (38:3), pp 699-728.
- Butler, M. 2011. *Clicktivism, Slacktivism, or “Real” Activism Cultural Codes of American Activism in the Internet Era*. University of Colorado at Boulder.
- Chan, J. and Ghose, A., 2014. Internet’s dirty secret: Assessing the impact of online intermediaries on HIV transmission. *MIS Quarterly* (38:4), pp 955-975.
- Chesney, T. 2006. “An empirical examination of Wikipedia’s credibility,” *First Monday* (11:11).

- Dellarocas, C. 2005. "Reputation mechanism design in online trading environments with pure moral hazard," *Information Systems Research* (16:2), pp 209-230.
- DeLuca, K. M., Lawson, S., and Sun, Y. 2012. "Occupy Wall Street on the Public Screens of Social Media: The Many Framings of the Birth of a Protest Movement," *Communication, Culture & Critique* (5:4), pp 483-509.
- Edwards, F., Howard, P. P. N., and Joyce, M. 2013. "Digital Activism and Non-Violent Conflict," Available at SSRN: <https://ssrn.com/abstract=2595115>
- Friedersdorf, C. 2015. "Will Black Lives Matter Be a Movement That Persuades?" *The Atlantic Online*.
- Garrett, R. K. 2009. "Echo Chambers Online?: Politically Motivated Selective Exposure Among Internet News Users," *Journal of Computer-Mediated Communication* (14:2), pp 265-285.
- Gibson, R.K. and McAllister, I., 2011. "Do online election campaigns win votes? The 2007 Australian "YouTube" election," *Political Communication* (28:2), pp.227-244.
- Giles, J. 2005. "Internet Encyclopedias go head to head," *Nature* (438:7070), pp 900-902.
- Hardin, G. 1968. "The tragedy of the commons," *Science* (162:3859), pp 1243-1248.
- Hippel, E. V., and Krogh, G. V. 2003. "Open Source Software and the "Private-Collective" Innovation Model: Issues for Organization Science," *Organization Science* (14:2), pp 209-223.
- Holman Rector, L. 2008. "Comparison of Wikipedia and other Encyclopedias for Accuracy, Breadth, and Depth in Historical Articles," *Reference Services Review* (36:1), pp 7-22.
- Kupferberg, N., and Protus, B. M. 2011. "Accuracy and completeness of drug information in Wikipedia: an assessment," *Journal of the Medical Library Association* (99:4), pp 310-322.
- Lavsa, S.M., Corman, S.L., Culley, C.M. and Pummer, T.L., 2011. "Reliability of Wikipedia as a medication information source for pharmacy students," *Currents in Pharmacy Teaching and Learning* (3:2), pp 154-158.
- Lim, S. and Simon, C., 2011. "Credibility judgment and verification behavior of college students concerning Wikipedia," *First Monday* (16:4).
- Majchrzak, A. 2009. "Comment: Where is the Theory in Wikis?" *MIS Quarterly* (33:1), pp 18-20.
- Miller, J. C., and Murray, H. B. 2010. "Wikipedia in court: When and how citing Wikipedia and other consensus websites is appropriate," *St. John's Law Review* (84), pp 633-656.
- Peoples, L.F., 2009. "The Citation of Wikipedia in Judicial Opinions," *Yale Journal of Law & Technology* (12), pp 1-51.
- Suchman, M. C. 1995. "Managing Legitimacy: Strategic and Institutional Approaches," *Academy of Management Review* (20:3), pp. 571-610.
- Tapscott, D., and Williams, A. D. 2008. *Wikinomics: How mass collaboration changes everything*. Penguin.
- von Krogh G. and Spaeth, S., 2007. "The open source software phenomenon: Characteristics that promote research," *The Journal of Strategic Information Systems* (16:3), pp 236-253.
- Wattal, S., Schuff, D., Mandviwalla, M., and Williams, C. B. 2010. "Web 2.0 and Politics: The 2008 US Presidential Election and an E-Politics Research Agenda," *MIS Quarterly* (34:4), pp 669-688.
- White, M. 2016. *The End of Protest: A New Playbook for Revolution*. Knopf Canada.

White, M. 2017. "Occupy and Black Lives Matter Failed. We can either Win Wars or Win Elections," *The Guardian online*

Xu, S. X., and Xiaoquan (Michael) Zhang. 2013. "Impact of Wikipedia on Market Information Environment: Evidence on Management Disclosure and Investor Reaction," *MIS Quarterly* (37:4), pp 1043-1068.

CHAPTER 2: THE SUSTAINABILITY OF POLYCENTRIC INFORMATION COMMONS

ABSTRACT

Research on various distributed online information systems—including blogging, crowdsourcing, media sharing, online communities, online reviews, open source software development, social media, wikis, peer-to-peer file sharing, and two-sided electronic markets—shows that the level of user engagement and overall activity in most systems eventually decline substantially. Here, we draw on Hardin’s theory of the tragedy of the commons and Ostrom’s theory of polycentric governance to introduce a unifying theory of polycentric information commons (PIC) that explains these phenomena. Further, our theory illuminates how polycentric governance principles, as manifested in system rules and infrastructure features, counterbalance various sustainability threats arising from unrestricted participation. By integrating previous research findings and offering new insights into information and governance practices, the theory, practically applied, can enhance the likelihood of sustained participation across diverse, decentralized online information systems. We conclude by discussing how researchers can use the theory in empirical investigations and how they can engage in theoretical elaborations.

Key Words: Theory development, tragedy of the commons, polycentricity, governance, decentralized online information systems, collective-action threats

2.1 Sustainability of Decentralized Systems Online

We increasingly depend on information systems (IS) with a high degree of volunteer user participation and consequential sustainability threats, including blogging, crowdsourcing, media sharing, online communities, online reviews, open source software (OSS) development, social media, wikis, peer-to-peer (P2P) file sharing, and two-sided electronic markets. Although many researchers have studied these decentralized online information systems and produced important insights over the past several years, the literature remains fragmented and light on theory (Majchrzak 2009; Singh et al. 2011). With the exception of the theory of social dynamics in online settings (Levina et al. 2014), which explains power relations across various user-generated content platforms, the few theories recently published are primarily phenomenon-specific (Bateman et al. 2011; Butler et al. 2014; Howison et al. 2014; Kane et al. 2012) and have limited applicability beyond their explicit context of origin. It is not surprising, then, that researchers are increasingly calling for “more highly socialized and multitheoretic explanations of community development” (Johnson et al. 2014) that “cover the interplay with institutions, goods, and the social practice” (Von Krogh et al. 2012). In response, our work integrates contemporary empirical evidence in IS research and theoretical insights from the literature on the sustainability of online communities (Butler 2001; Butler et al. 2014; Butler et al. 2012) with the literature on the governance of common pool resource (CPR) institutions (Ostrom 1990) to develop a theoretical framework aimed at explaining and predicting outcomes in decentralized online information systems.

While undoubtedly different in many regards, the online phenomena cited above share three important characteristics. First, relative to traditional information systems, decentralized online

information systems are (1) highly accessible⁴ to content consumers (typically at no cost) and (2) highly accessible to content producers who, in most cases, engage without payment. These features give rise to a third: (3) high-accessibility characteristics, which let individuals join for free and leave anytime, can lead to high volatility in consumer and producer participation. Although commercial (and sometimes nonprofit) organizations create and maintain the technical infrastructure of these systems, most of their content is not provided by system owners; rather, it is generated in an unordered, *decentralized* fashion by volunteer participants. This freedom of participation creates the conditions for sudden growth, but it also makes the systems vulnerable to sudden massive exits of content producers. This open access and high dependency on self-selecting individuals is the foundation on which we build our theory, guided by the following overarching research question: *How can decentralized online information systems mitigate the threats to sustainability that follow from their high degree of openness to participation?*

For centuries, ecologists, demographers, economists, sociologists, and political scientists have studied and debated questions related to the sustainability of various types of openly accessible resource systems. The discourse on the topic can be coarsely categorized into two prominent—yet rival—schools of thought. One school is best represented by ecologist Garret Hardin, and the other by political scientist Elinor Ostrom.

In 1968, Hardin published his influential thought experiment in which he describes the dissipation of medieval-type grazing commons resulting from unchecked individual-level overgrazing and pollution. “The tragedy of the commons,” as it came to be known, grew to be the leading paradigm in political science and economics, taught in college classes and public administration circles as a

⁴ Throughout the paper, we use the term “accessibility” in the sense of “hard to exclude people” as opposed to *technical* accessibility (i.e., internet access, ease of access, and so forth).

cautionary tale of what might happen to collective resources if strong, top-down institutions do not sufficiently curtail individual freedoms (Hardin 1968). According to the logic of the tragedy of the commons, decentralized online information systems will eventually decline and become unsustainable because of their high level of openness. Indeed, empirical evidence suggests that many online communities, after initial growth, have experienced a substantial decline in participation (as in Napster and MySpace) or ceased activity all together (Butler 2001; Butler et al. 2014; Butler et al. 2012; Hann et al. 2013; Ma et al. 2007; Moon et al. 2008; Ransbotham et al. 2011; Stewart et al. 2006; Zhang et al. 2013). Even high-profile systems such as Wikipedia, according to observers studying them, show signs of declining participation (Halfaker et al. 2011). The rise of cyber-archeology, which “digs” into failed digital communities and their digital artifacts, further illustrates the tendency of decentralized online information systems to decline or collapse (Harrison 2009; Jones 1997). Can the decline of these systems be viewed, at least partially, as a form of tragedy of the commons—that is, an inevitable outcome resulting from “too much” freedom in how people engage and interact with these technologies? If so, how can we then reconcile the supposed inevitability of decline with the fact that so many successful decentralized online information systems such as YouTube and Facebook manage to remain robust?

According to Ostrom, the tragedy of the commons oversimplifies reality, and its conclusion—that a central authority must significantly curtail individual freedoms—is highly problematic⁵ (Ostrom 1990). Ostrom and her colleagues analyzed thousands of cases of local resource governance arrangements (such as grazing grounds, fisheries, forests, and watersheds) and, while they did find instances of tragedies, they also found numerous examples of sustainable and well-functioning

⁵ Ostrom challenged Hardin’s notion of open access commons and instead maintained that Hardin in effect was describing a common property regime which, although permeable, is not ungovernable.

resource systems that operate without strong centralized authorities (Nagendra et al. 2012; Ostrom 1990; Schlager et al. 1999). More often than not, local communities, for hundreds of years in some cases, were found to be successful in governing CPRs with minimal oversight from central authorities (Ostrom 1990). That is, contrary to the tragedy of the commons, in most cases people communicate and work together to find sensible ways to share resources without endangering their long-term sustainability. Ostrom observed that the most resilient governance arrangements were those that dynamically managed boundary setting and mutual accountability through a high degree of inclusivity in decision-making. Increasing the number of independent decision-making centers generates system stability by reinforcing individual commitment to the CPR rather than delegating this responsibility to a narrow, elite group of decision makers. These boundary-setting and mutual-accountability mechanisms constitute local arrangements that keep wrongdoers at bay without compromising the freedoms of others. The mechanisms emerge from mutual adjustments among the involved autonomous actors, rather than relying primarily on central governance structures or private market forces; they thus represent a third mode of governance: *polycentricity*. Following this school of thought, could one major differentiating factor between successful and unsuccessful decentralized online information systems be rooted in the extent to which they embed polycentric governance principles in their design?

Drawing on these important distinctions and insights, we conceptualize decentralized online information systems as “information commons.” Although these various online systems differ in structure and objectives, they are similar in at least two important ways: (i) they are information systems, and (ii) they resemble Hardin’s commons in their openness. Adopting this as a frame, we suggest that the extent to which polycentric governance principles are embedded into those information commons explains why some systems decline and others flourish, even in the face of

many collective-action threats resulting from their inherent openness. Although polycentricity cannot explain everything we need to know about the sustainability of decentralized online information systems, we use it as a starting point for developing a foundational theory. While we build on Ostrom's insights, we acknowledge that decentralized online information systems are sociotechnical phenomena that do not entirely fit the definition of CPR systems. The difference in contexts necessitates a careful borrowing and adaptation of specific concepts to enhance the theory development beyond a simple overt reproduction that would fail to increase our understanding of the phenomena at hand (Whetten et al. 2009). Hence, in our theorizing, we rely on both our reasoning and our creative imagination (Bacharach 1989; Rivard 2014) supported by two literature streams: (i) the general literature on polycentric governance of shared resources, and (ii) the literature on decentralized online information systems (see the Appendix for details of our IS literature review).

We structure the paper in accordance with Zmud's (1998) recommendations to develop "pure" IS theory papers by defining the principle phenomena, explicating core concepts, and relating them to each other while articulating their *raison d'être*. Accordingly, we next specify the proposed theory's content and boundaries by defining *information commons* and specifying the many different IS phenomena that fit the definition. We subsequently conceptualize and define the core concepts underlying our theoretical development: *derived stakeholder value*, *sustainability*, *collective-action threats*, and *polycentric governance practices*. We then develop a conceptual model and associated propositions that relate the core concepts to each other and to their constituent constructs. We conclude by discussing the contributions of our theorizing and how the proposed theoretical framing can serve as a generative mechanism for further theoretical elaboration and empirical evaluation of information commons.

2.2 Information Commons

To facilitate analysis of their sustainability, we conceptualize decentralized online information systems as information commons—that is, a *highly accessible, self-rising information system in which stakeholders share an overarching goal*. To conceptually unify several streams of research, we address information commons in their generic form, acknowledging that future research should explore differences across their specific forms, such as whether an information commons is stand-alone, nested within traditional systems, or entangled with other information commons. To set the boundary conditions for our theorizing, we draw on Ostrom’s work to conceptualize information commons (Table 2.1) and describe contemporary phenomena that represent information commons (Table 2.2).

Both Hardin and Ostrom identify *accessibility* as a major defining characteristic of the commons and CPR systems, respectively (Hardin 1968; Ostrom 1990). Accordingly, relative to other information systems and platforms, information commons are highly accessible. Although the access is not necessarily unlimited, information commons are generally nonexclusive, allowing anyone who wishes—regardless of experience or credentials—a high degree of freedom of entrance and autonomy to supply and acquire digitalized information in the form of words, videos, images, sound, and code (Hann et al. 2013; Shah 2006; Stewart et al. 2006). The flip side of freedom of access is freedom of exit. Individual participants can easily leave the information commons at any time.

	CPR Concept		Adapted Concept	Examples	
	Common Pool Resource System	“...natural or man-made resource system that is sufficiently as to make it costly to exclude potential beneficiaries from obtaining benefits from its use” (E. Ostrom 1990, pp. 30)	Information Commons	Highly accessible, self-rising information system in which stakeholders share an overarching goal	Blogging, crowdsourcing, media sharing, online communities, online reviews, OSS development, P2P file sharing, two-sided electronic markets, social media, wikis
	Physical Resources	Natural and man-made	Information	Intangible; digitalized and man-made	Words, icons, videos, images, sound, code
Actors	Producers	“Anyone who actually constructs, repairs, or takes actions that ensure the long-term sustenance of a resource systems itself” (E. Ostrom 1990, pp. 31)	Producers	Architects and sponsors of the infrastructure that enable the system	The Mozilla foundation, Wikimedia foundation, Turner Broadcasting
	Providers	“...those who arrange for the provisioning of a CPR” (E. Ostrom 1990, pp. 31)	Providers	People who supply information to the system	Developers, uploaders, posters, content producers
	Appropriators	“...those who withdraw [resource units]” (E. Ostrom 1990, pp. 30)	Appropriators	People who acquire information from the system	Readers, downloaders, content consumers

Hardin warned about the dire consequences of accessibility; however, Ostrom, basing her assertion on a substantial body of empirical observations, argued that, in the majority of cases, a high degree of accessibility prompts the emergence of a self-rising order. Taking Ostrom’s position, we identify *self-rising order from within*—as opposed to top-down design—as the second characteristic of information commons. Prior IS research addresses this self-rising property in different forms of information commons. For instance, OSS development research has referred to

the property as “self-emerging” (Hann et al. 2013), while online community research calls it “inherently evolving” (Butler et al. 2014). Moreover, blogs have different centers of influence that emerge spontaneously (Chau et al. 2012; Howison et al. 2014), and online communities involve autonomous centers that emerge and disappear without any predefined pattern (Johnson et al. 2014). Although the self-rising property of information commons is widely recognized in extant literature, it is rather unconventional. Throughout history, traditional information systems—from the hieroglyphs in the Great Pyramid to printed newspapers—were produced by specialists working within hierarchical structures of authority. Hence, based on the self-rising property alone, information commons are distinctly unique types of information systems.

In these systems, people play various roles; IS research describes these roles using many different terms, including users, consumers, participants, contributors, developers, uploaders, bloggers, and posters. To further clarify our definition of information commons, we adapt Ostrom’s (1990) concepts to classify these principal stakeholders into three distinct categories: (i) *producers*—the architects and sponsors of the infrastructure that enable the system; (ii) *providers*—the people who supply information to the system; and (iii) *appropriators*—the people who extract information from the system. Of course, an individual may play more than one role; most people are both providers and appropriators of content on social media platforms, for example. Distinguishing among the principal actors involved in information commons also facilitates cross-case analyses (Darke et al. 1998). We expand further on these stakeholders in the next section.

Every CPR system has an overarching purpose, whether it is distributing water, managing grazing rights, determining tree-cutting schedules, restricting fishing activities, or allocating specific common resources. Similarly, every information commons has an overarching goal, which might be broad or specific, continuous or ephemeral, legal or illegal. While stakeholders may be

motivated by different considerations and goals, they inevitably share some overarching goal that ties them to the information commons. For instance, sellers and buyers in two-sided electronic markets such as eBay, Craigslist, Airbnb, and Uber are typically interested in maximizing their individual utility but share the overarching goal of *transacting*. Similarly, providers of code, articles, and posts may have different motivations and goals than appropriators, yet they both share the overarching goal of exchanging the information; otherwise, they would not be participating in the information commons.

Despite common characteristics, an important difference exists between information commons on the one hand, and Hardin's commons and Ostrom's CPRs on the other. Hardin and Ostrom both addressed physical resource systems with tangible natural or man-made resource units. This materiality makes the resource units subtractable—implying that their consumption depletes the CPR (Ostrom 1990). In contrast, information commons pertain to resources in digitalized form, implying that the resources in information commons cannot be depleted through overconsumption. If anything, information commons become more sustainable through increased consumption because of network effects that bring about the *comedy of the commons* (Rose 1986) rather than the *tragedy of the commons* (Hardin 1968). This difference in resources is a given and, in any case, Hardin's notion of the commons focuses on the *collective decline* resulting from unchecked individual actions, rather than on overconsumption per se. As a result, we assert that our use of the term “commons” is appropriate because the types of information systems that we theorize emerge from the actions of free individuals and make the systems susceptible to various collective-action threats afflicting CPRs. We expand more on these threats to information commons in subsequent sections.

Information commons are information systems. Hence, to bound the phenomena of information commons, we emphasize the representation and token views of information (McKinney Jr et al. 2010). Information can be viewed as *representative* of something that exists independently in the world as an object, is reflected through symbols and signs,⁶ and becomes meaningful only after a brain (whether organic or computerized) processes it. An image or a video is simply mediums transferring static and dynamic visuals with or without auditory sounds—all of which are meaningless without a brain to process them. Similarly, words on the screen, numbers, and symbols (be they emojis or lines of code) are also just meaningless shapes outside a human context (language) that can attach significance to them. Subscribing to the representation view, we can see that, even though information commons differ vastly in the ways in which information is transferred within them, they are nonetheless information systems. We can also see that the process of encoding information into symbols and signs and the subsequent interpretation of them is inherently vulnerable. The “pure” form of information typically gets compromised to a certain degree because of the limitations of symbol systems and of the processing capabilities of humans and computers.

The *token* view is also worth considering here. This view asserts that information, while abstract in general, acquires a certain tangible property when encoded into symbols and signs via an information system in which it can be further molded, moved, stored, repackaged, retrieved, and distributed (McKinney jr et al. 2010). The token view suggests that information, despite not being a physical resource, nonetheless is “governable.” When providers upload, delete, edit, transfer, and publish content, they engage in acts of information governance (as opposed to system governance,

⁶ Examples include language symbols, such as words, numbers, icons, and codes, as well as sensory signs, such as visual images, audible sounds, odorous smells, feelable textures, and tasteable flavors (as of now, the latter three are irrelevant for information commons).

which is the producers' domain). Thus, when we posit that information commons are "self-rising," we do not mean that they pop into existence out of nowhere, but rather that the content and information in them emerge in an unorganized fashion from the actions of individual providers.

In Table 2.2, we list exemplary contemporary phenomena that we conceptualize as *information commons*. Because these phenomena and the related literature are still emerging, each phenomenon has multiple definitions and names that at times overlap. For instance, we list Wikipedia as a *wiki*, but it may also be viewed as an *online community* or *crowdsourcing platform*. In this case, we categorize Wikipedia as a *wiki* because of its unique characteristic of live-time editing, which does not exist in most online communities or crowdsourcing platforms. The phenomena in Table 2.2 differ in many respects, but they all share the defining characteristics of information commons—that is, they are highly accessible, self-rising information systems in which stakeholders share an overarching goal, even if some of their individual goals differ and possibly conflict within the overarching goal's framework.

Phenomenon	Definition	Examples	Shared Goal
Blogging	A website that lets anyone write and publish a blog	Blogger, Tumblr, Twitter	Self-expression and information sharing
Crowdsourcing	A platform in which the contributions of numerous self-selected volunteers or part-time workers combine with those of others to achieve a greater result	Innocentive, Quirky, Threadless	A specific, context-dependent goal
Media sharing	A website that lets people upload and share their video clips and/or images with the public at large or invited guests	YouTube, Vimeo, DailyMotion, Instagram, Flickr, Imgur	Creation of a repository of videos and images
Online community	A website where people congregate online to discuss a subject or to introduce themselves for possible in-person meetings	TES, Mumsnet, HackerNews	Creation of a dedicated information resource

Online reviews	A platform on which reviews can be posted about people, businesses, products, or services	Yelp, TripAdvisor, Amazon, IMDB	Self-expression and peer signaling on quality of products and services
Open source software development	A website that software developers can use to control and manage free and open-source software development	SourceForge, GitHub, BerliOS	Making software publically available to study, change, and distribute
Peer-to-peer file sharing	A platform that lets users access media files such as books, music, movies, and games	Pirate Bay, Torrentz, isoHunt	Sharing of digital content
Two-sided electronic markets	A platform that lets individuals post and access information for making online or offline transactions	eBay, Craigslist, Airbnb, Uber	Transacting
Social media	Online forms of communication that individuals and companies use to share information with interested parties (friends, colleagues, customers, etc.)	Facebook, MySpace, Friendster	Maintaining social ties
Wikis	A website that lets users collaboratively edit its content and structure	LyricWiki, WikiAnswers, wikiHow, Wikipedia	Amassing a large body of information

2.3 Derived Stakeholder Value

Realizing the shared goal of an information commons requires that the involved stakeholders receive value⁷ from it. Accordingly, we define derived stakeholder value as *the benefits stakeholders gain from being involved with an information commons*. Although a significant body of IS research has examined IT's value by focusing on tangible aspects of organizational performance, the value of information commons cannot be adequately gauged through such traditional measures. In information commons, derived value is context dependent, can be tangible or intangible, is realized at both the individual and collective levels, and is likely to be multifaceted. Prior research reveals that IT users can derive different types of value including satisfaction, self-expression, enjoyment, and economic value in different contexts (Agarwal et al. 2000; Kohli et al.

⁷ The word *value* can mean either "a person's principles" or "worth." Here, we refer to value as worth, which is how it is commonly used in research on IT gains (typically referred to as the "value of IT").

2008; Melone 1990); all of these facets of value can be relevant when examining derived stakeholder value in information commons. Drawing on these insights, we distinguish between three types of actors involved in information commons and propose the following constructs related to derived stakeholder value: (i) producer value, (ii) provider value, and (iii) appropriator value.

2.3.1 Derived Producer Value

We define producer value as *the benefits a producer gains from architecting and maintaining the infrastructure of an information commons*. Information commons support may require an ongoing investment, such as to pay the salaries of graphic designers, programmers, and other personnel. Information commons' producers may also carry other costs, such as for domain names, servers, cloud storage, search engine optimization, and securing private information. While some producers (such as Wikipedia producers) are motivated by intangible factors such as prestige, satisfaction, and genuine interest in the overarching goal of the information commons, others are profit-seekers and thus measure much of the value they derive in terms of financial gain. Producers of two-sided electronic markets may derive economic value by charging providers and appropriators brokerage fees, but most information commons are free to use and supported through advertising revenue. For that reason, producers' value is directly dependent on appropriators (the segment advertisers are trying to reach) and indirectly on providers, who attract appropriators by supplying them with content. Of course, other economic systems also depend on the efficient alignment of supply and demand. The difference, though, is that in information commons, producers have little control over provision.

Table 2.3 Conceptualizing Stakeholder Value

Concept	Detention	Examples
Derived Stakeholder Value	The benefits stakeholders gain from being involved with the information commons	Benefits vary based on stakeholder type (see below)
Constructs		
Producer Value	The benefits a producer gains from architecting and maintaining the infrastructure of the information commons	Personal satisfaction, pride, profit
Provider Value	The benefits a provider gains from contributing content to the information commons	Recognition, personal satisfaction and enjoyment, rewards
Appropriator Value	The benefits an appropriator gains from accessing the content of the information commons	Satisfaction of a defined need for information (such as content, entertainment, software, or knowledge) or an undefined need for information (discovery and exploration)

2.3.2 Derived Provider Value

We define provider value as *the benefits a provider gains from contributing content to an information commons*. In some information commons, providers can expect to receive tangible value (as in crowdsourcing competitions and two-sided electronic markets), while in others, the value that providers derive is intangible. Past work finds that the value here is often a combination of personal and social fulfillment (Faraj et al. 2011; Hann et al. 2013; Levina et al. 2014; Ma et al. 2007; Moon et al. 2008; Tsai et al. 2014; Zhang et al. 2013). Past work also finds that providers are sensitive to barriers that increase the opportunity cost of participation (Butler et al. 2014), as well as that the intensity of provision varies across time and from individual to individual (Gu et al. 2007).

2.3.3 Derived Appropriator Value

We define appropriator value as *the benefits an appropriator gains from accessing the content of an information commons*. Appropriators turn to information commons to satisfy a specific need for information or an indeterminate need to explore (Aggarwal et al. 2013; Browne et al. 2007; Heer et al. 2005). The need for information is personal for each appropriator and varies greatly across different information commons—ranging from a need to learn about transaction opportunities or the availability of free software to a need to know what members of one’s social circle are doing. Often, appropriators seek to satisfy the need to explore or simply to be entertained. Regardless of the context, the derived value from satisfying the need for information and discovery determines the likelihood of the appropriator returning to the information commons. Like providers, appropriators are also sensitive to barriers that raise their opportunity cost of information retrieval and exploration (Gu et al. 2007; Jones et al. 2004). Examples of opportunity cost-raising barriers include cumbersome registration mechanisms, interfaces that are difficult to navigate, cybersecurity issues, and other factors that reduce the seamlessness of appropriation.

2.4 Sustainability

For stakeholders to continuously derive value from an information commons, that commons must be sustainable. Sustainability, or “the capacity to endure” (Davidson 2014), has long been examined in environmental, social, and economic contexts (Malhotra et al. 2013) and has been increasingly examined in the context of online communities, virtual teams, online forums, and smartphone applications (Adar et al. 2000; Butler 2001; Butler et al. 2014; Cheung et al. 2002; Ridings et al. 2010; Teo et al. 2003). Drawing on themes from CPR literature and research on the sustainability of emerging phenomena related to decentralized online information systems, we consider *sustainability* as the universal purpose of all information commons. Thus, understanding

sustainability can help explain why an information commons fails or succeeds in realizing its overarching goal, as well as help predict a commons' trajectory. Accordingly, we define sustainability as *the capacity of an information commons to continuously provide value to its stakeholders*. Ultimately, any information commons seeks to realize this overarching goal, whether it is achieved through continuous activity, such as the ongoing cataloging of all human knowledge or ongoing maintenance of social relationships, or in a more temporary fashion that completes the overarching goal and countermands the need for sustainability, such as solving a technical problem in an online forum. Drawing on insights from CPR and IS research, we put forward four constructs of sustainability: (i) provision, (ii) appropriation, (iii) revitalization, and (iv) equitability.

Concept	Definition	Examples
Sustainability	The capacity of the information commons to continuously provide value to its stakeholders	The micro-blogging platform Pownce is an information commons that is no longer sustainable due to low provision, appropriation, and revitalization
Constructs		
Provision	The extent to which providers continuously input information into the information commons	Posting and uploading of digital content on YouTube, Facebook, eBay, Craigslist, and Digg
Appropriation	The extent to which appropriators continuously consume information from the information commons	Viewing, reading, listening to, and downloading digital content on YouTube, Facebook, eBay, Craigslist, and Digg
Revitalization	The rate of information provision between new and disengaged providers	The rate of digital content contributions between new and inactive providers on YouTube, Facebook, eBay, Craigslist, and Digg
Equitability	The extent to which provision activities are distributed across a base of providers	The extent to which a broad rather than narrow base of individuals provide information on YouTube, Facebook, eBay, Craigslist, and Digg

2.4.1 Provision

Without continuous provision of resource units, CPR systems (Ostrom 1990) and information commons such as online communities (Butler 2001; Butler et al. 2014) eventually dwindle.

Provision activities in information commons rely on self-selecting individuals who come from vastly different backgrounds (Hann et al. 2013; Shah 2006; Stewart et al. 2006) and who inevitably incur the opportunity cost of forfeiting their time and energy for no tangible return (Butler et al. 2014). As a result, provision of information is diverse and volatile (Daniel et al. 2013; Gu et al. 2007; Ransbotham et al. 2011). Understanding provision activities is thus important for predicting whether an information commons is growing, stabilizing, or declining.

2.4.2 Appropriation

Appropriation is a fundamental activity in information commons. In contrast to subtractable physical resource systems, which are strained as appropriation grows (for instance, ever-increasing hunting can cause extinction), information commons depend on continuous appropriation to reinforce network effects (Susarla et al. 2012). In addition, studies find that providers in information commons are often motivated by continuous appropriation (Goes et al. 2014; Huberman et al. 2009; Jabr et al. 2013; McKinney Jr et al. 2010; Moon et al. 2008; Singh et al. 2014; Zhang et al. 2013) and that information commons production is often supported by appropriation—through subscriptions, advertising, or brokerage fees. As such, understanding appropriation predicts the likelihood of an information commons to continue offering value.

2.4.3 Revitalization

Sustainability of natural resource systems is often gauged by the rate of resource unit replacement (Ostrom 1990). Although information does not decay⁸ and is not subject to physical depletion, the representation view implies that information is interdependent with the agency transferring it (McKinney Jr et al. 2010). Thus, attraction-selection-attrition theory (Butler et al. 2014) suggests

⁸ Information can become outdated, which can be considered a form of decay.

that the balance between newly engaged and withdrawing information providers (that is, the balance between those who enter and those who are no longer active) is an important predictor of an information commons' capacity to continue generating value for its stakeholders. Accordingly, we focus on *information revitalization*—that is, the difference between contributions of content from new providers and those who have become disengaged. A positive balance between provider entrance and exit (that is, a net gain in the number of providers) accompanied by an overall decrease in provision indicates that the new providers are not as productive as the providers that left the system. In contrast, a negative balance between the entrance and exit of providers (a net loss in the number of providers) accompanied by an overall increase in provision indicates that the burden of supplying the information commons with content is shared among fewer providers.

2.4.4 Equitability

Equitability among providers is another important information practice. Although complete equitability is unrealistic and impractical, as some providers are naturally better suited for certain information provision tasks than others (Kuk 2006), it is important to remember that provision in information commons is inherently unpredictable as providers can exit at any time without notice. Thus, the greater the dependency on a few information providers, the more vulnerable the information commons is to attrition. Political scientists and economists use concentration indices—such as the GINI coefficient—as a proxy measure of inequality that endangers the sustainability of economies and municipalities (Pulselli et al. 2006; Rodrik 1999). Such concentration indices also can be used to understand the distribution of information-provision activities; understanding equitability can serve to predict the information commons' future capacity to provide value to its stakeholders.

Sustainability is as important for information commons as it is for organisms, ecosystems, and other artificial resource systems. On that account, the information commons are not unique. On the other hand, as highly accessible, self-rising information systems, information commons operate under conditions of uncertainty that make them more vulnerable to collective-action threats than traditional information systems.

2.5 Collective-Action Threats

Many factors can adversely or positively impact the sustainability of information commons. Ostrom identifies six institutional factors that include no less than 43 subfactors for analyzing socio-ecological system sustainability (Ostrom 2009). Factors known to reduce information commons' capacity to provide value to their stakeholders include government censorship, insufficient internet infrastructure, competition from other information commons, poor design, information overload, and lack of capital (Christine Roy et al. 2001; Rochet et al. 2003; Rosen and Purinton 2004; Sullivan 2012). Factors that increase information commons' capacity to provide value include uniqueness, virtuous network effects, and provision volume (Butler et al. 2014; Ellison et al. 2007; Prahalad et al. 2013). To advance research into these factors and maintain the thematic consistency of our theory development effort, we focus on collective-action threats to sustainability. Building on the notion of the tragedy of the commons, we conceptualize a collective-action threat as an *adverse aggregate effect that is caused by individual-level actions*.

Every resource system is subject to collective-action threats caused by the inherent tension between community interests and the pursuit of individual gain (Ostrom 1990). Similarly, information commons' inherent properties make them highly vulnerable. In these commons, the unchecked freedom of entry, exit, and action, and the lack of centralized authority to ensure that

rules are followed and processes carefully executed amplifies the potential for opportunistic behavior that can lead to negative consequences, including loss of trust, widespread desertion, and even crackdown by external authorities. Synthesizing observations from the literature on CPR management (Ostrom 1990) and the commons (Hardin 1968), with insights from the diverse body of research on decentralized online information systems, we identify five collective-action threats: (i) free-riding, (ii) congestion, (iii) pollution, (iv) violation, and (v) rebellion (see Table 2.5). Our theorizing of the five phenomena as collective-action threats is inspired by Hess and Ostrom’s (2003) conceptualization of information as a CPR; works on “open commons,” such as roads (Benkler et al. 2013); Hardin’s tragedy of the commons logic, which focuses on individual-level actions’ adverse effects on collective outcomes in an environment with few restrictions on such actions; and examples from the literature on decentralized online information system.

Concept	Definition	Examples
Collective-action Threats	Adverse aggregate effects that are caused by individual-level actions	Individual-level consumption as a cause of global warming; individual-level tax evasion as a cause of budget shortages
Constructs		
Free-riding	The extent to which appropriators evade information provision	Consuming but not contributing to online reviews platforms, discussion boards, online support communities, P2P file sharing
Congestion	The extent to which appropriation or provision clogs the information commons	Slowdowns in video streaming, online gaming networks, edit wars on Wikipedia
Pollution	The extent to which information fails to be aligned with the information commons’ overarching goal	Electronic word-of-mouth manipulation, article vandalism, false rumors spread on social media
Violation	The extent to which providers break internal and external morals, rules, and laws	Copyright violation, cyber bullying, spread of hate speech, Illegal pornography
Rebellion	The extent to which providers exit because of dissatisfaction with producer actions	User revolts in Wikipedia, Reddit, eBay, Uber, Twitter, Facebook, YouTube

2.5.1 Free-riding

We define free-riding as *the extent to which appropriators evade information provision*. Extracting resource units from a resource system without contributing to it on a large, continuous scale might endanger that system's sustainability (Ostrom 1990). The vast majority of appropriators never become active providers, and while information commons differ in their tolerance for free-riding, the extent to which appropriators avoid acting as providers will inevitably impact the information commons' sustainability prospects. For instance, if two information commons occupy the same niche, yet free-riding is substantially higher in one than the other, then—all else being constant—we can reasonably predict that the latter is more likely to be sustainable in the long run. Information provision carries an opportunity cost that differs vastly across individuals (Butler et al. 2014). If people do not think that the benefit of information provision outweighs the cost of forfeiting their time and energy, they are more likely to free-ride. Smaller information commons, in which each additional provider has a relatively high marginal effect, are especially vulnerable to free-riding. Free-riding behavior is cited as a problem in OSS development (Baldwin et al. 2006), P2P file-sharing networks (Adar et al. 2000; Hosanagar et al. 2010; Johar et al. 2011; Karakaya et al. 2009; Rodrigues et al. 2010), discussion forums (Gu et al. 2007; Wasko et al. 2005), and reputation mechanisms (Liu et al. 2004).

Although free-riding may adversely impact an information commons, some important caveats must be considered. Prior research finds that free-riding might be desirable in situations where certain providers are better suited to performing a specific task (Kuk 2006). In addition, an information commons with no free-riding would likely cause information overload and reduce the information value. Finally, evasion of information provision does not necessarily imply free-riding, as appropriators can help the information commons in other ways. For instance, some

appropriators might not contribute information, but may act as “ambassadors” and attract others to join (Susarla et al. 2012) or may engage in mundane but necessary tasks, such as editing and organizing content in wikis (Beck et al. 2015). When appropriators make contributions other than direct provisioning of information, they are playing a constructive role toward sustaining the commons and therefore should not be viewed as free-riders.

2.5.2 Congestion

We define congestion as *the extent to which appropriation or provision clogs an information commons*. CPRs are vulnerable to spikes in appropriation that, in extreme cases, can render the entire system highly ineffective. Unlike CPRs, information commons can experience congestion both in appropriation and in provision. Although information is a non-subtractable resource, the digital infrastructure that supports its provision and appropriation is a subtractable resource (Johar et al. 2011; Listanti et al. 2000). For instance, when Facebook was initially expanding its market penetration, it did so in a deliberately gradual manner to make sure it had enough server capacity to support growing demand to avoid congestion in appropriation (“How Facebook Became,” 2016). Other examples of information commons that are most susceptible to congestion include video-streaming platforms (Plissonneau et al. 2012; Setton et al. 2005; Wu et al. 2001) and P2P file-sharing networks (Johar et al. 2011). Appropriators might get frustrated with and even lose trust in information commons that suffer from frequent congestion. Loss of trust inevitably leads to a certain level of desertion that, if large scale, can undermine a commons. Another type of congestion problem that may frustrate appropriators is information overload caused by rapid provision (Rosen and Purinton 2004). A rapidly changing Facebook feed that makes it difficult for appropriators to keep up is an example of information overload caused by congestion in provision. Congestion in provision is a rather unique problem; it occurs because of the distinctive freedom of

access and action that information commons afford (similar to the problem academics experience as they try to get published). Congestion in provision is especially problematic in the most accessible information commons, which are open to immediate simultaneous provision from multiple providers (Aaltonen et al. 2015; Kittur et al. 2008).

At present, most information commons do not suffer from serious appropriation congestion. However, it is not yet clear what will happen if appropriation increases as projected and broadband infrastructure fails to keep pace (Bolcskel et al. 2001). Similarly, internet service providers' (ISPs) objective to undermine Net Neutrality can potentially increase congestion in appropriation for some information commons (Hahn et al. 2006).

2.5.3 Pollution

We define pollution as *the extent to which information fails to be aligned with an information commons' overarching goal*. Not unlike natural resources, information is subjected to contamination (data contamination in research is one example of pollution). According to the representation view of information, in its "pure" form, high-quality information should be true in relation to its source object, comprehensible, subjectively sincere, and socially legitimate (Habermas 1985). However, this high standard is rarely achieved, and all information is potentially subject to misrepresentation. Unintentional, occasional pollution from low-quality provision raises the search cost for appropriators (Gu et al. 2007) and decreases the information's overall value to appropriators (Rice 2012). Deliberate pollution, in the form of information manipulation and vandalism, can also undermine the trust in a commons and reduce appropriation. Information manipulation is cited as a problem in online review platforms (Anderson et al. 2014; Dellarocas 2005; Hu et al. 2012; Mayzlin et al. 2012), wikis (Kittur et al. 2008; Shachaf et al. 2010; Stvilia et

al. 2007), blogs (Schmierbach et al. 2012; Thorson et al. 2010), and two-sided electronic markets and online communities (Gibbs et al. 2010; Gu et al. 2007; Toma et al. 2008).

2.5.4 Violation

We define violation as *the extent to which providers break internal and external morals, rules, and laws*. Crime and the violation of social rules and norms demonstrate how the actions of relatively few individuals can cause substantial societal damage. This violation strains public resources, reduces productivity, and, in extreme situations, disintegrates entire communities (Covington et al. 1991). In information commons, high accessibility inevitably increases the potential for violation, as opportunistic actors can blend in and exploit the freedom of movement with relative ease. The literature is full of examples of violation activities in information commons: stalking, bullying, fraud, proliferation of hate speech, distribution of illegal pornography, coordination of hacking activities, coordination of terrorism, identity theft, intellectual property theft, and more (Beale et al. 2007; Gerstenfeld et al. 2003; Gross et al. 2005; Hinduja et al. 2010; Juvonen et al. 2008; Mishna et al. 2009; Svensson et al. 2012). In addition to violating social rules and norms, when widespread, such violations threaten information commons' sustainability in two distinct ways: (i) they cause information providers and appropriators to leave, thus reducing appropriation, provision, and, subsequently, revitalization and equitability; (ii) they prompt a challenge from powerful actors, such as ISPs or criminal or civil litigation systems. Depending on the nature of the offense, the impact of outside challenges on information commons might be small or large, ranging from a virtual slap on the hand to a fine, or, in some instances, a complete shutdown.

2.5.5 Rebellion

We define rebellion as *the extent to which providers exit because of dissatisfaction with producer actions*. Although most people do not actively participate in them, revolutions drive history. As such, they are in effect a collective-action phenomenon, where the actions of a relatively small number of individuals have an impact that extends beyond the local level (Muller et al. 1986). To date, the academic literature does not pay much attention to online rebellions, but that does not mean they do not occur. Anecdotal evidence suggests that many prominent information commons experienced some sort of user revolt at one time or another. Over the years, the mainstream media and the blogosphere have reported instances of rebellion in different information commons, including eBay, Digg, Wikipedia, YouTube, Twitter, Facebook, Instagram, Uber, and Reddit (Auerbach 2015; Clifford 2010; Feuer 2016; Graham-Felsen 2006; Gross 2012; Johnson 2009; Nizza 2007; Shih 2013; Tkacz 2011). While in-depth research on user rebellion is lacking, we can identify three recurring triggers: (i) unwelcomed changes to infrastructure features, (ii) unwelcomed changes to policies, and (iii) producers' heavy-handedness in dealings with core providers.

As with free-riding, congestion, pollution, and violation, rebellion is a potential issue in any situation involving people. In the information commons, however, these five basic threats are amplified by high accessibility. Given this, we now examine how information commons can remain sustainable in the face of such threats and thus continue to benefit their stakeholders.

2.6 Polycentric Governance Practices

To reduce the provenance and impact of collective-action threats and promote sustainability, traditional information systems rely on hierarchical structures. Such structures are characterized

by relative stability, compartmentalization, specialization, formal authority and procedures, and the use of behavioral and outcome controls. In contrast, systems characterized by open access are more likely to achieve long-term stability through polycentric governance practices that balance local-level autonomy and inclusivity with local-level boundary-setting and accountability mechanisms (Ostrom 1990). Drawing on these distinctions, we propose that *the extent to which an information commons integrates polycentric governance into its design determines how likely it is to fend off collective-action threats and thereby promote sustainability*.

The concept of polycentricity was originally developed by Polanyi (1951) to describe the continuous state of flux in science. Unconstrained by an intervening central authority, scientists, according to Polanyi, are free to exercise independent, original thought and engage in experimentation (Aligica et al. 2012). A decade later, the concept of polycentricity was adapted as an alternative to the movement toward greater centralization and consolidation of public services administrations in US metropolitan areas (Ostrom et al. 1961). The notion that multiple crosscutting and autonomous jurisdictions are better positioned to administer public services than a single centralized authority stood in stark contrast to the prevailing wisdom of political scientists and policymakers at the time. Defined as an arrangement in which independent elements mutually adjust to create orderly relationships within a larger system of rules (Ostrom 1972), polycentricity is characterized by self-emerging spontaneity, self-governing independence, and flexibility arising from resilience to experimentation with rules. Researchers have found polycentricity to be an effective alternative to top-down centralization in municipal governance (Ostrom 1972) and, subsequently, to characterize sustainable CPR systems (Ostrom 1990). People in resilient CPR systems are generally more engaged in the system's governance, actively participating in decision-making about adjusting rules and working together to monitor those who fail to abide by rules

agreed to by the majority of people. Ostrom (1990) identified eight principles underlying most flourishing CPRs that were absent in failing systems (see Table 2.6).

Principle	Explanation
1. Clearly defined boundaries	Individuals or households who have rights to withdraw units from the CPR must be clearly defined, as must the boundaries of the CPR itself.
2. Congruence between appropriation and provision rules and local conditions	Appropriation rules restricting time, place, technology, and/or quantity of resource units are related to local conditions and to provision rules requiring labor, material, and/or money.
3. Collective choice arrangements	Most individuals affected by the operational rules can participate in modifying the operational rules.
4. Monitoring	Monitors, who actively audit CPR conditions and appropriator behavior, are accountable to the appropriators or are the appropriators.
5. Graduated sanctions	Appropriators who violate operational rules are likely to be assessed graduated sanctions (depending on the seriousness and context of the offense) by other appropriators, by officials accountable to these appropriators, or by both.
6. Conflict-resolution mechanism	Appropriators and their officials have rapid access to low-cost local arenas to resolve conflicts among appropriators or between appropriators and officials.
7. Minimal recognition of rights to organize	The rights of appropriators to devise their own institutions are not challenged by external governmental authorities.
8. For CPRs that are parts of larger systems: Nested enterprises	Appropriation, provision, monitoring, enforcement, conflict resolution, and governance activities are organized in multiple layers of nested enterprises.

We draw on Ostrom's eight governance principles of successful CPR systems and on insights from empirical research on decentralized online information systems to synthesize polycentric governance practices applicable to the information commons context. To balance faithfulness to the reference literature and our theorizing, we examined points of overlap between Ostrom's principles with the different IS context, while avoiding artificial retrofitting as Whetten et al. (2009) recommend. To make the theoretical framework useful—and not overbearing—we consolidated some principles to arrive at four general polycentric governance practice constructs:

(i) boundary regulation, (ii) incremental adaptation, (iii) shared accountability, and (iv) provider recognition.

Concept	Definition	Examples
Polycentric Governance Practices	Practices that promote order, where independent elements make mutual adjustments to order relationships with one another within a general system of rules	In the physical world: local overlapping municipalities, the European Union, communes; Online: Wikipedia, eBay, Pirate Bay, YouTube
Constructs		
Boundary Regulation	The extent to which rules and technical infrastructure features afford information provision and appropriation consistent with the information commons' overarching goal	Restricting posting of content that does not meet the information commons' overarching goal; requiring providers to register
Incremental Adaptation	The extent to which changes in infrastructure and rules are gradually introduced and providers and appropriators are actively involved in shaping them	Gradually updating technical features; small-scale experimentation with new features and rules; seeking feedback from users on new rules and features
Shared Accountability	The extent to which rules and features afford peer monitoring and gradual sanctioning to support appropriate behavior and dispute resolution in an information commons	Peer monitoring mechanisms, such as mutual ratings and rankings, helpfulness cores, and flagging rule violators; gradual punishment of rule violators, from a warning to banning
Provider Recognition	The extent to which providers are acknowledged by peers, appropriators, and producers	Subscriptions; followers; digital status symbols, such as icons and avatars; "likes"; direct feedback

Polycentric governance is an abstract notion that manifests in an information commons' rules and technical infrastructure features. Rules underlie the governance of social activities by playing an integral part in ordering relationships, responsibilities, and expectations (Ostrom 1972). CPR research identifies three levels of rules—*constitutional*, *collective choice*, and *operational* (Ostrom 1990)—that are also present in information commons. Terms of use by an information commons' providers and appropriators exemplify constitutional rules. Policies that change in response to

petitioning or protest of providers and appropriators exemplify collective choice rules (for example, Facebook relaxed its “real name” policy in response to user protests). Operational rules are exemplified by an information commons’ day-to-day implicit and explicit norms of conduct, such as reciprocity (Mathwick 2002), sharing (Sharratt et al. 2003), trust (Ridings et al. 2002), and the language used in communications (Wilson et al. 2002). Without a mechanism to help enforce them, rules are no more than general directives or recommendations. Therefore, constitutional, collective choice, and operational rules are often reflected in an information commons’ infrastructure features, including the graphical interface and mechanisms for access (such as login), provision, appropriation, self-identity representation, communication, and peer monitoring (Bartelt et al. 2014; Jabr et al. 2013; Kane et al. 2012; Levina et al. 2014; Ma et al. 2007). Next, we detail the four practices of polycentric governance.

2.6.1 Boundary Regulation

Although CPR systems need boundaries primarily to control appropriation and avoid overuse (Ostrom 1990) (Principle 1, Table 2.6), information commons need boundaries to control provision, including the type of information allowed, provider conduct, and the degree to which providers can mask their identities. Typically, producers set these boundaries when they design the information commons, expressing the boundaries both in the commons’ constitutional rules and in its infrastructure features. However, drawing on Ostrom’s observations, we assert that provider engagement here is important: the more involved providers are in establishing boundaries, the more effective those boundaries are in regulating the information commons. The literature finds that setting effective boundaries based on overarching goals and other contextual characteristics of online communities and OSS projects is positively associated with outcomes (Bonaccorsi et al. 2003; Butler et al. 2012; Di Tullio et al. 2013; Hertel et al. 2003; Jarvenpaa et

al. 2011; Ren et al. 2007). Twitter's 140-character tweet restriction is an example of the alignment of Twitter's constitutional rules and technical features with its overarching goal of being a microblogging platform. In addition to regulating types of content, information commons must regulate anonymity, which is increasingly found to be associated with various types of illicit behavior (Christopherson 2007; Coffey et al. 2004; Suler et al. 1998). Extant research suggests that the tolerable degree of anonymity depends on the commons' overarching goal (Ren et al. 2007). Drawing on polycentricity theory and empirical findings, we assert that involving providers in decisions on boundary setting such as these increases the likelihood of sustainability.

2.6.2 Incremental Adaptation

A central notion in polycentricity theory is the idea that incremental, bottom-up experimentation with rules will likely lead to the discovery of better rules for governance (Ostrom et al. 1962) (Principles 2 and 3, Table 2.6). Because each information commons is different in its overarching goal, governance rules and infrastructure features must be specific to its evolving circumstances. It is particularly important that providers and appropriators—who experience these dynamics close up—influence decisions about rules and features. As the literature finds, autonomy of provision spurs creativity (Bishop 2007; Hertel et al. 2003; Lee et al. 2012; Roberts et al. 2006), and such creativity should not be restricted to content provision alone; it should also be leveraged to incrementally adapt an information commons' rules and infrastructure features to support its overarching goal. Although we still lack substantial empirical evidence on the impact of provider involvement in adjusting the rules and infrastructure features of an information commons, past IS literature on the benefits of user involvement in systems design (Ives et al. 1984; Kujala 2003) and recent literature on the sustainability of startups (Blank 2013; Ries 2011) suggest that open communication and feedback between producers and participants can increase the speed and

quality of adaptations to rules and infrastructure features.

2.6.3 Shared Accountability

Effective, low-cost, local-level conflict-resolution mechanisms that gradually enforce sanctions on rule violators, together with peer monitoring, are associated with CPR sustainability (Ostrom 1990) (Principles 4, 5 and 6, Table 2.6). In many cases, community members can effectively identify perpetrators long before outside police authorities get involved (Bennett et al. 2009). Thus, a system of accountability arising bottom-up from within the community can be a robust alternative to top-down policing. Similarly, research finds peer monitoring in information commons to be an effective governance mechanism (Chua et al. 2007; Feller et al. 2008; Gu et al. 2007; Wall et al. 2007). As long as the offense is not severe, sanctioning of rule violators should be carried out in a gradual manner (Ostrom 1990). In information commons, this gradual sanctioning typically starts with a warning, then escalates step-by-step to temporary—and in rare cases, permanent—blocking of access. Drawing on polycentricity theory and empirical findings, we assert that, in an information commons, rules and infrastructure features that enable shared accountability among community members increase the likelihood of sustainability.

2.6.4 Provider Recognition

In information commons, provision activities typically carry no monetary payoff and are instead motivated by an array of personal and social factors. Although producers cannot easily impact the providers' personal motivation, they can enhance the social experience by incorporating various peer-recognition features. For some time now, IS researchers have recognized that providers are “strongly driven by status and status seeking, and that status sentiments are more likely to sustain virtual communities” (Lampel et al. 2007). As such, research finds that recognition—whether from peers, appropriators, or producers—is a major driver of continuous provider participation. Hence,

direct feedback, rating schemes, helpfulness scores, “likes,” number of followers or subscribers, and digital status symbols such as icons and badges have a lasting positive impact on provision (Goes et al. 2014; Jabr et al. 2013; Levina et al. 2014; Moon et al. 2008; Ren et al. 2007; Von Krogh et al. 2012). Despite differences in overarching goals and contexts across information commons, it is safe to assume that rules and infrastructure features that facilitate recognition increase provider satisfaction, which in turn increases the likelihood that providers will remain involved in the information commons.

2.7 Theorizing Relationships

Our discussions thus far have focused on the *terms*—that is, the concepts and their elaboration into constituent constructs—of the theory of PIC (see Tables 2.1, 2.3–2.5, and 2.7). We now build on these terms to theorize the *relationships* between concepts and their constituent constructs⁹ as summarized in the conceptual model in Figure 2.1.

We draw on the logic of profile constructs and ideal profiles (Doty and Glick 1994; Venkatraman 1989) and on the literature on decentralized online information systems to conceptualize the relationship between sustainability and its underlying constructs—that is, continuity of provision, continuity of appropriation, revitalization, and equitability. We suggest that these four constituent constructs contribute in *non-substitutive and mutually reinforcing ways* to achieving the sustainability necessary for providing continuous value to stakeholders (whether providers, appropriators, or producers). Accordingly, we suggest that an ideal sustainability profile for an

⁹ We follow the notion of “ladder of abstraction” (Van de Ven, 2007, p. 115) in differentiating between concepts and constructs. Concepts are semantically defined but cannot be operationalized, while constructs are mid-level abstractions and may be operationalized.

information commons is one with high values for each of its constituent constructs, as explained in the following.

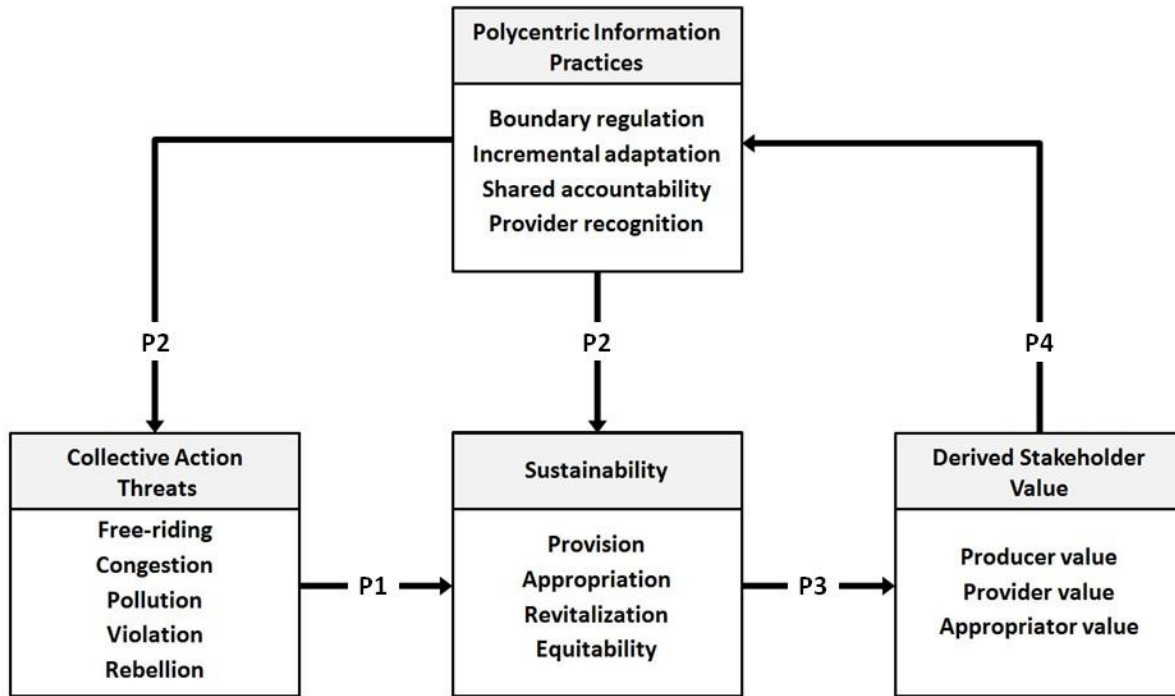


Figure 2.1 Conceptual Model

Continuous appropriation and provision are fundamental forces that feed each other (Huberman et al. 2009; Jabr et al. 2013; Moon et al. 2008; Zhang et al. 2013) in the same way that supply and demand are fundamental to all economic activities involving resource exchange. The need to revitalize provision to sustain the information commons might not be as immediately evident, as it depends on the provider attrition rate (Butler et al. 2014). However, this revitalization is as essential to determining the sustainability prospects of an information commons as the resource replacement rate is to determining the sustainability of natural and economic systems (Ostrom 1990). Providers in information commons typically enter and exit the commons at a rapid rate, which, when positive, generally injects new energy and content (Ransbotham et al. 2011). As long as revitalization in provision remains flat or is positive, the information commons is likely to

remain sustainable. However, if the rate between exit of old providers and entry of new providers is negative for a prolonged period of time, the likelihood of the commons remaining sustainable will decrease. Because provision carries an opportunity cost, the smaller the provider pool, the greater the relative burden on the remaining providers; this can cause some of them to leave, further aggravating the problem. Revitalization goes hand in hand with equitability. The need for equitable distribution of provision might not be immediately evident; after all, a few providers can often sustain a software development project or a discussion forum for some time. Further, equitability is rarely perfect—some providers are, by nature, more likely to be more involved than others. Still, we assert that the efforts of a few providers cannot sustain most information commons for an extended period of time because it increases the systemic vulnerability to provider dropout. Thus, low revitalization or low equitability is likely to lead to reduced provision which, in turn, is likely to reduce appropriation. Hence, we make the following core premise in our theorizing:

Premise: *Sustainability of an information commons requires that continuous provision and appropriation as well as revitalization and equitability in provision operate as a system of complements.*

We draw on the nature of collective action in relation to individual action (Morgeson and Hofmann 1999), the tragedy of the commons (Hardin 1968), and the literature on decentralized online information systems to theorize how collective-action threats that impact information commons' sustainability emerge from individual actions. We suggest that information commons' inherent openness creates conditions for individuals to engage in free-riding, congestion, pollution, violation, and rebellion. These actions are motivated by individual reasons that may be at odds with the common good. While less-inclusive information systems can rely on hierarchical mechanisms to control such actions by individuals, providers in information commons are largely unaccountable to the system producers. An information commons' producer cannot threaten to fire

underperforming providers or incentivize them with pay increases and promotions. For good and for bad, an information commons is a virtual space in which people—who often feel anonymous (regardless of how anonymous they really are)—upload, post, publish, broadcast, and code content and information freely. Just as multiple small individual actions can create remarkable aggregate outcomes (Benkler et al. 2006; Tapscott et al. 2008), self-interested actions by individuals can propagate to create adverse aggregate effects, including large-scale free-riding, congestion, pollution, violation, and rebellion that threaten the long term sustainability of an information commons. Even if a small proportion of providers and appropriators mistakenly or deliberately violate the rules—whether those rules pertain specifically to the commons or violate the law—the aggregate effect becomes larger than the sum of individual parts in commons that involve hundreds, thousands, or even millions of actors. Because the amount of content provisioned and appropriated and the number of providers and appropriators differs considerably across information commons, the threshold at which self-interested, contrary-to-common-good individual actions emerge as collective-action threats is not clear-cut. However, given that open access is characteristic of information commons, the potential exists for collective-action threats to emerge from such individual actions.

Research shows that collective-action threats may adversely impact provision, appropriation, revitalization, and equitability in information commons. For instance, pollution from low-quality provision raises the search cost for appropriators (Gu et al. 2007) and decreases their capacity to derive value (Rice 2012); this can lead to reduced appropriation. As another example, when free-riding increases, revitalization and equitability are likely to decline (Adar et al. 2000; Butler et al. 2014; Johar et al. 2011; Karakaya et al. 2009). Past research also offers evidence of the adverse impact of congestion and violation on provision and appropriation activities (Beale et al. 2007;

Finn 2004; Johar et al. 2011; Setton et al. 2005; Wu et al. 2001). At this stage, we could not identify research on the negative impact of rebellion, but substantial anecdotal evidence exists on the toll user revolts take on provision and appropriation (Auerbach 2015; Clifford 2010; Feuer 2016; Graham-Felsen 2006; Gross 2012; Johnson 2009; Nizza 2007; Shih 2013; Tkacz 2011). Hence, we propose:

Proposition 1: *Collective-action threats, which emerge from the aggregation of the free-riding, congestion, pollution, violation, and rebellion actions of individual providers and appropriators, create vulnerabilities for an information commons' sustainability.*

Next, we draw on Ostrom's observations on the association between the eight governance principles (Table 2.6) and CPR resilience (Ostrom 1990) and the literature on decentralized online information systems to theorize that boundary regulation, incremental adaption, shared accountability, and provider recognition constitute a set of polycentric governance practices that reduce collective-action threats in information commons while increasing their sustainability. For instance, achieving a compromise between providers and producers on content boundaries is likely to reduce rebellion, which can improve the sustainability of the information commons (Jarvenpaa et al. 2011). Establishing and enforcing boundaries can reduce pollution and help focus provision activities for an information commons that has close-ended overarching goals, such as in OSS (Bonaccorsi et al. 2003; Hertel et al. 2003). Indeed, failure to set boundaries on the type of information allowed into an online community can lead to rebellion and the subsequent exit of some providers (Butler et al. 2012; Ren et al. 2007). Further, incremental adaptation of rules and infrastructure features can reduce instances of pollution, violation, and congestion in provision (Aaltonen et al. 2015; Dellarocas 2005). For instance, Wikipedia's adoption of the three-revision rule helped mitigate pollution and congestion threats in provision (Aaltonen et al. 2015). As part

of the infrastructure, algorithms that incorporate feedback from other providers (peer monitoring and shared accountability) can substantially reduce pollution and violation in P2P file-sharing networks (Kamvar et al. 2003). When producers unilaterally initiate major changes to the infrastructure's graphic design and technical features or change the content or anonymity boundaries, a significant proportion of providers are likely to feel dissatisfied and leave (e.g., Gross 2012; Johnson 2009; Nizza 2007). Incremental adaptation that takes stakeholder feedback into account is likely to reduce dissatisfaction and rebellion (Halfaker et al. 2011). Shared accountability—as manifested through ratings, flagging, and “report abuse” and “report problem” features—can also reduce violation and pollution (Diakopoulos et al. 2011; Jøsang et al. 2007; Resnick et al. 2000). Identity features that act as recognition features and reduce anonymity have a positive effect on provision and appropriation (Forman et al. 2008; Ma et al. 2007). Recognition from peers, appropriators, or producers has a lasting effect on sustainability outcomes as it motivates continuous provision (Goes et al. 2014; Levina et al. 2014; Ma et al. 2007) and helps entice appropriators to join the ranks of providers, subsequently reducing free-riding and increasing revitalization. Hence, we propose that:

Proposition 2: *Polycentric governance practices of boundary regulation, shared accountability, incremental adaptation, and provider recognition (a) reduce collective-action threats and (b) increase the sustainability of an information commons.*

Finally, we draw on literature from CPR governance (Ostrom 1990) and decentralized online information systems to suggest that the sustainability of an information commons creates the conditions for deriving stakeholder value, which in turn impacts the evolution of the commons' polycentric governance practices. Provision of content and information allows appropriators to derive value by fulfilling their need for information (Aggarwal et al. 2013; Browne et al. 2007;

Heer et al. 2005). It also allows producers to derive economic value from having a volunteer “workforce” provide the bulk of the content (Ramaswamy et al. 2010). Appropriation allows producers to derive value from selling ads, while also motivating providers who, in part, derive value from feeling needed (Burke et al. 2009). Revitalization and equitability help an information commons maintain the flow of provision over time and allow appropriators and producers to continue to derive value. In contrast, when sustainability decreases, the stakeholders’ capacity to derive value also decreases, creating a spiral effect that endangers the long-term sustainability of the information commons.

The need to change the information commons’ polycentric governance practices relates inversely to the level of derived stakeholder value. On the one hand, when all stakeholders derive value, producers do not need to make urgent changes to the information commons’ rules and infrastructure features supporting boundary regulation, incremental adaptation, shared accountability, and provider recognition. In such cases, producers have the opportunity to experiment with incremental modifications to enhance the overall experience for providers and appropriators. On the other hand, the urgency to change polycentric information practices increases when the value derived diminishes for one or more stakeholder groups. The importance of receiving and adapting to stakeholder feedback is recognized as crucial for the sustainability of organizations (Anderson 1999; Ries 2011); we suggest it is similarly important for information commons. As the creators of the information commons, producers ultimately control changes to the system rules and infrastructure features. Thus, they should create the conditions for monitoring the extent to which appropriators and providers derive value from participation to preserve their capacity to derive value. For instance, Angie’s List, a pioneer of online consumer peer reviews, made major changes to its boundary regulation on access in an attempt to reverse a long-running

trend of decline in its provision and appropriation (Tuttle 2016). It took Angie's List a long time to realize that when providers and appropriators do not derive value, they go somewhere else (to yelp.com, for instance) and might never return. Had Angie's List included an effective feedback process on the value derived by its providers and appropriators, its producers might have been able to act more urgently to stop its decline by adjusting its access boundaries. Hence, we propose:

Proposition 3: *A sustainable information commons allows stakeholders to continuously derive value.*

Proposition 4: *Feedback on value derived by stakeholders impacts the evolution of an information commons' polycentric governance practices.*

2.8 Discussion

While the openness and decentralization of PIC can spur incredible outcomes (Benkler et al. 2006; Tapscott et al. 2008), they also inevitably create systemic vulnerabilities that cannot be fully curtailed by traditional means of hierarchical command and control. Information governance that relies on voluntarily participation is a major paradigm shift—one that challenges previous assumptions of organized production and its underlying mechanisms (Baldwin et al. 2011; Benkler et al. 2006; Füller et al. 2006). As the dust of the first wave of mostly empirical research is settling and researchers provide evidence of the effectiveness of crowdsourcing, open source, and other forms of peer production (Benkler et al. 2006; Tapscott et al. 2008), the need for new theories capable of explaining the observed outcomes becomes more pronounced (Arazy et al. 2011; Johnson et al. 2014; Majchrzak 2009; Von Krogh et al. 2012). Most online forums, communities, and OSS projects are, in fact, inactive (Butler et al. 2012; Ren et al. 2007; Stewart et al. 2006; Wiertz et al. 2007), suggesting that the majority of initiatives become unsustainable following a phase of initial growth. In our theorization, we therefore focused on the common denominator—

the existential question of sustainability—to develop the theory of PIC by iterating between insights from literature on CPR governance and from decentralized online information systems.

Building on the notions of the tragedy of the commons and polycentricity, we focused on identifying the sustainability threats that impact information commons with few restrictions on participation. The aim of our theoretical framework is to *explain* outcomes related to sustainability and governance in information commons as well as to *predict* (Gregor 2006) the likelihood of achieving sustainability and the subsequently derived stakeholder value. As such, our merging of insights from CPR governance literature and empirical evidence from a diverse body of decentralized online information systems research under a single theoretical platform contributes to existing literature by (i) unifying multiple related empirical insights under one overarching framework; (ii) conceptualizing decentralized online information systems as *information commons*, which, in turn, helps us better understand questions related to their sustainability in the face of collective-action threats; and (iii) explicating the polycentric governance practices that reduce the threats' prevalence and mitigate their impact.

We are not the only ones to use the CPR governance literature to study intangible resources. Previous research has convincingly employed this framing to examine governance of radio spectrums, budgets, and even academic databases (Brubaker 1997; Evans 2005; Hess et al. 2003; Kranich et al. 2008; Soroos 1982). From these efforts, we have learned that it is implausible and inappropriate to directly translate all CPR and commons research to advance IS theory. Instead, we have endeavored to achieve a balance between remaining faithful to the reference literature (Whetten et al. 2009), our own “creative imagination” (Bacharach 1989. pp, 498), and the body of literature on decentralized online information system to which we aim to contribute. Hence, although digital information is vastly different from physical resources—especially when it comes

to subtractability and cost of distribution—information commons are nonetheless susceptible to sustainability threats that are traditionally more associated with physical systems, including free-riding, pollution, congestion, violation, and rebellion.

Besides being used to study intangible resources, the CPR governance and collective-action literature has also been sporadically employed to frame IS studies, including knowledge exchange in electronic communities of practice (Wasko et al. 2004), governance of Wikipedia (Forte et al. 2009), congestion in P2P networks (Johar et al. 2011), governance of OSS development (Markus 2007; O'Mahony 2003), and adoption and growth of regional health information infrastructures (Constantinides et al. 2014). What sets our work apart from these previous studies is our ambition to create a new theoretical platform for studying a diverse body of decentralized online information systems.

Thus, we position the theory of PIC beside other emerging theories aimed at explaining and predicting outcomes in decentralized online information systems, such as the theory of social dynamics in online fields (Levina et al. 2014). While the latter builds on Bourdieu's notion of social fields and provides a theoretical framing for studying power relations across various user-generated content platforms, our theory builds on Ostrom's work and focuses on governance and sustainability. Both theories draw inspiration from well-established reference literature to tailor new, context-specific frameworks for advancing our understanding of emerging online technologies from different, yet complementary, vantage points. We also see the theory of PIC as a theoretical expansion of previous empirical and theoretical work of IS scholars who examined the sustainability of online communities; their work provides important insights on the role of community size, communication patterns, member attrition, and the opportunity cost of participation (Butler 2001; Butler et al. 2014; Butler et al. 2012).

The theoretical framework we present here opens new avenues of research and further theoretical expansion. Future research can look more deeply at the role of rules and infrastructure features in information commons governance. The web design literature provides a good starting point for identifying and classifying the different features (Andrews 2002; Zhang et al. 2000). We need to better understand how infrastructure features reflect governance rules and how they facilitate a balance between autonomy and restrictiveness. Theoretical and empirical work that takes the next step, opening the black box of infrastructure components to see how they reflect governance practices that reduce sustainability threats' prevalence and impact, can help us better understand the technical aspects of information commons.

Future research might also more deeply examine the nuanced differences between types of actors in an information commons. Although stakeholders share a certain overarching interest in the context of an information commons, producers, providers, and appropriators are by no means homogenous. Some researchers already distinguish between actors according to their engagement levels, splitting providers and appropriators into "core" and "peripheral" actors (Gu et al. 2007; Scacchi 2004; Setia et al. 2012). Those distinctions are important for understanding dynamics in PIC; for instance, engaging all providers in decision-making about changes is much harder than engaging only the core providers, who are also more likely to provide valuable feedback. Differentiating types of appropriators can also enhance our understanding of sustainability. For example, a distinction can be made between casual and intensive appropriation, as well as between patterns of passive content appropriation and active ambassadorship that promotes social contagion effects (Susarla et al. 2012) that increase an information commons' sustainability through enhanced revitalization. Similarly, differentiating between types of producers can help us better understand their role in governance. Many information commons are produced by profit-

seeking owners who, besides ensuring the system's sustainability in the face of collective-action threats, work to establish a profitable business model to achieve financial sustainability. The tension between these two goals should be further examined. For instance, appropriators and providers in an information commons might perceive advertisements as pollution and personal data mining as a violation. In addition to expanding our understanding of the nuances in stakeholder groups, we also suggest that future research examines the important differences within each stakeholder category based on their network characteristics. Examining the co-evolution of the network structure of stakeholders with the governance of an information commons can provide insights into how network characteristics constrain or enable the emergence of polycentric governance and the continuous derivation of value by producers, providers, and appropriators.

Finally, our conceptualization of information commons may also be applied to study different types of organizational systems, including open-access enterprise-sponsored systems, crowdsourcing initiatives created for soliciting ideas or specific tasks from information providers (such as MyStarbucksIdea), support forums created by organizations to allow users to help each other, and enterprise-based social networks that are generally highly open for participation and likely to encounter collective-action threats (which might reduce their sustainability and potential for providing value) and thus can benefit from integrating polycentric governance practices.

The main objective of the PIC theory is to help frame future inquiry in a broad research field with too few theoretical alternatives. The contextual differences between various information commons, and the inevitable differences between study objectives, necessitate the operationalization of measures specific to each research; pollution in P2P file-sharing, for example, manifests differently than pollution in user-generated online review systems. Given this, the different concepts and constructs we developed and linked should be adapted in future

qualitative and quantitative studies aimed at explaining and predicting (Gregor 2006) sustainability and subsequently derived stakeholder value in information commons. Our generative lexicon of concepts can be used for framing both process and variance studies (Van de Ven 2007). Process studies might examine how policies and infrastructure features that reflect polycentric governance practices are adjusted over time. Process studies can also enhance our understanding of different information commons' life cycles—from their inception to maturity—through either a positivist prism, focusing on measuring variables, or through an interpretive prism, aiming at explaining the phenomena of interest (Van de Ven 2007). Case studies and netnography (Kozinets 2002) can be used to enhance our understanding of how stakeholders derive value, how sustainable information practices unfold, and how polycentric governance practices emerge in response to or anticipation of unfolding collective-action threats.

The proposed theoretical framing can also be used to investigate changes in sustainable information practices, collective-action threats, and stakeholder value following a change in an information commons' policies and infrastructure features. Such discoveries can be made through lab and field experiments, as well as through selective examination of naturally occurring quasi experiments—such as examining outcomes before and after a major change that reflects polycentric governance practices. In addition, many of the proposed constructs can be operationalized to be observed directly and used in inferential analysis of relationships between variables. For instance, it is possible to examine the relationship between free-riding and revitalization by measuring traffic and provision activities. Latent constructs can also be further developed to measure stakeholder perceptions psychometrically. At this early stage, we wish to avoid boxing the nascent theoretical framework into any particular research methodology; we urge

future researchers to freely adapt and dissect the proposed concepts and constructs to fit their specific lines of inquiry.

The role of science is to zoom in and break down phenomena into small components. Periodically, however, it must zoom out to examine how the puzzle pieces fit together on a grander scale. Our main goal was to take many pieces from a growing body of mostly empirical literature and try to order them in a way that facilitates a certain degree of cross-fertilization between different streams of research related to the overarching phenomenon of information commons. As such, we hope that the resulting theoretical framework will prove general enough to invite further theoretical expansion and specific enough to be immediately useful for framing empirical studies examining research questions related to the governance and sustainability of diverse types of decentralized online information systems.

2.9 References

- Aaltonen, A., and Lanzara, G. F. 2015. "Building governance capability in online social production: insights from Wikipedia," *Organization Studies* (36:12), pp 1649-1673.
- Adar, E., and Huberman, B. A. 2000. "Free riding on Gnutella," *First Monday* (5:10).
- Agarwal, R., and Karahanna, E. 2000. "Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage," *MIS quarterly* (24:4), pp 665-694.
- Aggarwal, R., and Singh, H. 2013. "Differential Influence of Blogs Across Different Stages of Decision Making: The Case of Venture Capitalists," *Mis Quarterly* (37:4), pp 1093-1112.
- Aligica, P. D., and Tarko, V. 2012. "Polycentricity: from Polanyi to Ostrom, and beyond," *Governance* (25:2), pp 237-262.
- Anderson, P. 1999. "Perspective: Complexity theory and organization science," *Organization Science* (10:3), pp 216-232.
- Anderson, E. T., and Simester, D. I. 2014. "Reviews without a purchase: Low ratings, loyal customers, and deception," *Journal of Marketing Research* (51:3), pp 249-269.
- Andrews, D. C. 2002. "Audience-specific online community design," *Communications of the ACM* (45:4), pp 64-68.
- Arazy, O., Nov, O., Patterson, R., and Yeo, L. 2011. "Information quality in Wikipedia: The effects of group composition and task conflict," *Journal of Management Information Systems* (27:4), pp 71-98.
- Auerbach, D. July. 6, 2015. "What the Reddit Rebellion is Really About," *Slate*. Retrieved from http://www.slate.com/articles/technology/bitwise/2015/07/reddit_amageddon_what_it_s_really_about.html
- Bacharach, S. B. 1989. "Organizational theories: Some criteria for evaluation," *Academy of Management Review* (14:4), pp 496-515.
- Baldwin, C., and von Hippel, E. 2011. "Modeling a paradigm shift: From producer innovation to user and open collaborative innovation," *Organization Science* (22:6), pp 1399-1417.
- Baldwin, C. Y., and Clark, K. B. 2006. "The architecture of participation: Does code architecture mitigate free riding in the open source development model?," *Management Science* (52:7), pp 1116-1127.
- Bartelt, V. L., and Dennis, A. R. 2014. "Nature and Nurture: The Impact of Automaticity," *MIS Quarterly* (38:2), pp 521-538.
- Bateman, P. J., Gray, P. H., and Butler, B. S. 2011. "Research note-the impact of community commitment on participation in online communities," *Information Systems Research* (22:4), pp 841-854.
- Beale, A. V., and Hall, K. R. 2007. "Cyberbullying: What school administrators (and parents) can do," *The Clearing House: A Journal of Educational Strategies, Issues and Ideas* (81:1), pp 8-12.
- Benkler, Y., and Frischmann, B. M. 2013. "Commons and growth: The essential role of open commons in market economies," *University of Chicago Law Review* (80), pp 1499-2071.
- Benkler, Y., and Nissenbaum, H. 2006. "Commons-based peer production and virtue*," *Journal of Political Philosophy* (14:4), pp 394-419.
- Bennett, T., Holloway, K., & Farrington, D. P. 2009. "A review of the effectiveness of neighbourhood watch," *Security Journal*, (22:2), pp 143-155.

- Bishop, J. 2007. "Increasing participation in online communities: A framework for human-computer interaction," *Computers in Human Behavior* (23:4), pp 1881-1893.
- Blank, S. 2013. "Why the lean start-up changes everything," *Harvard Business Review* (91:5), pp 63-72.
- Bolcskel, H., Paulraj, A. J., Hari, K. V., Nabar, R. U., and Lu, W. W. 2001. "Fixed broadband wireless access: state of the art, challenges, and future directions," *Communications Magazine, IEEE* (39:1), pp 100-108.
- Bonaccorsi, A., and Rossi Lamastra, C. 2003. "Altruistic individuals, selfish firms? The structure of motivation in Open Source software," *First Monday* (9:1).
- Browne, G. J., Pitts, M. G., and Wetherbe, J. C. 2007. "Cognitive stopping rules for terminating information search in online tasks," *MIS Quarterly* (31:1), pp 89-104.
- Brubaker, E. R. 1997. "The tragedy of the public budgetary commons," *Independent Review* (1:3), pp 353-370.
- Burke, M., Kraut, R., and Joyce, E. 2009. "Membership claims and requests: Conversation-level newcomer socialization strategies in online groups," *Small Group Research* (41:1), pp 4-40.
- Butler, B. S. 2001. "Membership size, communication activity, and sustainability: A resource-based model of online social structures," *Information Systems Research* (12:4), pp 346-362.
- Butler, B. S., Bateman, P. J., Gray, P. H., and Diamant, E. I. 2014. "An Attraction-Selection-Attrition Theory of Online Community Size and Resilience," *Mis Quarterly* (38:3), pp 699-728.
- Butler, B. S., and Wang, X. 2012. "The cross-purposes of cross-posting: Boundary reshaping behavior in online discussion communities," *Information Systems Research* (23:3-part-2), pp 993-1010.
- Chau, M., and Xu, J. 2012. "Business intelligence in blogs: Understanding consumer interactions and communities," *MIS Quarterly* (36:4), pp 1189-1216.
- Cheung, G. W., and Rensvold, R. B. 2002. "Evaluating goodness-of-fit indexes for testing measurement invariance," *Structural equation modeling* (9:2), pp 233-255.
- Christine Roy, M., Dewit, O., and Aubert, B. A. 2001. "The impact of interface usability on trust in web retailers," *Internet Research* (11:5), pp 388-398.
- Christopherson, K. M. 2007. "The positive and negative implications of anonymity in Internet social interactions:"On the Internet, Nobody Knows You're a Dog"," *Computers in Human Behavior* (23:6), pp 3038-3056.
- Chua, C. E. H., Wareham, J., and Robey, D. 2007. "The role of online trading communities in managing internet auction fraud," *MIS Quarterly*, pp 759-781.
- Clifford, C. Jan. 20, 2010. "EBay Sellers' Rebellion: The Aftermath," *CNN Money*. Retrieved from http://money.cnn.com/2010/01/20/smallbusiness/ebay_aftermath/
- Coffey, B., and Woolworth, S. 2004. "'Destroy the scum, and then neuter their families:' the web forum as a vehicle for community discourse?," *The Social Science Journal* (41:1), pp 1-14.
- Constantinides, P., and Barrett, M. 2014. "Information Infrastructure Development and Governance as Collective Action," *Information Systems Research* (26:1), pp 40-56.
- Covington, J., and Taylor, R. B. 1991. "Fear of crime in urban residential neighborhoods," *The Sociological Quarterly* (32:2), pp 231-249.

- Daniel, S., Agarwal, R., and Stewart, K. J. 2013. "The effects of diversity in global, distributed collectives: A study of open source project success," *Information Systems Research* (24:2), pp 312-333.
- Darke, P., Shanks, G., and Broadbent, M. 1998. "Successfully completing case study research: combining rigour, relevance and pragmatism," *Information Systems Journal* (8:4), pp 273-289.
- Davidson, K. 2014. "A typology to categorize the ideologies of actors in the sustainable development debate," *Sustainable Development* (22:1), pp 1-14.
- Dellarocas, C. 2005. "Reputation mechanism design in online trading environments with pure moral hazard," *Information Systems Research* (16:2), pp 209-230.
- Di Tullio, D., and Staples, D. S. 2013. "The governance and control of open source software projects," *Journal of Management Information Systems* (30:3), pp 49-80.
- Diakopoulos, N., and Naaman, M. Year. "Towards quality discourse in online news comments," Proceedings of the ACM 2011 conference on Computer supported cooperative work, ACM2011, pp. 133-142.
- Doty, D. H., and Glick, W. H. 1994. "Typologies as a unique form of theory building: Toward improved understanding and modeling," *Academy of Management Review*, (19:2), pp. 230-251.
- Ellison, N. B., Steinfield, C., and Lampe, C. 2007. "The benefits of Facebook "friends:" Social capital and college students' use of online social network sites," *Journal of Computer-Mediated Communication* (12:4), pp 1143-1168.
- Evans, P. 2005. "The new commons vs. the second enclosure movement: comments on an emerging agenda for development research," *Studies in Comparative International Development* (40:2), pp 85-94.
- Faraj, S., Jarvenpaa, S. L., and Majchrzak, A. 2011. "Knowledge collaboration in online communities," *Organization Science* (22:5), pp 1224-1239.
- Feller, J., Finnegan, P., Fitzgerald, B., and Hayes, J. 2008. "From peer production to productization: A study of socially enabled business exchanges in open source service networks," *Information Systems Research* (19:4), pp 475-493.
- Feuer, A, Feb. 19, 2016. "Uber Drivers Up Against the App," *The New-York Times*. Retrieved from http://www.nytimes.com/2016/02/21/nyregion/uber-drivers-up-against-the-app.html?_r=0
- Finn, J. 2004. "A survey of online harassment at a university campus," *Journal of Interpersonal Violence* (19:4), pp 468-483.
- Forman, C., Ghose, A., and Wiesenfeld, B. 2008. "Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets," *Information Systems Research* (19:3), pp 291-313.
- Forte, A., Larco, V., and Bruckman, A. 2009. "Decentralization in Wikipedia governance," *Journal of Management Information Systems* (26:1), pp 49-72.
- Füller, J., Bartl, M., Ernst, H., and Mühlbacher, H. 2006. "Community based innovation: How to integrate members of virtual communities into new product development," *Electronic Commerce Research* (6:1), pp 57-73.
- Gerstenfeld, P. B., Grant, D. R., and Chiang, C. P. 2003. "Hate online: A content analysis of extremist Internet sites," *Analyses of Social Issues and Public Policy* (3:1), pp 29-44.

- Gibbs, J. L., Ellison, N. B., and Lai, C.-H. 2010. "First comes love, then comes Google: An investigation of uncertainty reduction strategies and self-disclosure in online dating," *Communication Research*, p 0093650210377091.
- Goes, P. B., Lin, M., and Au Yeung, C.-m. 2014. "'Popularity Effect' in User-Generated Content: Evidence from Online Product Reviews," *Information Systems Research* (25:2), pp 222-238.
- Graham-Felsen, S. Sep. 7, 2006. "The Facebook Rebellion," *CBS News*. Retrieved from: <http://www.cbsnews.com/news/the-facebook-rebellion/>
- Gregor, S. 2006. "The nature of theory in information systems," *MIS Quarterly* (30:3), pp 611-642.
- Gross, D. Dec. 19, 2012. "Instagram Backtracks After user Privacy Revolt," *CNN*. Retrieved from <http://www.cnn.com/2012/12/18/tech/social-media/instagram-terms-users/>
- Gross, R., and Acquisti, A. Year. "Information revelation and privacy in online social networks," Proceedings of the 2005 ACM workshop on Privacy in the electronic society, ACM2005, pp. 71-80.
- Grover, V., and Kohli, R. 2012. "Cocreating IT Value: New Capabilities and Metrics for Multifirm Environments," *MIS Quarterly* (36:1), pp 225-232.
- Gu, B., Konana, P., Rajagopalan, B., and Chen, H.-W. M. 2007. "Competition among virtual communities and user valuation: The case of investing-related communities," *Information Systems Research* (18:1), pp 68-85.
- Habermas, J. 1985. "The theory of communicative action, Volume 2: Lifeworld and system: A critique of functionalist reason,".
- Hahn, R. W., and Wallsten, S. 2006. "The economics of net neutrality," *The Economists' Voice* (3:6).
- Halfaker, A., Kittur, A., and Riedl, J. Year. "Don't bite the newbies: how reverts affect the quantity and quality of Wikipedia work," Proceedings of the 7th international symposium on wikis and open collaboration, ACM2011, pp. 163-172.
- Hann, I.-H., Roberts, J. A., and Slaughter, S. A. 2013. "All are not equal: An examination of the economic returns to different forms of participation in open source software communities," *Information Systems Research* (24:3), pp 520-538.
- Hardin, G. 1968. "The tragedy of the commons," *Science* (162:3859), pp 1243-1248.
- Harrison, R. 2009. "Excavating Second Life Cyber-Archaeologies, Heritage and Virtual Communities," *Journal of Material Culture* (14:1), pp 75-106.
- Heer, J., and Boyd, D. Year. "Vizster: Visualizing online social networks," IEEE Symposium on Information Visualization, 2005. INFOVIS 2005., IEEE 2005, pp. 32-39.
- Hertel, G., Niedner, S., and Herrmann, S. 2003. "Motivation of software developers in Open Source projects: an Internet-based survey of contributors to the Linux kernel," *Research Policy* (32:7), pp 1159-1177.
- Hess, C., and Ostrom, E. 2003. "Ideas, artifacts, and facilities: information as a common-pool resource," *Law and Contemporary Problems* (66:1/2), pp 111-145.
- "How Facebook Became an Advertising Behemoth". April, 8. 2016. *The Economist*. Retrieved from: <http://www.economist.com/blogs/economist-explains/2016/04/economist-explains-4>
- Hinduja, S., and Patchin, J. W. 2010. "Bullying, cyberbullying, and suicide," *Archives of suicide research* (14:3), pp 206-221.

- Hosanagar, K., Han, P., and Tan, Y. 2010. "Diffusion models for peer-to-peer (p2p) media distribution: On the impact of decentralized, constrained supply," *Information Systems Research* (21:2), pp 271-287.
- Howison, J., and Crowston, K. 2014. "Collaboration Through Open Superposition: A Theory of the Open Source Way," *MIS Quarterly* (38:1), pp 29-50.
- Hu, N., Bose, I., Koh, N. S., and Liu, L. 2012. "Manipulation of online reviews: An analysis of ratings, readability, and sentiments," *Decision Support Systems* (52:3), pp 674-684.
- Huberman, B. A., Romero, D. M., and Wu, F. 2009. "Crowdsourcing, attention and productivity," *Journal of Information Science* (35:6), pp 758-765.
- Ives, B., and Olson, M. H. 1984. "User involvement and MIS success: A review of research," *Management Science* (30:5), pp 586-603.
- Jabr, W., Mookerjee, R., Tan, Y., and Mookerjee, V. 2013. "Leveraging Philanthropic Behavior for Customer Support: The case of user support forums," *MIS Quarterly* (38:1), pp 187-208.
- Jarvenpaa, S. L., and Lang, K. R. 2011. "Boundary management in online communities: Case studies of the Nine Inch Nails and ccMixter music remix sites," *Long Range Planning* (44:5), pp 440-457.
- Johar, M., Menon, S., and Mookerjee, V. 2011. "Analyzing sharing in peer-to-peer networks under various congestion measures," *Information Systems Research* (22:2), pp 325-345.
- Johnson, B. June. 26, 2009. "YouTube Users Revolt Over New Design," *The Guardian*. Retrieved from <https://www.theguardian.com/technology/blog/2009/jun/26/youtube-redesign>
- Johnson, S. L., Faraj, S., and Kudaravalli, S. 2014. "Emergence of Power Laws in Online Communities: The Role of Social Mechanisms and Preferential Attachment," *MIS Quarterly* (38:3), pp 795-808.
- Jones, Q. 1997. "Virtual-Communities, Virtual Settlements & Cyber-Archaeology: A Theoretical Outline," *Journal of Computer-Mediated Communication* (3:3), pp 0-0.
- Jones, Q., Ravid, G., and Rafaeli, S. 2004. "Information overload and the message dynamics of online interaction spaces: A theoretical model and empirical exploration," *Information Systems Research* (15:2), pp 194-210.
- Jøsang, A., Ismail, R., and Boyd, C. 2007. "A survey of trust and reputation systems for online service provision," *Decision Support Systems* (43:2), pp 618-644.
- Juvonen, J., and Gross, E. F. 2008. "Extending the school grounds?—Bullying experiences in cyberspace," *Journal of School Health* (78:9), pp 496-505.
- Kamvar, S. D., Schlosser, M. T., and Garcia-Molina, H. Year. "The eigentrust algorithm for reputation management in p2p networks," Proceedings of the 12th international conference on World Wide Web, ACM2003, pp. 640-651.
- Kane, G. C., Alavi, M., Labianca, G. J., and Borgatti, S. 2012. "What's different about social media networks? A framework and research agenda," *MIS Quarterly* (38:1), pp 275-304.
- Karakaya, M., Korpeoglu, I., and Ulusoy, Ö. 2009. "Free riding in peer-to-peer networks," *Internet Computing, IEEE* (13:2), pp 92-98.
- Kittur, A., Suh, B., and Chi, E. H. Year. "Can you ever trust a wiki?: impacting perceived trustworthiness in wikipedia," Proceedings of the 2008 ACM conference on Computer supported cooperative work, ACM2008, pp. 477-480.

- Kohli, R., and Grover, V. 2008. "Business value of IT: An Essay on expanding research directions to keep up with the times," *Journal of the Association for Information Systems* (9:1), p 23.
- Kozinets, R. V. 2002. "The field behind the screen: Using netnography for marketing research in online communities," *Journal of Marketing Research* (39:1), pp 61-72.
- Kranich, N., and Schement, J. R. 2008. "Information commons," *Annual Review of Information Science and Technology* (42:1), pp 546-591.
- Kujala, S. 2003. "User involvement: a review of the benefits and challenges," *Behaviour & Information Technology* (22:1), pp 1-16.
- Kuk, G. 2006. "Strategic interaction and knowledge sharing in the KDE developer mailing list," *Management Science*, (52:7), pp 1031-1042.
- Lampel, J., and Bhalla, A. 2007. "The role of status seeking in online communities: Giving the gift of experience," *Journal of Computer-Mediated Communication* (12:2), pp 434-455.
- Lee, D., Kim, H. S., and Kim, J. K. 2012. "The role of self-construal in consumers' electronic word of mouth (eWOM) in social networking sites: A social cognitive approach," *Computers in Human Behavior* (28:3), pp 1054-1062.
- Levina, N., and Arriaga, M. 2014. "Distinction and status production on user-generated content platforms: Using Bourdieu's theory of cultural production to understand social dynamics in online fields," *Information Systems Research* (25:3), pp 468-488.
- Listanti, M., Eramo, V., and Sabella, R. 2000. "Architectural and technological issues for future optical Internet networks," *IEEE Communications Magazine* (38:9), pp 82-92.
- Liu, J., and Issarny, V. 2004. "Enhanced reputation mechanism for mobile ad hoc networks," *Trust management*, Springer, pp. 48-62.
- Ma, M., and Agarwal, R. 2007. "Through a glass darkly: Information technology design, identity verification, and knowledge contribution in online communities," *Information systems research* (18:1), pp 42-67.
- Majchrzak, A. 2009. "Comment: Where is the theory in wikis?," *Mis Quarterly* (33:1), pp 18-20.
- Malhotra, A., Melville, N. P., and Watson, R. T. 2013. "Spurring impactful research on information systems for environmental sustainability," *MIS Quarterly* (37:4), pp 1265-1274.
- Markus, M. L. 2007. "The governance of free/open source software projects: monolithic, multidimensional, or configurational?," *Journal of Management & Governance* (11:2), pp 151-163.
- Mathwick, C. 2002. "Understanding the online consumer: A typology of online relational norms and behavior," *Journal of Interactive Marketing* (16:1), pp 40-55.
- Mayzlin, D., Dover, Y., and Chevalier, J. A. 2012. "Promotional reviews: An empirical investigation of online review manipulation," National Bureau of Economic Research.
- McKinney Jr, E. H., and Yoos, C. J. 2010. "Information about information: A taxonomy of views," *MIS Quarterly* (34:2), pp 329-344.
- Melone, N. P. 1990. "A theoretical assessment of the user-satisfaction construct in information systems research," *Management Science* (36:1), pp 76-91.
- Milgrom, P., and Roberts, J. 1995. "Complementarities and fit strategy, structure, and organizational change in manufacturing," *Journal of Accounting and Economics* (19:2), pp 179-208.

- Mishna, F., Saini, M., and Solomon, S. 2009. "Ongoing and online: Children and youth's perceptions of cyber bullying," *Children and Youth Services Review* (31:12), pp 1222-1228.
- Moon, J. Y., and Sproull, L. S. 2008. "The role of feedback in managing the Internet-based volunteer work force," *Information Systems Research* (19:4), pp 494-515.
- Morgeson, F. P., and Hofmann, D. A. 1999. "The structure and function of collective constructs: Implications for multilevel research and theory development," *Academy of Management Review* (24:2) pp 249-265.
- Muller, E. N., and Opp, K.-D. 1986. "Rational choice and rebellious collective action," *American Political Science Review* (80:02), pp 471-487.
- Nagendra, H., and Ostrom, E. 2012. "Polycentric governance of multifunctional forested landscapes," *International Journal of the Commons* (6:2).
- Nizza, M. May. 2, 2007. "The Day Digg Users Revolted," *The New-York Times*. Retrieved from http://thelede.blogs.nytimes.com/2007/05/02/the-day-the-digg-users-revolted/?_r=0
- O'Mahony, S. 2003. "Guarding the commons: how community managed software projects protect their work," *Research Policy* (32:7), pp 1179-1198.
- Ostrom, E. 1990. *Governing the commons*, Cambridge university press.
- Ostrom, E. 2009. "A general framework for analyzing sustainability of social-ecological systems," *Science* (325:5939), pp 419-422.
- Ostrom, V. 1972. "Polycentricity," in: Annual Meeting of the American Political Science Association. Washington, DC.
- Ostrom, V., Tiebout, C. M., and Warren, R. 1961. "The organization of government in metropolitan areas: a theoretical inquiry," *American Political Science Review* (55:04), pp 831-842.
- Payne, A. F., Storbacka, K., and Frow, P. 2008. "Managing the co-creation of value," *Journal of the Academy of Marketing Science* (36:1), pp 83-96.
- Plissonneau, L., and Biersack, E. Year. "A longitudinal view of HTTP video streaming performance," Proceedings of the 3rd Multimedia Systems Conference, ACM2012, pp. 203-214.
- Polyani, M. 1951. "The logic of liberty," *University of Chicago, Chicago, IL*.
- Prahalad, C. K., and Ramaswamy, V. 2004. "Co-creation experiences: The next practice in value creation," *Journal of Interactive Marketing* (18:3), pp 5-14.
- Prahalad, C. K., and Ramaswamy, V. 2013. *The future of competition: Co-creating unique value with customers*, (Harvard Business Press).
- Pulselli, F. M., Ciampalini, F., Tiezzi, E., and Zappia, C. 2006. "The index of sustainable economic welfare (ISEW) for a local authority: A case study in Italy," *Ecological Economics* (60:1), pp 271-281.
- Ramaswamy, V., and Gouillart, F. J. 2010. *The power of co-creation: Build it with them to boost growth, productivity, and profits*, Simon and Schuster.
- Ransbotham, S., and Kane, G. C. 2011. "Membership turnover and collaboration success in online communities: Explaining rises and falls from grace in Wikipedia," *MIS Quarterly* (35:3), pp 613.
- Ren, Y., Kraut, R., and Kiesler, S. 2007. "Applying common identity and bond theory to design of online communities," *Organization Studies* (28:3), pp 377-408.
- Resnick, P., Kuwabara, K., Zeckhauser, R., and Friedman, E. 2000. "Reputation systems," *Communications of the ACM* (43:12), pp 45-48.

- Rice, S. C. 2012. "Reputation and uncertainty in online markets: an experimental study," *Information Systems Research* (23:2), pp 436-452.
- Ridings, C., and Wasko, M. 2010. "Online discussion group sustainability: Investigating the interplay between structural dynamics and social dynamics over time," *Journal of the Association for Information Systems* (11:2), p 95.
- Ridings, C. M., Gefen, D., and Arinze, B. 2002. "Some antecedents and effects of trust in virtual communities," *The Journal of Strategic Information Systems* (11:3), pp 271-295.
- Ries, E. 2011. *The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses*, Crown Books.
- Rivard, S. 2014. "Editor's comments: the ions of theory construction," *MIS Quarterly* (38:2), pp iii-xiv.
- Roberts, J. A., Hann, I.-H., and Slaughter, S. A. 2006. "Understanding the motivations, participation, and performance of open source software developers: A longitudinal study of the Apache projects," *Management Science* (52:7), pp 984-999.
- Rochet, J. C., and Tirole, J. 2003. "Platform competition in two-sided markets," *Journal of the European Economic Association* (1:4), pp 990-1029.
- Rodrigues, R., and Druschel, P. 2010. "Peer-to-peer systems," *Communications of the ACM* (53:10), pp 72-82.
- Rodrik, D. 1999. "Where did all the growth go? External shocks, social conflict, and growth collapses," *Journal of Economic Growth* (4:4), pp 385-412.
- Rose, C. 1986. "The comedy of the commons: custom, commerce, and inherently public property," *The University of Chicago Law Review* (53:3), pp 711-781.
- Sarker, S., Sarker, S., Sahaym, A., and Bjørn-Andersen, N. 2012. "Exploring value cocreation in relationships between an ERP vendor and its partners: a revelatory case study," *MIS Quarterly* (36:1), pp 317-338.
- Scacchi, W. 2004. "Free and open source development practices in the game community," *Software, IEEE* (21:1), pp 59-66.
- Schlager, E., and Ostrom, E. 1999. "Property rights regimes and coastal fisheries: an empirical analysis," *Polycentric Governance and Development: Readings from the Workshop in Political Theory and Policy Analysis*, The University of Michigan Press, pp. 99-104.
- Schmierbach, M., and Oeldorf-Hirsch, A. 2012. "A little bird told me, so I didn't believe it: Twitter, credibility, and issue perceptions," *Communication Quarterly* (60:3), pp 317-337.
- Setia, P., Rajagopalan, B., Sambamurthy, V., and Calantone, R. 2012. "How peripheral developers contribute to open-source software development," *Information Systems Research* (23:1), pp 144-163.
- Setton, E., Yoo, T., Zhu, X., Goldsmith, A., and Girod, B. 2005. "Cross-layer design of ad hoc networks for real-time video streaming," *Wireless Communications, IEEE* (12:4), pp 59-65.
- Shachaf, P., and Hara, N. 2010. "Beyond vandalism: Wikipedia trolls," *Journal of Information Science* (36:3), pp 357-370.
- Shah, S. K. 2006. "Motivation, governance, and the viability of hybrid forms in open source software development," *Management Science* (52:7), pp 1000-1014.
- Sharratt, M., and Usoro, A. 2003. "Understanding knowledge-sharing in online communities of practice," *Electronic Journal on Knowledge Management* (1:2), pp 187-196.

- Shih, G. Dec. 13, 2013. "Twitter Backtracks on Block Feature After Users Revolt," Reuters. Retrieved from <http://www.reuters.com/article/us-twitter-abuse-idUSBRE9BC03520131213>
- Singh, P. V., Sahoo, N., and Mukhopadhyay, T. 2014. "How to Attract and Retain Readers in Enterprise Blogging?," *Information Systems Research* (25:1), pp 35-52.
- Singh, P. V., Tan, Y., and Youn, N. 2011. "A hidden Markov model of developer learning dynamics in open source software projects," *Information Systems Research* (22:4), pp 790-807.
- Soroos, M. S. 1982. "The commons in the sky: the radio spectrum and geosynchronous orbit as issues in global policy," *International Organization* (36:03), pp 665-677.
- Stewart, K. J., and Gosain, S. 2006. "The impact of ideology on effectiveness in open source software development teams," *Mis Quarterly*, pp 291-314.
- Stvilia, B., Gasser, L., Twidale, M. B., and Smith, L. C. 2007. "A framework for information quality assessment," *Journal of the American Society for Information Science and Technology* (58:12), pp 1720-1733.
- Suler, J. R., and Phillips, W. L. 1998. "The bad boys of cyberspace: Deviant behavior in a multimedia chat community," *CyberPsychology & Behavior* (1:3), pp 275-294.
- Sullivan, J. 2012. "A tale of two microblogs in China," *Media, Culture & Society* (34:6), pp 773-783.
- Susarla, A., Oh, J.-H., and Tan, Y. 2012. "Social networks and the diffusion of user-generated content: Evidence from YouTube," *Information Systems Research* (23:1), pp 23-41.
- Svensson, M., and Larsson, S. 2012. "Intellectual property law compliance in Europe: Illegal file sharing and the role of social norms," *New Media & Society* (14:7), pp 1147-1163.
- Tapscott, D., and Williams, A. D. 2008. *Wikinomics: How mass collaboration changes everything*, (Penguin).
- Teo, H.-H., Chan, H.-C., Wei, K.-K., and Zhang, Z. 2003. "Evaluating information accessibility and community adaptivity features for sustaining virtual learning communities," *International Journal of Human-Computer Studies* (59:5), pp 671-697.
- Thorson, K., Vraga, E., and Ekdale, B. 2010. "Credibility in context: How uncivil online commentary affects news credibility," *Mass Communication and Society* (13:3), pp 289-313.
- Tkacz, N. Jan. 20, 2011. "The Spanish Fork: Wikipedia's Ad-Fuelled Mutiny," *Wired*. Retrieved from <http://www.wired.co.uk/news/archive/2011-01/20/wikipedia-spanish-fork>
- Toma, C. L., Hancock, J. T., and Ellison, N. B. 2008. "Separating fact from fiction: An examination of deceptive self-presentation in online dating profiles," *Personality and Social Psychology Bulletin* (34:8), pp 1023-1036.
- Tsai, H.-T., and Bagozzi, R. P. 2014. "Contribution Behavior in Virtual Communities: Cognitive, Emotional, and Social Influences," *MIS Quarterly* (38:1), pp 143-163.
- Van de Ven, A. H. 2007. *Engaged scholarship: A guide for organizational and social research*, (Oxford University Press on Demand).
- Venkatraman, N. 1989. "The concept of fit in strategy research: Toward verbal and statistical correspondence," *Academy of Management Review* (14:3), pp 423-444.
- Von Krogh, G., Haefliger, S., Spaeth, S., and Wallin, M. W. 2012. "Carrots and rainbows: Motivation and social practice in open source software development," *MIS Quarterly* (36:2), pp 649-676.

- Wall, D. S., and Williams, M. 2007. "Policing diversity in the digital age Maintaining order in virtual communities," *Criminology and Criminal Justice* (7:4), pp 391-415.
- Wasko, M. M., and Faraj, S. 2005. "Why should I share? Examining social capital and knowledge contribution in electronic networks of practice," *MIS Quarterly*, pp 35-57.
- Wasko, M. M., Faraj, S., and Teigland, R. 2004. "Collective action and knowledge contribution in electronic networks of practice," *Journal of the Association for Information Systems* (5:11), p 2.
- Webster, J., and Watson, R. T. 2002. "Analyzing the past to prepare for the future: Writing a," *MIS Quarterly* (26:2), pp 13-23.
- Whetten, D. A., Felin, T., and King, B. G. 2009. "The practice of theory borrowing in organizational studies: Current issues and future directions," *Journal of Management* (35:3), pp 537-563.
- Wiertz, C., and de Ruyter, K. 2007. "Beyond the call of duty: Why customers contribute to firm-hosted commercial online communities," *Organization Studies* (28:3), pp 347-376.
- Wilson, S. M., and Peterson, L. C. 2002. "The anthropology of online communities," *Annual Review of Anthropology*, pp 449-467.
- Wu, D., Hou, Y. T., Zhu, W., Zhang, Y.-Q., and Peha, J. M. 2001. "Streaming video over the Internet: approaches and directions," *Circuits and Systems for Video Technology, IEEE Transactions on* (11:3), pp 282-300.
- Zhang, C., Hahn, J., and De, P. 2013. "Research Note-Continued Participation in Online Innovation Communities: Does Community Response Matter Equally for Everyone?," *Information Systems Research* (24:4), pp 1112-1130.
- Zhang, P., and Von Dran, G. M. 2000. "Satisfiers and dissatisfiers: A two-factor model for website design and evaluation," *Journal of the American Society for Information Science* (51:14), pp 1253-1268.
- Zmud, B. 1998. "Editor's Comments Volume 22 Iss. 3," *MIS Quarterly* (22:3), p 1.

APPENDIX: Reviewed IS Research

In the 1980s, Elinor Ostrom and her colleagues at Indiana University's Workshop in Political Theory and Policy Analysis set out to inform CPR theorization by collecting, sorting, and analyzing more than 5,000 empirical field studies of local resource management arrangements around the world. The scrutinized field studies did not use CPR or polycentricity language, and they came from many different research domains. To unify this diverse body of empirical work from various research domains, the workshop scholars chiefly focused on identifying the structural characteristics of the observed resource system, the attributes and behaviors of its members, the rules it used, and the reported outcomes (Ostrom 1990). This grounded theory approach was complemented by the use of existing taxonomy and terminology from polycentricity research on public-goods governance (Ostrom et al. 1961; Ostrom 1972). The resulting framework became the theoretical foundation that was used over the next 30 years in CPR governance research (also known as collective-action research). The framework was constantly refined by new insights from the field, but its core remained intact.

On a much smaller scale, we emulated this process. Using core taxonomy and terminology from CPR governance research to guide our effort, we systematically reviewed 73 studies of crowdsourcing, social media, online communities, electronic word of mouth, peer-to-peer networks, and open source software (OSS) development published in *Information Systems Research* and *MISQ* between 2005 and 2014. We specifically focused on extracting and translating insights on rules and their use, stakeholders and their derived value, evidence on collective-action threats, evidence of Ostrom's eight governance principles, infrastructure and design features (*environmental conditions* in CPR research), and outcomes (direct effects and indirect

externalities). We knew that we would not be able to reach the scope of Ostrom and her colleagues; we therefore explicitly focused our effort on the leading IS research journals to provide a solid foundation. To frame our discussion, we also branched to other sources for additional evidence and perspectives. Although our primary goal is theory development and not a literature review, we did use literature review methodologies to find relevant research on decentralized online information systems. We followed Jane Webster and Richard Watson's (2002) guidelines for performing effective literature reviews and examined reference sections to identify many other important papers published in *ISR* or *MISQ* prior to 2005, as well as in other IS (and related fields') journals. To avoid saturation, we cast our net wide (but not too wide), using standard search engines such as Ebscohost and Google Scholar to find additional relevant material. This inclusive strategy gave us additional evidence and insights to substantiate the many theoretical claims and propositions presented in this manuscript.

Table A: Overview of the Reviewed IS Research				
Paper	Focus	Findings	Adaptation	Coding
Aggarwal & Singh 2013	Examines the impact of technology blog content on venture capitalist investor decision-making	Blogs impact decision-making in the initial screening stage, but less so in subsequent stages. Blog exposure gave better negotiation leverage to the entities featured.	Information commons reduce appropriators' search costs and impact their decision-making. In some contexts, information commons increase competition between appropriators for information, leading to increased transaction costs.	appropriation; appropriator value; exploration
August et al. 2013	Examines different strategies available to firms that sponsor OSS development	When contributors are efficient, sponsors should pursue an open software license and focus on capitalizing through support services sales. When contributors are inefficient, sponsors should pursue a proprietary license and sell each software unit	Producers of information commons are often profit motivated and must determine the best strategy when seeking a balance between attracting effective providers and preventing free-ridership from competitors who benefit	production; producer value; provision; free-riding

		separately to prevent contributors from spreading the software code to competitors.	from sitting on the sidelines.	
Bapna et al. 2008	Examines how much surplus consumers receive from eBay auctions	Consumers extract a median surplus of at least \$4 per eBay auction. In 2003, eBay's auctions generated at least \$7.05 billion in total consumer surplus.	Transaction information commons directly benefit appropriators.	appropriation; appropriator value
Bartelt & Dennis 2014	Examines how instant messaging and discussion boards are utilized and subsequently impact quality of outcomes	Genre rules can have powerful effects that significantly impact team behavior and performance.	Operational rules are as important as infrastructure for the sustainability of information commons. Language used and built into a commons' features influence its operational rules, and those rules become increasingly entrenched in the fabric of the commons over time.	operational rules; infrastructure; incremental adaptation
Bateman et al. 2011	Presents a model to unify conflicting explanations as to why people engage in discussion forums	Members may have psychological bonds to a particular online community based on need, affect, and obligation. Each form of community commitment has a unique impact on each behavior, with need-based commitment predicting thread reading; affect-based commitment predicting reply posting and moderating behaviors; and obligation-based commitment predicting only moderating behavior.	The need for information drives appropriation. The need to make an impact drives provision. A sense of obligation to the commons drives certain providers to monitor their peers.	provision; provider value; appropriation; appropriator value; operational rules; shared accountability
Browne et al. 2007	Examines why people stop searching for information online	Mental list and single criterion rules are used more often when people search for relatively straightforward information (such as product information); people use magnitude threshold and representational stability	Appropriation of information depends on the appropriator's task. More complex and abstract tasks, such as diagnosis, adhere to the representational stability rule. Other rules for stopping information searches are: mental list, difference threshold,	appropriation; appropriation value

		rules when searching for complex information.	single criterion, and magnitude threshold.	
Butler & Wang 2012	Examines the extent to which boundary reshaping in an online discussion community impacts member dynamics and responsiveness	An inherent tension exists between content boundary management and reshaping the community. Reshaping behaviors within a discussion forum affect member dynamics and community responsiveness in both positive and negative ways.	Information commons content boundaries impact information governance. Flexible boundaries afford more engaged provisioning, while greater engagement leads to changes in the commons that potentially drive other providers away.	boundary regulation; provision; operational rules; constitutional rules; provision; pollution; rebellion
Butler et al. 2014	Presents a model of key latent constructs influenced by technology choices and the possible causal paths by which they dynamically effect communities	Community size and resilience directly affect the community's sustainability over time. The lower the participation costs and the more focused the issue, the better chances for sustainability.	Critical mass of provision and appropriation directly impact the sustainability of the information commons. The lower the opportunity cost of provision and appropriation and the more focused the subject, the more likely the commons is to be sustainable.	provision; appropriation; boundary regulation
Chan & Ghose 2014	Examines the connection between Craigslist and spread of HIV	Entry of Craigslist is related to a 15.9% increase in HIV cases. The analysis suggests that the site entry produces an average of 6,130–6,455 cases of HIV infection in the United States each year. Analyses reveal that nonmarket-related casual sex is the primary driver of the increase in HIV cases.	Example of a direct effect of a polycentric information commons that is also a social externality.	appropriation; appropriation value
Chau & Xu 2012	Proposes a framework for gathering business intelligence from blogs by automatically collecting and analyzing blog contents and bloggers' interaction networks	Networks of bloggers have different centers of influence. The networks are decentralized and do not exhibit pattern structures, such as star and hierarchical structures. These implicit communities have been formed spontaneously.	Polycentric information commons are decentralized and spontaneously arising.	provision; polycentricity

Chua et al. 2007	Examines how community members monitor fraud on auction sites	Communities monitor for fraudulent behavior. Engaged community members can often detect suspicious activity better than outside official authorities. Tactics are: direct intervention against those who commit fraud, notification of other community members, and involvement of outside authorities.	Opportunistic providers can pollute information to gain at the expense of other members in the information commons. Members of the information commons often combat pollution directly and by involving outside authorities.	pollution; violation; provision; shared accountability
Claussen et al. 2013	Examines how rule modification by Facebook that rewarded quality app developers with access to Facebook's users impacted application development	The rule change led to the development of new applications with significantly higher user satisfaction.	Producers have control over rules and infrastructure design that impact information provision and subsequent appropriation. Loosening control and increasing polycentricity, for example, can lead to improved provision and more satisfied appropriators.	production; provision; infrastructure; constitutional rules; incremental adaptation; boundary setting
Daniel et al. 2013	Examines the impact of diversity among developers on engagement and market success	Variation in participants' contribution-based reputation is positively associated with success. Diversity in the spoken language and country of participants has a negative impact on community engagement, but a positive effect on market success. Dispersion in project participant roles positively influences community engagement and market success. Diversity's impact on market success is conditional on the project stage; it is most important at the very beginning and toward the end.	Polycentric governance of information facilitates greater diversity that, in turn, enhances outcomes. Diversity is especially important in the birth stage and when reaching maturity. During the growing stage, diversity is less desirable.	provision; polycentricity

Dellarocas 2005	Presents a theoretical analysis of reputation mechanism design in trading environments with pure moral hazard	Reputation mechanisms on eBay mitigate moral hazard and adverse selection, thus increasing efficiency in markets. No one-size-fits-all set of guidelines exists for reputation mechanism design. Depending on the context, reputation mechanisms can act as a sanctioning tool or as a signaling mechanism.	An externality of reputation information commons is increased efficiency in markets. No one-size-fits-all set of guidelines exists for reputation mechanism design. Depending on the context, reputation mechanisms can act as a sanctioning tool or as a signaling mechanism.	pollution; producers; shared accountability
Dewan & Ramapras 2012	Examines the relationship between music blogs and music sales	Music sharing reduces the sales of songs.	P2P sharing causes an economic externality.	appropriation; appropriator value
Dewan & Ramapras 2014	Examines the impact of music blogs on music consumption	Niche music receives substantially more attention in blogs. This increased exposure leads to more consumption of niche music.	Polycentric information commons produce more diverse information; that, in turn, leads to a decrease in the influence of monocentric information sources.	appropriation; appropriator value; exploration
Duan et al. 2009	Examines whether individuals exhibit a herd behavior pattern when choosing software online	Individuals do exhibit herd behavior when choosing software; online reviews have particular impact on the demand for lesser known products.	The opportunity cost of information search impacts the intensity of appropriation. Online reviews improve the odds of marginal products and services at the expense of more established products and services; this displacement can be viewed as an externality.	appropriation
Feller et al. 2008	Examines how a network of firms manages OSS development	Infrastructure impacts social interactions. Shared goals and norms drive collaboration, along with the ability to impose collective sanctions on those who violate the shared norms.	Polycentric governance of information commons is a socio-technical process in which social interactions that drive governance of information depend on the physical architecture of the commons. Shared norms and goals drive collaboration between providers.	production; provision; boundary setting; shared accountability

Fitzgerald 2006	Challenges the notion that open source is driven by volunteers and the notion of collective intelligence	The open source model moved toward greater “professionalism” with more planning and less spontaneous and paid contributors.	Open source moves from open access regime to common property regime.	regimes
Forman et al. 2008	Examines the relationship between identity disclosure and perceived usefulness of reviews	Community norms impact disclosure of personal identity. Disclosure of identity of review contributors increases the perceived helpfulness of the review. In addition, prevalence of identified reviewers increases sales.	Information commons’ operational rules and norms impact providers’ choice to disclose their identity. Appropriators tend to value information from non-anonymous providers. A culture of transparency adds value to the information commons.	operational rules; provision; appropriation; appropriator value
Ghose 2009	Examines how online reviews impact seller and product uncertainty, consequently impacting information asymmetry	Reputation signals do not mitigate seller or product uncertainty.	This contradicts other findings. Online reviews do not mitigate information asymmetry; information does not relate a positive externality.	externalities
Goes et al. 2014	Examines the change review writers undergo as users subscribe to them	As review writers become more popular, they produce more reviews and more objective reviews; however, their numeric ratings also systematically change and become more negative and more varied.	Information providers are motivated by demand from appropriators. As time passes, they learn to provide more objective information.	provision; provider value; appropriation
Goh et al. 2013	Compares the impact of both user-generated content and marketed content on purchasing behavior	User-generated content is significantly more influential in driving purchasing behavior.	Polycentric information can have a direct economic impact on its subject source.	direct effect; appropriation; appropriator value

Gu et al. 2007	Examines how users value virtual communities and how virtual communities differ in their value propositions	Quality of information is more important than quantity for the growth of virtual investment communities.	Information quality is more important than quantity because of the cognitive cost associated with processing information. High-quality information is more likely to lead to externalities.	appropriation; appropriator value
Gu et al. 2012	Examines the difference in impact of reviews on sales of a “high-involvement” product (cameras) between reviews found on the retailer site and external review sources	External sources play a greater role in impacting sales of high-involvement products compared to reviews posted on the retailer site.	Appropriators triangulate information from different information commons prior to making major purchasing decisions.	appropriation
Hahn et al. 2008	Examines how OSS development teams are formed and how individuals make decisions about which teams to join	Past ties between initiator and collaborator impact present participation.	Past relationships between core providers and peripheral providers play a role in the formation of provision network on a polycentric information commons.	provision
Han et al. 2012	Examines the business and strategic value of an open, collaborative organizational form in which participating firms co-create economic value by jointly developing and co-marketing of IT innovations	Participating firm’s market valuation increased; not participating benefited rivals due to knowledge spillover. Heterogeneity is not associated with value creation. Development of radical innovation adds more value to the firm. Limited access and decision authority is better than unlimited access and decision authority.	Polycentric governance of information creates synergy, especially when rules and boundaries are well defined. Stakeholders’ level of heterogeneity is irrelevant to quality. Information created spills over— and other firms benefit as well.	production; boundary setting
Hann et al. 2013	Examines whether open source programmers get higher wages due to their experience in open source	Credentials obtained from programming OSS are associated with up to an 18% increase in wages.	Volunteer information providers can potentially benefit financially.	provision; provider value

Hardaway & Scamell 2012	Proposes changes in the academic research publication process to incorporate greater transparency, inclusivity, and openness in reviews as well as publication decisions.	An inclusive review process that is open to participants (but not too open) will improve the quality of published research and the fairness of the publication process.	Polycentric information commons of academic information will increase the suitability and quality of academic research over time.	production; producer value; provision; provider value
Hildebrand et al. 2013	Examines how feedback from other users impacts the satisfaction and creativity of individuals designing their own products	Early feedback from community members reduces the boldness of the design and reduces creativity.	Providers check and balance each other to a point of reducing creativity and individual expression.	provision
Howison & Crowston 2014	Presents a theory about building artifacts out of information and the organizational affordances of those artifacts	OSS development is more about work redesign than about how to run an effective virtual team, and more about task work in context than generic team work processes. Layerability, open distribution, irrevocable contribution, and time are conditions that contribute to open source success.	Polycentric information commons is individual rather than team driven. Polycentric information commons are characterized by compartmentalization and incremental improvements that can take time. Information providers respect the autonomy of other providers.	provision; production; incremental adaptation
Hsien-Tung & Bagozzi 2014	Examines the factors driving contribution in virtual communities	Social identity and group norms drive decisions to contribute.	Information providers in the commons assume a certain collective identity relating to the information's subject matter.	provision; provider value; shared goal; provider recognition
Jabr et al. 2014	Examines the factors driving contribution in online technical support forums	Peer feedback mechanisms are associated with better quality and technical support compared to quantity feedback (just a count of contribution). The results indicate that problem solvers are partially motivated by the need to be recognized by their peers.	Information providers value the feedback from other providers and are partially motivated by it. Existence of peer feedback mechanisms thus increase the likelihood of long-term sustainability of the information commons.	provision; provider value; provider recognition; infrastructure

Jarvenpa a & Majchrzak 2010	Commentary on how trust asymmetry, deception, and novelty impact collaboration online	Participants manage vigilant interactions in different ways.	Conflict between providers impacts relationships in information commons.	provision; violation; shared accountability
Johar et al. 2011	Examines how P2P congestion impacts propensity to share	Sharing increases as the network grows, reducing congestion. The community can reduce the impact of free-riding by increasing individual sharing.	Information is a resource that can be subjected to congestion and free-riding.	free-riding; congestion; provision
Johnson et al. 2014	Examines the emergence of power law distributions via the mechanisms of preferential attachment, least efforts, direct reciprocity, and indirect reciprocity	No single unitary motivation drives participation in online communities but many different—sometimes competing, sometimes complimentary—motivations exist. New participants can become popular at any time, replacing previously popular participants as social centers.	Polycentric governance of information is dynamic, and new autonomous centers of authority emerge and disappear constantly without a pattern. Multiple individual motivations drive information provision.	provision; provider value; polycentricity
Kane et al. 2014	Outlines a set of theoretically distinct questions raised by the introduction of social media in and by organizations to formulate a research agenda	Social media introduces questions of platform design into social network analysis. Social media gives users capabilities that they do not possess in traditional offline social networks; these capabilities include visualizing network structure and searching for content in a network without using relational ties.	The architectural features of the polycentric information commons directly impact stakeholders and indirectly impact the prospects of long-term sustainability. Polycentric information commons characterized by transparency and open access to information.	production; appropriation;
Levina & Arriaga 2014	Presents an analytical lens for studying social status production processes across different user-generated content platforms	Content production is a social process that occurs in different specific contexts in which different status signals play a role.	Provision of information is a context-dependent social process.	provision; provider value; provider recognition

Li & Hitt 2010	Tests whether online reviews more closely reflect a product's quality or its price value	Online review are biased towards the product value—that is, ratings reflect price more than they reflect value.	An example of unintentional information pollution.	provision; appropriation; pollution
Ludwig et al. 2014	Examines to what extent communication-style alignment impacts participation quantity and quality	Stronger trend of alignment leads to greater participation quantity and quality. At a community level, greater synchronicity in the linguistic style across all community members fosters individual members' participation behavior.	Operational rules (norms) of communication impact provision.	operational rules; provision
Lukyanenko et al. 2014	Examines the impact of modeling decisions on information quality in amateur content-creation settings	Participants provide more accurate information when classifying phenomena at a more general level. In addition, we found greater overall accuracy when participants could provide free-form data compared to a condition in which they selected from constrained choices.	Polycentric information provision benefits from freedom of action.	provision; autonomy
Ma & Agarwal 2007	Examines a feature that supports reputation and identity management, as well as archives of past interactions, to assess the impact on online community members' satisfaction and willingness to contribute	IT artifacts that allow reputation management improve perceived identity verification, cause greater satisfaction, and increase contributions by members.	Appropriators and peer providers prefer to interact with those who have an online persona (not complete transparency or complete anonymity) and also need to have an online persona to feel validated by others. Interaction with those who have an online persona leads to greater satisfaction with the commons and increases the likelihood of provision.	provision; provider value; appropriation; operational rules; peer recognition
Ma et al. 2014	Examines the impact of pre-release movie piracy on movie sales	Sales decrease approximately 20% due to piracy.	Example of an economic externality that impacts the film industry.	appropriation; appropriator value; externalities
McLure Wasko & Faraj 2005	Examines why people contribute to	People contribute their knowledge when they perceive that it enhances their professional	Provision of information is driven by mostly personal reasons and not by social motivations.	provision; provider value free-riding

	discussion forums	reputations, when they have the experience to share, and when they are structurally embedded in the network. Surprisingly, contributions occur without regard to expectations of reciprocity from others or high levels of commitment to the network.	Polycentric information commons tolerate free-riding.	
Meservy et al. 2014	Examines how individuals filter knowledge encountered in online forums	Peripheral cues (source expertise and validation) have a greater influence on knowledge filtering decisions than does the content quality of the solution.	Appropriators rely on heuristics when evaluating information. Provider's status and perceived expertise are used as a proxy for quality of information.	appropriation
Moon & Spraul 2008	Examines the role of feedback on participation in online communities	In forums that incorporate a systematic quality feedback system, answer providers contribute more often and technical problem resolution is more effective. Volunteers who produce higher-quality contributions have longer participation duration, and participation duration is positively associated with community maintenance contributions.	Feedback from appropriators and peer providers motivates provision and increases its overall quality.	provision; provider value; provider recognition
Nan & Lu 2014	Examines patterns of crisis management arising unintentionally from decentralized and spontaneous actions in an online community during an earthquake	Fluctuation of message content themes in this online community served to energize continuous input from ordinary organization members. Self-organization is a viable source of organizational order that complements the traditional centralized plan-and-control approach.	Polycentric information governance is characterized by diverse provision patterns that attract further diverse provision, resulting in synergy.	polycentricity; provision

Kumar & Benbasat 2006	Examines how an infrastructure that facilitates consumer reviews and recommendations impacts the perceived usefulness and social presence of the website incorporating them	Infrastructure that facilitates consumer reviews and recommendations impacts positively the perceived usefulness and social presence of the website incorporating them.	Producers of platforms that facilitate provision of eWOM benefit because appropriators of eWOM information see the websites as more useful and pleasing.	appropriation; appropriator value
Oh et al. 2013	Examines the factors causing rumors on twitter	Source ambiguity is the most important factor in creating rumors, followed by personal involvement; anxiety is the least yet marginally important rumor-causing factor.	Information source ambiguity impacts information provision.	provision; pollution
Park et al. 2013	Examines the impact of information on expectation and trading behavior in the context of a stock investment discussion forum	Investors tend to cluster with similar investors and generally do not benefit from the diversity offered by the community.	Appropriators do not always benefit from the diversity of information found in the information commons, as different biases lead them to primarily seek information that suits their preconceived knowledge.	appropriation
Pavlou & Dimoka (2006)	Examines the impact of written reviews on trust in sellers and the price premiums they receive from their online reputation	Written reviews explain 50% of the variance in price premiums, suggesting that reviews act as an information asymmetry-mitigating mechanism.	Polycentric governance of reputation information mitigates information asymmetry in markets, subsequently increasing market efficiency.	appropriation; shared accountability
Ransbotham & Kane 2011	Examines longitudinally how retention and turnover of contributors impacts the status of Wikipedia articles	Membership retention relates in a curvilinear fashion to effective collaboration: it's important in initial stages, but becomes less important and turnovers become more important; moderate levels of membership turnover are desirable.	Polycentric information provision is dynamic as providers replace previous core providers in a continuous manner.	provision; polycentricity
Ray et al. 2014	Develops the concept of engagement in online communities as	Engagement is a key factor driving knowledge contribution, but not necessarily satisfaction. Engagement and	Providers and appropriators must feel a sense of engagement to gain a sense of collective purpose that leads to	provision; provider value; appropriation

	key variable for driving knowledge contribution and diffusion across the community through word of mouth. Engagement and satisfaction are parallel mediating forces at work in online communities	satisfaction together drive positive word of mouth and subsequent growth.	information provision and appropriation. Engagement motivates providers of information and, together with satisfaction, is associated with being a member of the information commons and drives positive word of mouth, contributing to its growth.	
Ren et al. 2012	Examines how identity and communication features impact the attachment of online community members over time	Group identity features that allow members to see profiles and facilitate repeated exposure to the group activities impact traffic twice as much as interpersonal bond features.	Providers and appropriators are more motivated by group identity than personal ties.	provision; provider value; appropriation; provider recognition
Rice 2012	Examines how reputation is evolving and how it affects transactions	Reputation mechanisms are significant for purchase decisions regarding specific sellers, while specific reputation information is associated with the overall decision of whether to transact in the marketplace. When reputations are increasingly noisy, buyers are less likely to react negatively to poor ratings and are more likely to give sellers the benefit of the doubt.	Polycentric governance of reputation information mitigates information asymmetry in markets, subsequently increasing markets efficiency. Pollution of reputation information confuses appropriators and decreases efficiency.	appropriation; pollution; shared accountability
Setia et al. 2012	Examines the role of peripheral developers in the creation of OSS	Peripheral developers are instrumental, especially when the project reaches maturity.	While core information providers are important at the beginning, as time passes, peripheral providers become increasingly important.	provision
Shi et al. 2014	Examines patterns of information diffusion on Twitter	Unidirectional (weak) rather than bidirectional social ties exhibit the fastest pattern of information diffusion as users pass content to other users.	Information diffusion is the fastest through encounter dynamics of provision and appropriation in information commons.	provision; appropriation; relationship dynamics

Singh et al. 2011	Examines to which extent individuals involved in OSS development learn from their own experience and from interacting with peers	Learning from peers is most important for developers. Developers benefit from discussions they initiate in the middle of the project; at the beginning and end of the project, developers benefit more from participating in discussions initiated by others.	Providers learn both from other providers (especially in the middle of the project) as well as from their own experiences.	provision; operational rules
Singh et al. 2013	Examines how text characteristics attract readers of blogs	Sentimental textual characteristics (negative, positive, or controversial tone) attract and retain readers, while quality textual characteristics (grammar, readability, and comprehensiveness) have no impact on attracting readers, but do impact retention.	Appropriators care more about the content of the information than about how well it presented; however, presentation plays a role in retaining appropriators.	appropriation; appropriator value
Smith & Telang 2009	Looks at the effect of piracy on sales of movies	Giving away content in one channel can stimulate sales in a paid channel if the free content is sufficiently differentiated from its paid counterpart.	Piracy (information commons) can have a positive effect of sales if the content is sufficiently differentiated.	externalities; direct effect
Stewart & Gosain 2006	Examines the role of ideology in OSS	Generally, team members' adherence to the tenets of the OSS community ideology impact trust and communication in teams, which positively relates to effectiveness.	Provision outcomes in relationship dynamics are impacted by ideological alignment between providers.	provision; provider value; shared goal
Stewart et al. 2006	Develops and tests a model of the impacts of license restrictiveness and organizational sponsorship on two indicators of success: user interest in, and development activity on, OSS development projects	Users are mainly attracted to OSS projects initiated by noncommercial interests that do not employ restrictive licenses. Identity of the sponsor of the projects impacts developer participation.	Ideological synchronization between producers and providers is a condition for an information commons' ability to attract providers and achieve sustainability. Providers prefer constitutional rules that are flexible.	provision; provider value; production; producer value; constitutional rules; rebellion

Turel et al. 2011	Examines how user cognition and ultimately usage intentions toward an information technology are distorted by addiction to the technology	The level of online auction addiction distorts the way the IT artifact is perceived.	Over appropriation can lead to addiction.	appropriation
von Krogh et al. 2012	Reviews the literature on the motivations of OSS developers	Current research on the motivation of software developers does not go deep enough.	Provision of information is motivated by multiples variables.	provision; provider value
Wen et al. 2013	Examines the effect of copyright infringement lawsuits both on the open source development projects being sued and on related projects that employ the disputed technology	Projects that exhibit high technology overlap with the disputed OSS experienced a 15% greater decline in user interest and 45% less developer activity than projects in the control group; OSS projects that are intended for business and specific to the disputed OSS platform had a 34% greater decline in user interest and 86% less developer activity than the control group.	Challenge from the outside by powerful interests threatens the sustainability of a polycentric information commons. The impact is stronger on providers than on appropriators.	provision; appropriation
Xu & Zhang 2013	Examines the impact of firm information on Wikipedia on insider trading	Information on Wikipedia causes firm bosses to disclose negative news about the firm's financial performance faster, reducing the potential gain from insider trading.	An example of market externality of an information commons.	appropriation; externalities
Yan & Tan 2014	Examines whether social support exchanged in an online healthcare community benefits patients' mental health	Patients benefit from learning from others, and their participation in the online community helps them improve their health and better engage in their disease self-management process. Evidence exists that informational support is the most prevalent type in the online healthcare community, but	An example of a direct effect of a polycentric information commons; appropriation can be psychologically beneficial.	appropriation; appropriator value

		emotional support plays the most significant role in helping patients move to a healthier state.		
Ye et al. 2014	Examines how sellers respond to changes in the design of eBay's reputation systems	Coercing buyers to revoke their negative feedback through retaliation enables low-quality sellers to manipulate their reputations and masquerade as high-quality sellers. Low-quality sellers exerted additional efforts to improve their reputation scores after the system changed.	Opportunistic information providers pollute the commons to advance their interests. Change in constitutional rules by producers changed the balance of power between providers and appropriators, and subsequently reduced pollution.	provision; appropriation; pollution; violation boundary regulation; incremental adaptation, shared accountability
Zhang et al. 2013	Examines the factors that influence online innovation and community members' continued participation	Both users and software modifiers are influenced by community response, but users seem to be influenced by it more. Participation is primarily motivated by software needs, the enjoyment of programming, the desire to improve the software, the opportunity to learn about others' improvements to the software, and the gratification of helping others and giving back to the community.	Engagement of appropriators leads to continuous appropriation. Providers are partially motivated by feedback from appropriators, but mostly by personal reasons.	provision; provider value; provider recognition; appropriation

CHAPTER 3: HARNESSING DIGITAL ACTIVISM FOR INSTITUTIONAL CHANGE: A COLLECTIVE-ACTION PERSPECTIVE

ABSTRACT

Interest in how information technology is impacting societal issues is growing. Although considerable evidence shows that online communities and social media spur social activism online and offline, we know little about whether—and under what conditions—this digital activism translates into institutional change. To address this void, we draw on Polycentric Information Commons theory to advance knowledge on how partisan actors can harness digital activism to create institutional change. We assert that the success of such efforts depends on the actors' capacity to use a dedicated information commons to generate legitimate collective action. To inform this theorizing, we examine the case of Cabotagestudien, in which partisan actors used social media and smartphone-based crowdsourcing to mobilize and drive collective action that successfully fended off attacks from other institutional actors and changed haulage regulations and supply chain practices in Northern Europe.

Key Words: Digital activism, institutional change, collective action, polycentric information commons, legitimacy

3.1 Decentralized Systems Online and Societal Change

Information systems (IS) scholars increasingly identify a need for research on the impact of information technology (IT) on societal issues (Saunders 2007; Walsham et al. 2007, Majchrzak et al. 2016). In recent years, researchers have published many important studies on topics such as the digital divide and poverty (Agarwal et al. 2009; Hsieh et al. 2008), IT in developing countries (Braa et al. 2004; Leonardi et al. 2016; Venkatesh, and Sykes 2013), the impact of E-governance on corruption (Srivastava et al. 2016), and IT's impact on environmental sustainability (Elliot 2001; Melville 2010). Increasingly, researchers are examining how online-based technologies such as crowdsourcing and online communities are used in promoting societal issues (Ketter et al. 2016; Selander and Jarvenpaa 2016). It is increasingly evident that the internet can be useful for driving activism online and for mobilizing activists in the physical world, however, we still lack theoretical understanding on how it can be purposefully harnessed to change institutional dynamics in substantial and lasting ways. Following calls for greater theoretical grounding in research on the impact of technology on societal issues (Majchrzak et al. 2016), we closely examine how online technologies can be harnessed to promote institutional change.

In the late 1990s, the internet began transitioning from a novel broadcast media that business and government organizations used to pass content to passively browsing users, to a decidedly more open medium in which participating users actively create content, write posts and blogs, upload images and videos, and even develop code (Tapscott and Williams 2008). This transition challenged long-standing economic theories of production by demonstrating that loosely organized private individuals can produce outcomes previously thought possible only by hierarchical systems with clear incentive structures (Tapscott and Williams 2008). For two decades now, various academics and other interested observers have studied these new open access systems that rely on

self-selecting individuals for content and information. This research area is rather broad and touches on many questions, including how we might better understand user motivations (Von Krogh et al. 2012), underlying dynamics (Levina and Arriaga 2014), network effects (Ganley and Lampe 2009), performance of specific tasks (Majchrzak and Malhotra 2013), and direct impacts on user well-being (Yan and Tan 2014). A small but growing subset of this research stream focuses on how open access systems might affect broader societal outcomes, including business (Andriole 2010), financial markets (Xu and Zhang 2013), political campaigns (Wattal et al. 2010), and public health (Goh et al. 2016).

Evidence demonstrates that online communities and social media can be a fertile ground for spurring social activism on- and offline (Enjolras et al. 2013; White 2016). Still, we lack a theoretical understanding of whether, and under what conditions, this digital activism translates into *institutional change*—that is, a high-impact transformation of existing social, political, and economic arrangements. Researchers have noted that digital activism campaigns online, while effective in achieving short-term impacts, typically lose steam rather quickly and generally fail to generate more permanent change in the institutions of society (Obar et al. 2012; Selander and Jarvenpaa 2016). Although institutional change is not necessarily the main objective of digital activism online, evidence from other campaigns that openly challenge institutional order and rely heavily on online technologies for mobilization show that they are rarely successful. As a case in point, we might consider the well-known Occupy Wall-Street and Arab Spring movements which, according to their founders, failed to achieve institutional change (White 2016, 2017). Given the growing visibility—and shortcomings—of digital activism, it becomes increasingly important to understand, theoretically, what might be holding it back and how it can be more effectively leveraged to drive high-impact change.

Polycentric Information Commons (PIC) theory focuses on the fundamental question of how to sustain open access systems in the face of degenerative collective-action threats that are amplified because of the systems' relatively high degree of openness (Mindel et al. 2018). The theory asserts that the integration of polycentric governance practices in the design of the rules and features of these systems may help reduce their vulnerability. Based on the logic of collective action, the theory provides an analytical framework for examining governance issues related to the sustainability and governance of open access systems. Still, PIC lacks some important nuances. While Mindel et al. (2018) assert that information commons must be sustainable to realize their overarching goal for the benefit of their stakeholders, they acknowledge that information commons differ in their goals and that further theorizing is needed to gain more nuance on the different "types" of information commons (Mindel et al. 2018).

Building on previous literature on digital activism and the blueprints of PIC theory, we theorize on information commons that seek to drive institutional change in society. As a complementary lens, we draw on institutional theory on legitimacy (Suchman 1995; Zimmerman and Zeitz 2002) and on change through collective action (Hargrave and Van de Ven 2006; Seo and Creed 2002) to propose that digital activism's success depends on its capacity to achieve appropriate sustainability in order to drive *legitimate* collective action. As basis for this theorizing, we examine a case in which actors harnessed social media and smartphone-based crowdsourcing to mobilize and drive collective action that caused institutional change.

3.2 Digital Activism

In the early 1990s, social activists were already creating websites and using email chains to pass information, solicit support, and coordinate activities (Butler 2011). Although the new medium

made some aspects of outreach and coordination easier, it did not radically change the long-existing traditional and centralized model of activism. Things began to change, however, as the internet gradually evolved from a broadcast to a participatory medium that enables private individuals to add content, organize as online communities, and connect with each other on social network platforms (Von Krogh et al. 2012). This shift let activist organizations more easily tap into large existing networks (Obar et al. 2012), as well as more easily recruit new activists, raise funds, and communicate the message to the general public.

Defined as an “organized public effort, making collective claim(s) of target authority(s), in which civic initiators or supporters use digital media” (Edwards et al. 2013. pp. 10), *digital activism* can be split into two types¹⁰: (i) digital online activism, and (ii) digital offline activism. Although these types of activism may overlap in terms of tactics and objectives, the key distinction is *spatial*. Online activism mostly occurs on the internet among dispersed individuals, while offline activism uses open access systems to mobilize people in the physical world. Both types have successful at mobilizing people for various causes, but each has its limitations.

3.2.1 Digital Online Activism

In the not-so-distant past, to raise funds and promote the change of rules, laws, or policies, activists spent long hours standing on street corners to gather signatures, going door to door to solicit support, and waiting in well-stretched lines in bureaucratic offices to obtain documents and permits. However, in recent years, open access systems significantly lowered the barriers of entry to the activism arena (Van Laer and Van Aelst 2010). Nowadays, anyone—regardless of gender,

¹⁰ At present, we do not address hacker activism, or *hacktivism*, which is a form of illicit digitally enabled online activism occurring mostly on the dark web.

race, sexual orientation, or age—can rather easily engage in promoting various causes through their devices, as long they can access the internet. Thus, instead of braving the elements outdoors for hours to gather a few dozen signatures, open access systems afford activists the opportunity to effectively reach numerous individuals anywhere.

The literature on digital online activism is split between cautious optimism regarding its potential and a dose of sober realism on its actual impact. There is consensus, however, that digital online activism makes it much easier and faster to raise awareness about various causes, collect signatures, and raise funds. Perhaps one of the most memorable recent examples of digital online activism is the 2014 “Ice Bucket Challenge”—a social-media-propelled campaign in which citizens and celebrities from around the world dumped icy water on themselves to raise awareness of amyotrophic lateral sclerosis (ALS), a rare and not very well-known neurological disease affecting human muscles. Within only a few months, the viral Ice Bucket campaign increased ALS society donations five-fold in comparison to the entire previous year (Diamond 2014). While the Ice Bucket Challenge no doubt helped the ALS cause (Gebelhoff 2015), the campaign was not universally popular; some argued that it diverted attention from other equally deserving causes (Hiltzik 2014), while others claimed that it was mostly an attention-seeking gimmick and not real activism. Approximately 17 million people posted online videos and pictures of themselves being doused in icy water, but only a tiny fraction—an estimated 4–5%—bothered to donate to ALS organizations. As one observer noted, people collectively spent more on ice and bucket purchases than on actual ALS donations (Kosinski 2014).

Still, despite the skepticism about the motivations of online activists, research suggests that involvement in such campaign leads to increased civic mindfulness and can generate other positive externalities (Kristofferson et al. 2013). For instance, researchers found that taking part in online

campaigns increases the likelihood that participants join charitable organizations (Lee and Hsieh 2013). Researchers also find that online activism gives those who usually do not take part in activism an opportunity to participate (Bonilla and Rosa 2015). Some research asserts that online activism is a useful tactical tool for promoting good causes (Karpf 2010), and that even small gestures can be powerful when combined (Christensen 2011; Vie 2014).

The growing body of research examining online activism primarily focuses on its short-term impact on campaigns and specific events (Agarwal et al. 2014; Ghobadi and Clegg 2015; Wattal et al. 2010). Although recent research has begun to examine factors that impact participation in long-term online activism (Selander and Jarvenpaa 2016), the question of whether online activism can go beyond increasing awareness, petitioning, and fundraising to bring about actual institutional change is presently unresolved (Lim 2013; Obar et al. 2012).

3.2.2 Digital Offline Activism

Open access systems have become a powerful medium for mobilizing people in the physical world (Gerbaudo 2018). Notable examples include the Anti-Globalization, Occupy Wall Street, and Black Lives Matter movements, as well as the Arab Spring protests (Carney 2016; Ghonim 2012; Gleason 2013; Juris 2005; Kelly Garrett 2006). Over the past few years, scholars of those and other digitally enabled social movements have uncovered their power and limitations.

Today, the rapid spread of messages and the sheer number of people receiving them through social networks make mobilizing masses of people unequivocally cheaper and faster than in the past (Eltantawy and Wiest 2011; Obar et al. 2012; Theocharis et al. 2015). Further, as with digitally enabled online activism, the reduced cost of coordination gives traditionally disenfranchised groups—including minorities and the economically disadvantaged—an opportunity to organize

and express themselves (Anduiza et al. 2014; Ghonim 2012). In a matter of hours, a call on social media can translate into a flash crowd of protestors occupying a physical space, be it a street, a shopping center, a highway, or an airport terminal (Doubek 2017).

Although digital offline activism may give a voice to the marginalized, such movements too often succumb to the echo-chamber effect—that is, rapidly growing discontent spreading like wildfire among people echoing to each other similar beliefs (Garrett 2009), making the movement an outlet for expression of frustration (Aouragh and Alexander 2011; Gillan 2009). Outside the echo-chamber, some people who may be privately sympathetic to the cause can grow dismayed at the aggressive tone or plainly agitated because of the disruptions caused by protests (Tankersley 2011; White 2016). To the activists in the eye of the storm who are intensely involved with the protest, the mobilization may appear revolutionary, while the rest of us may simply see an angry crowd blocking roads and making our daily commute worse than usual.

Even when movements gain widespread public support, they often fail to achieve concrete gains due to lack of consensus on the next steps following the initial mobilization (Gladwell 2010; Taub 2016; White 2016). This often results from an absence of leaders who are capable of unifying the many decentralized parts enough to articulate a plan of action that the movement can get behind (Gladwell 2010; Taub 2016). Thus, paradoxically, the major advantage of digitally enabled offline activism—that is, its decentralized emergence—becomes its Achilles heel. Too many independent centers are devoid of a common direction, and thus ineffective in achieving significant concrete gains beyond the action of mobilization itself. Leadership is essential for social movements (Morris and Staggenborg 2004), and when it is fragmented, it is very difficult having productive dialogue with established institutions (White 2016). As the events of Occupy Wall-Street illustrate, weathering the storm—that is, holding steady until protesters get tired and go home—is in effect

the default tactic that established institutions use when challenged on the streets (White 2016). Even when protests are successful at overthrowing entrenched institutions, the lack of leadership becomes a problem in the aftermath. The events of the Arab-Spring revolution in Egypt illustrate how successful digital activism mobilization failed to achieve institutional change¹¹. The lack of unified leadership on the side of the activists following the collapse of the Mubarak regime allowed entrenched political and military institutions to reclaim power (Howard and Hussain 2011; Gunitsky 2015).

So, when examining digital on- and offline activism, we see their advantages as well as their limitations in promoting institutional change. Digital online activism is too “soft”—that is, it is somewhat ineffective beyond low-risk actions (Lim 2013) such as raising awareness and funds. Moreover, some perceive it as a form of lazy *clicktivism* (Gladwell 2010). On the other hand, digital offline activism is too “hard”—that is, it is effective at punching a hole in public opinion, but too aggressive and disjointed to facilitate lasting change (White 2016).

3.3 Theoretical Framing

To better understand how digital activism may be harnessed to promote change in the institutional foundations of society, we merge two theoretical perspectives—PIC (Mindel et al. 2018) and institutional change and legitimacy (Hargrave and Van de Ven 2006; Seo and Creed 2002; Suchman 1995)—to offer a detailed case study analysis of *Cabotagestudien* (CS). Using collective action as a shared perspective that merges these two theoretical lenses, we were able to examine

¹¹ Calhoun (2013) notes that it is too early to judge the true long-term impact of offline activism movements, but even some of the prominent figures leading Occupy Wall-Street and the Arab Spring concede failure due to the lack of coordinated leadership in the aftermath of mobilization (White 2016, 2017).

this case with a focus on the mutually constitutive duality (Feldman and Orlikowski 2011; Jones et al. 2004) between the digital activism and the institutional context it aims to change.

3.3.1 Digital Activism as Polycentric Information Commons

PIC theory (Mindel et al. 2018) examines open access systems online that integrate user-generated content and information as “common grounds.” According to this lens, social media and other open access systems online that enable digital activism may be conceptualized as “information commons,” that is, as manifestations of collective action taking place in cyberspace. This is a departure from prior research on digital activism, which primarily examines the enabling online systems as communication tools (Harlow and Harp 2012; Obar et al. 2012; Segerberg and Bennett 2011).

When examining the use of new technologies to promote change, digital activism literature broadly distinguishes between two types of actors: institutional and non-institutional (Bimber et al. 2012; Selander and Jarvenpaa 2016). This distinction works well when examining the dynamics between a single, identifiable organization (an institutional actor) and everyone else outside of this organization (non-institutional actors). The challenge, however, is that activism often involves multiple organizations and actors entangled as part of an “institutional field,” which obscures the institutional/non-institutional dichotomy. In the dichotomous view, for example, a social movement organization dedicated to a cause is an institutional actor, while the individuals accessing the organization’s social media page to make comments or acquire information are non-institutional actors. However, people accessing the organization’s information may be politicians, reporters, business owners, and other actors who belong to various organizations and whose individual actions impact institutional change. To avoid the confusion, we appropriate PIC

theory's distinction between three types of actors relevant to digital activism as follows: (1) *producers*: the architects and organizers of the digital activism platform; (2) *providers*: the people who supply information by posting content; and (3) *appropriators*: the people who acquire information for various reasons (Mindel et al. 2018).

While collective action is typically viewed as a generative mechanism (Hargrave and Van de Ven 2006), PIC theory sees it as potentially degenerative to the point that it might undermine the system's sustainability (Mindel et al. 2018). In the same way that commons in the physical world face sustainability pressures as the aggregate result of smaller-scale actions (Ostrom 1990), information commons in digital activism are vulnerable because of their high degree of open access, widespread anonymity, and lack of hierarchical control (Mindel et al. 2018). Thus, as with common pool physical resources such as forests or grazing grounds, these information commons are likely to experience pressure due to pollution, congestion, violations, and rebellion.

Low barriers for participation in digital activism increase the odds that a certain portion of the uploaded content may be polluted by inaccurate or even deliberately manipulated information. *Information pollution*—defined as “the extent to which information is contaminated and fails to align with the information commons overarching goal” (Mindel et al. 2018, p. 617)—is a problem for platforms that invite users to contribute content and information (Gu et al. 2007; Hu et al. 2012; Kittur et al. 2008). Further, prior research finds that the spread of misinformation online reinforces echo chamber effects, which can lead to increased polarization and other negative spillover effects (Del Vicario et al. 2016; Kata 2012; Silverman and Singer-Vine 2016).

Digital activism can also suffer from activity spikes that jam the system. Such *congestion*—that is, “the extent to which appropriation and provision clog the information commons” (Mindel et al. 2018, p. 616)—may cause the system to crash and create information overload. Digital activism

research finds that simultaneous uploads of posts can create messy feeds and cause confusion and information overload, which in turn may lead to fatigue among activists (Branagan 2013; Kelly Garrett 2006). Frequent crashes caused by congestion may frustrate existing and potential activists alike; moreover, it may prevent important messages from spreading, subsequently muffling the potential impact of digital activism.

Because digital activism that touches on political issues is likely to be controversial and heated (Lee 2005), it can be a fertile ground for harassment, cyberbullying, hate speech, and trolling (Li 2005; Warzel 2016). Defined as “the extent to which providers in an information commons violate rules and laws” (Mindel et al. 2018, p. 617), *violations* such as harassment may cause some activists to quit and may even attract pressure from outside actors. Although digital activism literature does not address harassment specifically, plenty of anecdotal evidence suggests that it is indeed a problem (Larkin 2016).

In digital activism, rifts between producers and providers may cause the latter to desert or rebel. *Rebellion*—that is, “the extent to which providers in an information commons are dissatisfied and exit it” (Mindel et al. 2018, p. 617)—jeopardizes the supply of content (Clifford 2010; Graham-Felsen 2006; Gross 2012). So, while digital activism can grow exponentially because of low participation barriers, it can also experience rapid participation decline. Donor reactions to reports that Red Cross stashed money donated to help the 2010 Haiti earthquake victims (Sullivan 2015) exemplifies how rebellion can adversely affect digital activism.

To reduce the prevalence and impact of collective-action threats in digital activism, PIC theory proposes four mutually reinforcing governance practices that may help actors balance openness and order to achieve stability. These practices (Table 3.1) may be embedded in the system’s

technical features and design, as well as in its official and unofficial rules and norms of conduct (Mindel et al. 2018).

Concept	Definition	Construct	Definition	References
<i>Collective-action threats</i>	Adverse aggregate effects that are caused by individual-level actions during digital activism	Pollution	The extent to which information is contaminated and fails to align with the overarching goal of digital activism	Kata 2012; Kelly Garrett 2006; Lampel and Bhalla 2007; Mindel et al. 2018; Stien 2016
		Congestion	The extent to which appropriation and provision of information clog the system that enables digital activism	
		Violations	The extent to which providers of information violate rules and laws related to digital activism	
		Rebellion	The extent to which providers of information are dissatisfied and exit a digital activism effort	
		Boundary regulation	The extent to which rules and technical infrastructure features regulate anonymity and content during digital activism	
<i>Polycentric governance practices</i>	Digital activism practices that promote order through independent adjustments as part of an emergent system of rules	Incremental adaptation	The extent to which actors gradually introduce changes in infrastructure features and rules during digital activism	Mindel et al. 2018; Ostrom 1990
		Shared accountability	The extent to which rules and features afford peer monitoring and gradual sanctioning of violators of rules and norms during digital activism	
		Provider recognition	The extent to which providers of information are acknowledged by peers, appropriators, and producers during digital activism	

Boundaries that align with digital activism’s overarching goal reduce the prevalence and impact of collective-action threats. *Boundary regulation*—defined as “the extent to which rules and technical infrastructure features of the information commons regulate anonymity and content” (Mindel et al. 2018, p. 619)—may therefore help reduce collective-action threats in information commons (Butler and Wang 2012; Di Tullio and Staples 2013; Hertel et al. 2003; Jarvenpaa and

Lang 2011; Ren et al. 2007) and thereby improve digital activism outcomes. For example, rules and design features, such as login requirements, can reduce the adverse effect of anonymity, while boundaries on content can prevent the discourse from getting overly heated.

Information providers and appropriators are sensitive to changes they perceive as too sudden or radical. Given this, *incremental adaptation*—that is, “the extent to which changes in infrastructure features and rules are gradually introduced” (Mindel et al. 2018, p. 619)—may reduce the likelihood of alienating digital activism participants and, as a consequence, reduce instances of rebellion and violations. PIC theory would suggest that a digital activism effort is more likely to be sustained when it involves information providers in governance and incorporates community feedback into an incremental adaptation of the system’s rules and features.

It is important that participants have a sense of collective responsibility toward digital activism efforts. This self-emerging *shared accountability*—defined as “the extent to which rules and features afford peer monitoring and gradual sanctioning of violators of rules and norms of the information commons” (Mindel et al. 2018, p. 619)—can improve outcomes in various open access systems (Chua et al. 2007; Wise et al. 2006). Consistent with this insight, we suggest that actors adopt shared accountability and peer monitoring to reduce collective-action threats to their digital activism efforts.

The more content providers feel recognized for their efforts by other stakeholders, the more likely they are to continue participating and contributing (Lampel and Bhalla 2007). Given this, *provider recognition*—defined as “the extent to which providers are acknowledged by peers, appropriators and producers” (Mindel et al. 2018, p. 619)—is an important source of motivation (Von Krogh et al. 2012). Recognition may come as direct congratulatory messages and posts, or as badges, icons,

“likes,” and other features that tell providers that their efforts are noticed. Provider recognition is an important driver of continued engagement during digital activism, which often succumbs to drops in engagement level (Van Laer and Van Aelst 2010).

We build on these PIC theory concepts (Table 3.1) to empirically examine and theorize about how digital activism may be harnessed for institutional change. At the same time, we observe that PIC theory is generally concerned about the sustainability of information commons, whereas digital activism’s primary objective is to drive action and achieve some form of change. Given this, we suggest that digital activists not only sustain the information commons but also collectively achieve a legitimacy threshold (Zimmerman and Zeitz 2002) as a basis for maneuvering other partisan actors and the complex institutional landscape in which they operate.

3.3.2 Digital Activism for Institutional Change

Institutions—that is, the significant practices, relationships, or organizations in a society¹²—implicitly and explicitly impact most aspects of life. Hence, how institutions are altered is of great interest to social scientists from multiple disciplines, including management research (Hargrave and Van de Ven 2006; Seo and Creed 2002), political science (Streeck and Thelen 2009), and information systems research (Mignerat and Rivard 2009; Orlikowski and Barley 2001). *Institutional change*—defined as “a difference in form, quality, or state over time in an institution” (Hargrave and Van de Ven 2006, p. 866)—can be manifested on the regulatory level, the practice level, or both. As Table 3.2 shows, the *regulatory* level appertains to the formal laws, rules, and policies that govern interactions among institutional actors, groups, and individuals, while the *practice* level appertains to the informal behavioral norms of conduct.

¹² Merriam-Webster

As partisan actors harness digital activism to change regulations and practices in a specific institution, they must maneuver the broader *institutional field* in which entrenched institutions are likely to oppose the changes to preserve the status quo. For that reason, it is very difficult for actors to spur institutional change on their own. On the other hand, a collective action in which multiple actors maneuver simultaneously, even if not in unison, is less risky for the involved individuals and has a better chance to succeed (Hargrave and Van de Ven 2006). As outlined above, examining the role of digital activism in collective action, institutional researchers have found that it is an effective mechanism to reach and mobilize great numbers of individuals, but also that the intended societal impact is rarely achieved or sustained.

Institutional theories posit that *legitimacy* is the key means by which institutional actors obtain and maintain resources and ultimately exert influence (Oliver 1991; Zimmerman and Zeitz 2002). Accordingly, we suggest that the muffled impact of digital online and offline activism is largely the result of weak legitimacy. Although legitimacy has many definitions (Deephouse and Suchman 2008), it generally concerns an institutional actor's congruence with social laws, norms, and values; an institutional actor is legitimate within its field if its actions are recognized as legal, aligned with existing practices, and socially acceptable to its stakeholders and related institutional actors. In his seminal work, Suchman (1995) identifies three broad types of legitimacy: pragmatic, moral, and cognitive.

Without the continuous support of its most immediate stakeholders, an institutional actor is unlikely to survive. For that reason, institutional actors must secure *pragmatic legitimacy* by providing value to their immediate stakeholders (Suchman 1995). Pragmatic legitimacy is based on a symbiotic dependency between an actor's actions and the interests of an institution's immediate stakeholders; it is the most basic form of legitimacy that must be achieved. In the

context of digital activism, the campaign—whether it aims to raise funds, collect signatures, increase awareness, or get people on the streets—must first and foremost gain initial support from first adopters. Once it gains traction, network effects (Ackland and O’neil 2011) propel further growth in participation, thereby stabilizing pragmatic legitimacy.

To gain and maintain legitimacy, the actions of institutional actors must be generally perceived as ethical by stakeholders and other institutional actors (Suchman 1995). As such, *moral legitimacy* gives institutional actors the credence required to shore up support and fend off potential challengers. Digital activism that is perceived to be “slacktivism” or “too aggressive” is less likely to secure broad moral legitimacy.

Table 3.2 Institutional Theory Applied to Digital Activism

Concept	Definition	Construct	Definition	References
<i>Institutional change</i>	A difference in form, quality, or state over time in the institution targeted by digital activism	Institutional regulations	Formal laws, rules, and policies that govern the behaviors and actions within the institution targeted by digital activism	Hargrave & Van de Ven 2006
		Institutional practices	Informal behavioral norms of conduct within the institution targeted by digital activism	
<i>Legitimate action</i>	An action that is desirable, proper, or appropriate within an institution of regulations and practices related to digital activism	Pragmatic legitimacy	The extent to which the actions of an actor are perceived to enhance the interests of an institution’s immediate stakeholders	Suchman 1995
		Moral legitimacy	The extent to which the actions of an actor are perceived to be the “right thing to do” within an institution	
		Cognitive legitimacy	The extent to which the actions of an actor are perceived to be comprehensible, factually valid, and “inevitable” within an institution	

Institutional actors’ activities also must be perceived as comprehensible and rational by stakeholders and other institutional actors in order to be “taken-for-granted” (Suchman 1995). This high standard of *cognitive legitimacy* is not based on instinctive self-interest (pragmatic legitimacy) or subjective evaluation (moral legitimacy), but rather on a rational consideration that

an institutional action is inevitable because it is superior to other alternatives. As such, cognitive legitimacy is the highest form of legitimacy and the most difficult to establish.

3.4 Research Method

Using the *explanatory case study approach* (Benbasat et al. 1987; Yin 2008), we examine CS, a research collaboration between a logistics researcher from Lund University in Southern Sweden and the Swedish Transportations Workers Union, that turned into a digital activism campaign with widespread ramifications. CS is a unique contemporary case in which institutional actors effectively leveraged on- and offline digital activism—through social media and smartphone-based crowdsourcing—to change road haulage regulations and practices in Northern Europe. Through data crowdsourcing via smartphones, CS generated evidence on existing EU legislations’ adverse effects to prompt political, legislative, and industry changes far beyond what traditional activism had achieved in the preceding 15 years. As such, CS offers important practical and theoretical lessons on digital activism’s potential to promote institutional change.

We gathered case study data from three primary sources: interviews with key actors, electronic records of communication among actors, and third-party reports. Interviews are essential information sources for case study research as they provide the individual perspectives of key actors in their own words (Yin 2008). Following the guidelines of Eisenhardt and Graebner (2007) for case-study-based theory development, we sought a plurality of perspectives and interviewed the actors who had the most knowledge of CS and its impact. Plurality of perspectives reduced the narrative bias that might potentially impact subsequent theorization. As Table 3.3 shows, we conducted 18 semi-structured interviews with key stakeholders: the principal producer of CS who initiated the crowdsourcing study; the mathematician who analyzed the crowdsourced data; the union leaders who financially supported the study and used the findings to push for institutional

change; several truck drivers who were highly involved in data collection; a high-ranking politician who used the crowdsourced data for campaigning and later for promoting new legislation; a manager of a large retail transportation fleet who instituted changes to practice based on the crowdsourced findings; CS supporters and the legal counsel for an opposition interest group who lobbied against CS.

Table 3.3 Interviews

18 Semi-structured Interviews (in-person, unless indicated by "T" for telephone interviews)

#	Key Informants	Position	# Interviews
1	University Researcher	Principal producer of CS	3
2	Mathematician	Developer of CS data triangulation algorithm	1(T)
3	Union Leader 1	The former head of the Swedish Transport Union	1
4	Union Leader 2	Assistant to the union leader in charge of operations	1
5	Union Leader 3	Assistant to the union leader in charge of lobbying	1
6	Driver 1	Leader of the protest movement "Sweden Comes to Halt"	1
7	Driver 2	Moderator of the CS Facebook page	1
8	Driver 3	High-volume data collector for CS	1
9	Driver 4	High-volume data collector for CS	1
10	Driver 5	Opinion leader in the online trucking community	1
11	Driver 6	Opinion leader in the online trucking community	1(T)
12	High-Ranking Politician	Former minister who used CS to promote new legislation	1
13	Industry Leader	Transport boss of a large Swedish retailer	1(T)
14	Legal Council	Lawyer of a large logistics association (chief opponent of CS)	1
15	Company Owner	Retired haulage company owner (supporter of CS)	1
16	Whistle-blower	Former owner of a haulage firm (provided hacked data)	1(T)

Social media posts and communiqués are increasingly used as data sources in case study research (Wattal et al. 2010; Yates and Paquette 2011) as they provide a wealth of mostly unfiltered insights into the views, opinions, and conversational dynamics of actors in real-time. The electronic time stamping also helps create a chronology that lets researchers trace changes over time. We reviewed a total of 95 email messages, 47 communication streams on Facebook instant messaging, and 439 Facebook posts to help develop a more comprehensive account of our case. Case study researchers also often use third-party reports to obtain an outsider perspective on the events in question (Myers

1997). We therefore collected a total of 621 media reports on the crowdsourcing project, as well as 7 official reports related to the case.

In accordance with Eisenhardt (1989), we aimed to develop a theory-informed narrative for the case study. We followed prior case study research to distill insights from the multiple qualitative data sources into a cohesive interpretation by analyzing key episodes and events (Newman and Robey 1992; Lyytinen and Newman 2008). We used temporal bracketing to organize the key events into five thematically cohesive episodes: episode one (September 2012–March 2013) focuses on the events related to the instigation of institutional change; episode two (April 2013–July 2013) focuses on the events related to the emerging collective action; episode three (May–September 2013) centers on the enactment of polycentric governance practices; episode four (May 2013–March 2014) examines how the collective action further cascaded; and episode five (November 2014–February 2015) focuses on how information-provider rebellion was avoided. To provide context for the five episodes, we first described the intuitional field and summarized the resulting institutional changes on both the regulation and practice levels.

To develop our inductive theory, we followed the recommendations by Eisenhardt (1989) and Eisenhardt and Graebner (2007). First, we used the theoretical framing (Tables 3.2 and 3.3) to identify collective-action threats, polycentric governance practices, legitimate digital activism, and institutional changes across the various episodes. The first author analyzed all of the material based on the theoretical framework's concepts and constructs, while the second and third authors reviewed the analyses results. Based on their feedback, the first author revisited the data and, when needed, conducted follow-up interviews with the CS's principle producer. Following three iterations of this analysis, critical review, data review, and follow-up interview, we achieved the analysis provided below. Next, relying on analytical generalization (Lee and Baskerville 2003;

Yin 2008), we used the analysis in conjunction with existing literature to develop a series of propositions; we then tied these propositions together in a conceptual model that theorizes on the relationship between legitimate digital activism, institutional change, collective-action threats, institutional resistance, and polycentric governance.

3.5 Case Study Analysis

Here, using PIC and institutional theories to frame the analysis, we provide a detailed account of the case study (summarized in Table 3.4). We begin by describing the institutional field in which the digital activism operated and conclude with the resulting regulatory and practice changes in the targeted road haulage institutions in Northern Europe.

3.5.1 Institutional Field

In the 1990s, The European Union (EU) integration accelerated and debates over the impact of opening borders to newly admitted Eastern European member countries began to dominate European politics (Richardson and Mazey 2015; Schimmelfennig 2001). One such debate focuses on the impact of EU Regulation directive 1072/2009 (2009), which concerns domestic road haulage regulations. To reduce CO₂ emissions, Regulation 1072/2009 grants the right to long-distance international road haulage operators to stay in a country for up to a week and carry up to three domestic loads. This practice is known as *cabotage*.

Regulation 1072/2009's intention was to increase the fill-rate of large trucks so that they could be used economically and not travel empty. The regulation was hailed as a good example of how environmental mindfulness can be applied in a practical manner. Indeed, data submitted by member state statistics bureaus to Eurostat¹³ showed that, following the regulation's enactment,

¹³ Eurostat is the Luxemburg-based EU administration branch dedicated to compiling statistical data.

fewer empty trucks were operating at any given moment across the EU. The policy's success prompted EU legislators to schedule further cross-border haulage deregulations.

However, as Regulation 1072/2009 took effect, haulers from newly admitted Eastern EU member states began inundating the domestic markets of Western European countries (Hilal 2008; Kummer et al. 2014). Reports emerged of widespread exploitation of cheaper Eastern European drivers by haulage companies (Hilal 2008). Anecdotal reports also suggested that cabotage regulations—that is, the provisions limiting local carries to a maximum of three per week—were largely ignored, and that underpaid Eastern European drivers were driving around with empty containers in the hope of getting assignments. A retired haulage company owner explained to us the chief reason that haulage companies were ignoring the cabotage rules: *“I [and other haulage companies] could get four or five Romanian drivers for the price of one Swedish driver.”* On top of that, he said that *“nobody was enforcing the rules”* limiting cabotage to three assignments per week. The availability of cheap labor launched a race to the bottom, driving haulage companies to let go of West European drivers while recruiting more East Europeans to compete for transportation assignment bids. The big-box retailers and other haulage customers, who benefitted from lower transportation costs, simply turned their heads, according to the retired haulage company owner. So, in effect, EU cross-border transport deregulation gave rise to institutional exploitation of East European drivers, harmed rather than improved environmental sustainability, and disrupted the livelihood of West European drivers.

3.5.2 Instigating Institutional Change (September 2012–March 2013)

Witnessing its members hurting, the Swedish Transportation Union pushed the first domino to challenge the status quo. Established in 1897, the Union is tasked with protecting the rights and

interest of its 60,000 transport industry members. To maintain pragmatic legitimacy in the eyes of its stakeholders, the Union engages in different activities, including negotiating collective bargaining agreements on working conditions with various commercial and governmental organizations, and lobbying for or against policies impacting its members. Naturally, the Union opposed Regulation 1072/2009 prior to its enactment and continued arguing against it, but without success. In political circles, Regulation 1072/2009 was seen as environmentally friendly, economically efficient, and undisruptive. Although the Union had considerable anecdotal evidence on widespread disregard for cabotage laws and exploitation of Eastern European drivers, the evidence was largely dismissed by the political establishment as merely isolated incidents and rumors. As one Union leader explained it to us: *“We needed real data and statistics to get the attention of policy makers.”*

To challenge the status quo, the Union decided to sponsor scientific research that could contradict or at least seriously question Eurostat’s data to show that Regulation 1072/2009 was not nearly as good for the environment. The Union leadership contacted a logistics researcher and asked him to examine Regulation 1072/2009’s environmental impact. The researcher first contacted various haulage companies and retailers with their own transportation fleets to request access to their internal data logs on transportation assignments. Not a single company agreed to release its data, however. The researcher and his team then began manually recording the movements of trucks in and out of Gothenburg port, the major entry point of foreign trucks into Sweden. However, while this data collection approach showed a disproportionate traffic volume of Eastern European trucks, it said nothing about their environmental impact.

To gather the requisite empirical observations, the researcher considered crowdsourcing techniques. Although relying on crowds of volunteers to collect data had been attempted in other

research contexts—including emergency response, noise pollution mapping, medical research, and map making (Goodchild and Glennon 2010; Heipke 2010; Stevens and D’Hondt 2010)—it had not been used to map moving targets on such a vast geographical scale. To facilitate this mapping, the research team designed a simple smartphone application that let users record truck license plates and automatically extract the time-stamp and coordinates from the phone’s GPS.

At this point, the team’s main challenge was to legitimize the data crowdsourcing to motivate users to participate. The Union issued an announcement through its webpage, newsletter, and official Facebook page with the goal of getting its members to download the app and participate in data collection. Initially, the call was largely ignored, and hardly anyone downloaded the app. Moreover, several truckers prominent on the Facebook community page “Trucker’s Paradise”—with more than 12,000 members, most of whom are Swedish truck drivers—openly ridiculed the idea, questioning both its practicality and usefulness. As one truck driver put it: *“The entire thing seems like a waste of time.”*

The researcher, henceforth referred to as the “producer” of the information commons¹⁴, personally contacted every driver who expressed skepticism about the data crowdsourcing initiative. Using Facebook messenger as his main tool of communication, he urged the skeptics to participate by empathetically addressing the basic collective need of Swedish truck drivers: to secure their livelihood. As the producer noted to a skeptic truck driver: *“You have nothing to lose and a lot to gain.”* Although still hesitant, the drivers that he contacted gave their word that they would participate in data collection and that they would urge other truckers to join in as well.

¹⁴ The entity leading the information commons—setting its overarching goal and organizing its rules and infrastructure, called a “producer” in PIC theory.

To coordinate the data collection efforts, the producer created a new Facebook group page titled “Cabotagestudien.” Although the study’s stated goal was to examine the environmental and social impact of Regulation 1072/2009, the chosen name of the project and the associated Facebook page purposefully zoomed in on the issue of cabotage—specifically, the practice of local carries by East European trucks in the Nordic countries. The producer and his team reasoned that it was easier to rouse the crowd to participate in data collection if it spoke to their basic concerns about the deteriorating conditions in the industry, job security and accurate statistics rather than to an abstract notion of environmentalism.

In summary, the implicit rebellion, on social media, of truck drivers who initially refused to partake in data crowdsourcing was threatening the pragmatic legitimacy of the digital offline activism (Table 3.4). However, through direct recognition of key opinion leaders in the trucking community, the producer secured initial pragmatic legitimacy to support CS. The resulting partnership between the university researcher (producer), the Transport Union leadership, and the trucking community’s online opinion leader signaled an initial shift in the institutional field.

3.5.3 Emerging Collective Action (April 2013–July 2013)

The Union’s engagement of the producer to obtain data that could be used to challenge Regulation 1072/2009 suddenly expanded into a collective action. Hundreds of truck drivers from rival factions—union members, anti-union private trucking group members, and independent operators alike—downloaded the app and used it to tag the license plates of trucks. As one non-union member truck driver put it when committing to the data crowdsourcing: *“I hate the Union, but I will tag.”*

The volume of participation and the stream of truck observations steadily grew over the six weeks of the first crowdsourcing run, providing thousands of data points used to triangulate truck movement (Figure 3.1). For the first time ever, an observation-based account of truck movement was available. As Figure 3.2 shows, the producer and his team created a website to show information appropriators the movement patterns of trucks on an interactive map.

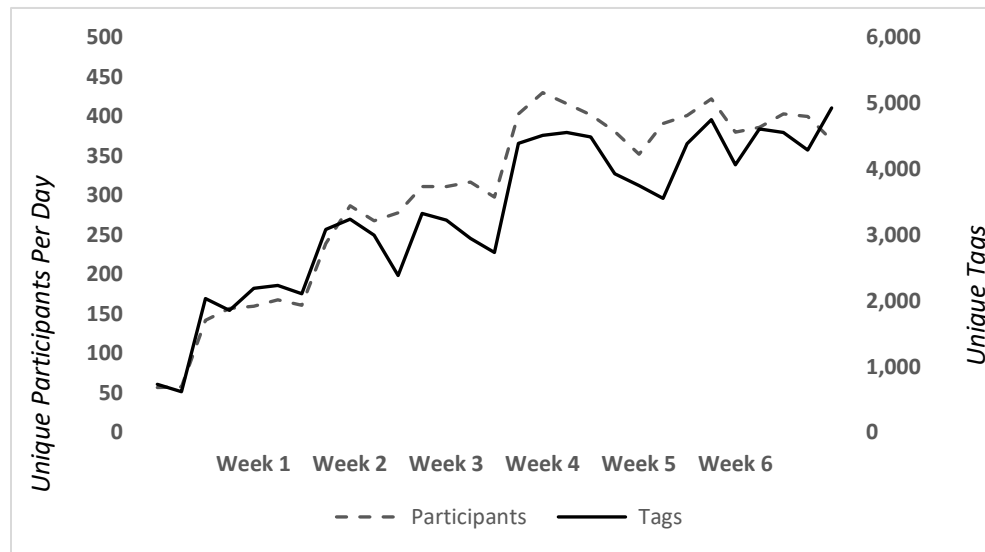


Figure 3.1 First Crowdsourcing Run, April–June 2013

The observed movement patterns largely confirmed the suspicion that East European trucks did not exit the country as required, but instead drove around scrambling for local assignments. Moreover, the observations suggested that East European trucks were often traveling without cargo. In addition to tagging trucks, some drivers began recording and posting videos showing the horrid unsanitary living conditions of East European drivers in their makeshift camps. In just six weeks of data crowdsourcing and on t=other on the ground activism, CS could provide evidence of the adverse consequences of Regulation 1072/2009: rampant breaking of cabotage practices and widespread marginalization of East European drivers.

The openness of the information commons—that is, the study’s Facebook page and the crowdsourcing app—almost immediately gave rise to collective-action threats. After maps of truck movement patterns were made accessible to the public online, increased internet traffic caused congestion in appropriation and frequent crashes of the website. More concerning, the production team began noticing an increase in nonsensical observations submitted to the data pool. While some of that information pollution was attributed to simple human error in punching numbers, the appearance of observations of trucks on open water and other suspicious data suggested that the app also was being deliberately abused. It was unclear what proportion of this deliberate data pollution was the result of juvenile vandalism and what was sabotage by those who opposed the CS objectives, but we do know that the latter occurred. For example, we interviewed the legal counsel for an organization supporting Regulation 1072/2009, and he openly bragged about tagging the license plates of trucks he knew for a fact were on assignments in other parts of Europe. The concern at the time was that data pollution might compromise the accuracy of the mathematical triangulation used to map truck movements, subsequently hurting CS’s cognitive legitimacy.

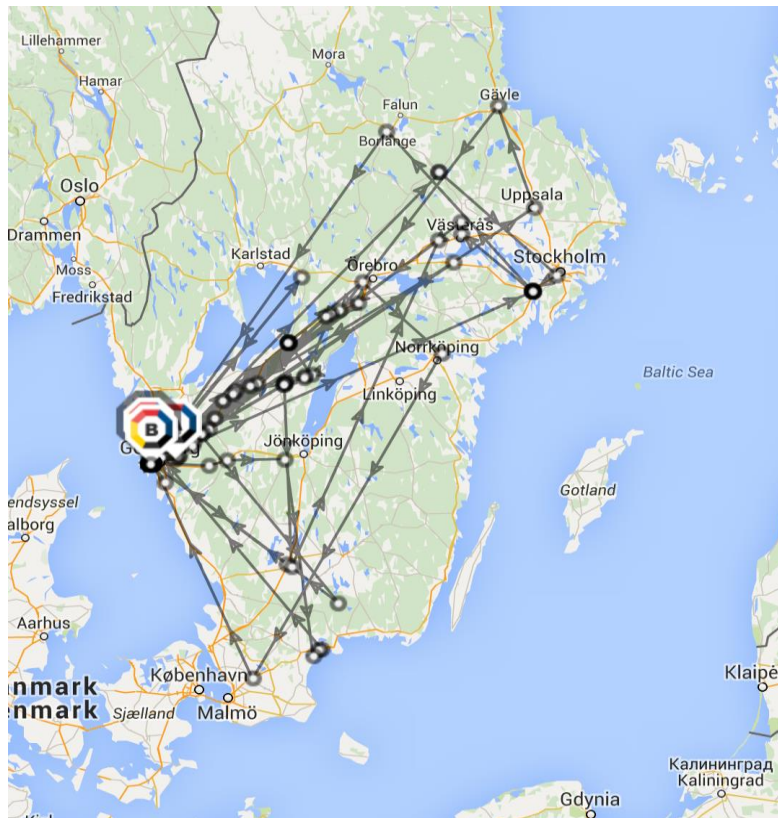


Figure 3.2 Example Movement Patterns of a Truck Violating Cabotage Regulations

In addition to pollution, violations began emerging, threatening the digital activism’s moral legitimacy. After the findings became public and exposed the scope of the problems that Regulation 1072/2009 caused, an increasing number of frustrated Swedish truck drivers began openly expressing vicious—and at times racist and xenophobic—sentiments toward the East European drivers on the study’s Facebook page. One truck driver, for example, posted the angry message: *“The [expletive] Polish drivers are stealing our jobs!”* Although the East European drivers were themselves often victims of institutional exploitation (Hilal 2008), some Swedish drivers perceived them to be the problem. In our interviews, the principle producer, Union leaders, and high-ranking politicians all expressed concern that such antagonistic rhetoric could alienate public opinion and other political allies.

In summary, collective action emerged as an increasing number of truckers and others downloaded the data-crowdsourcing app (Table 3.4). In response to growing interest in the digital activism, the producers launched a website that showed visitors the preliminary results of the data analysis on an interactive map; however, the website frequently crashed because of congestion in appropriation. Further, as PIC theory predicts, other collective-action threats—including violations (in the form of xenophobic comments) and pollution (in the form of invalid tags)—began emerging.

3.5.4 Enacting Polycentric Governance (May 2013–September 2013)

The first version of the app and the associated website were rather rudimentary, and the production team realized that several incremental adaptations were needed to respond to the collective-action threats following the first data crowdsourcing run. To deal with the study website’s frequent crashes, the principle producer acquired additional server capacity to increase data availability to the appropriators. At first, most of the people accessing the website were the Swedish drivers involved in data collection, but other institutional actors—including the media, politicians, and various transport industry actors—quickly began accessing the findings as well.

To combat data pollution, the production team took several steps. First, it devised detailed protocols for cleaning data prior to analysis. Second, the team required app users to register. The team was initially concerned that eliminating anonymity might reduce participation. However, after deliberations that included feedback from core information providers, the team decided that data accuracy (cognitive legitimacy) was more important than user anonymity. Rather than block polluters, the CS mathematician used the polluted data to further calibrate the triangulation algorithm. As the mathematician noted, paradoxically, “*knowing which tags were deliberately*

manipulated made it easier to detect other unintentionally polluted observations and thus focus the analysis on the untainted data.”

To combat violations on the Facebook page, the production team and other participants tried to post conciliatory messages to urge people to focus on the data collection and avoid xenophobic posts. For example, one truck driver commented on another trucker’s xenophobic post on the study Facebook page as follows: *“Please stop with the hate posts. They [the East European drivers] are not the problem.”* However well intended, those calls for civility were largely ignored. The principle producer then contacted core providers to discuss the harmful impact of xenophobic rhetoric on public perception of the study. Some providers volunteered to act as peer reviewers, and norms of shared accountability began emerging as these fellow drivers systematically urged peers to avoid using derogatory language. Subsequently, though not eliminated, the volume and viciousness level of xenophobic posts dropped substantially.

To encourage the crowd, the production team also started to post updates on the data collection effort’s progress, along with links to media stories, on the Facebook page. More than hundred such posts—ranging from two to five paragraphs long—were made over a two-year period. Additional recognition came from the Scandinavian trucker magazines as well as other supporters, who began offering modest prizes to top information providers. In addition, the producer team created a ranking system that included medals and other achievement icons to recognize top data providers. The principle producer also personally contacted top data providers to commend their effort. Although, initially, concerns about the future of their livelihood was the main motivation for truck drivers to participate, the various forms of recognition helped to prevent the flame from sizzling and strengthened the needed pragmatic legitimacy.

In summary, following the emergence of collective action and the collective-action threats, CS producers enacted a series of polycentric responses: they incrementally adapted server capacity, set a boundary on anonymity, recruited responsible activists to act as Facebook page moderators, and added a rating system to recognize top data taggers. These improvements helped to increase the moral and cognitive legitimacy of the digital activism.

3.5.5 Cascading Collective Action (May 2013–March 2014)

Following the release of the crowdsourced data, multiple dormant actors in the institutional field began taking steps to protect and enhance their own legitimacy. Although such maneuvers were largely uncoordinated, they nonetheless culminated in an expanding collective action.

During the first data crowdsourcing run in May 2013, CS received increasing interest from politicians and the media. For example, the Swedish infrastructure minister stated: *“We need to investigate the matter [cabotage] to address the issues. I myself downloaded the application and used it.”* The minister also called for the creation of a committee to reexamine Regulation 1072/2009’s impact. At the same time, reports on CS began appearing on national news coverage, including a featured story on the primetime evening news broadcast. Further, politicians opposing rapid EU expansion began citing CS to bolster their legitimacy with potential backers and voters during the 2014 Swedish Parliamentary elections. The high-ranking politician we interviewed stated that CS provided an *“important window to the larger issue of the impact of EU rapid expansion on blue-collar labor in Sweden.”*

As the publicity of CS grew, major transport buyers began paying greater attention to their transportation sourcing practices, and some even allied themselves with CS—presumably to bolster their own legitimacy. For instance, following the first data crowdsourcing run, one of

Sweden's largest retailers released its data logs to the production team. This data was instrumental for determining the environmental impact of the transportation fleet, which further bolstered the cognitive legitimacy of CS. In our interview, the transportation boss of the retailer praised CS for *“raising the awareness of the issue of exploitation of East European drivers and rampant law breaking.”* Public opinion pressure following the explosion of the issue in the media and political arena likely pressured the retailer to release the data logs.

In September 2013, following CS's initial success, the Union provided additional funding. When a second data crowdsourcing run was launched in October 2013, the response from the crowd was even stronger, and for the first time CS surpassed a thousand daily app users (Figure 3.3). This time around, it took only four weeks to gather enough observations to gain meaningful triangulation of movement patterns. The second data crowdsourcing run's success spurred additional media coverage and political posturing, and some haulers began seeing the value in being associated with the increasingly high-profile study. Three haulers opened their data logs to the producer's team. In exchange for the data, the haulers asked for a certification for being honest players, presumably to improve their legitimacy. Gaining access to the haulers' data further increased the cognitive legitimacy of the crowdsourced data; the production team could now clearly show that the crowdsourced observation closely aligned with haulers GPS logs. Even the mathematician was somewhat surprised by how well the data was aligned.

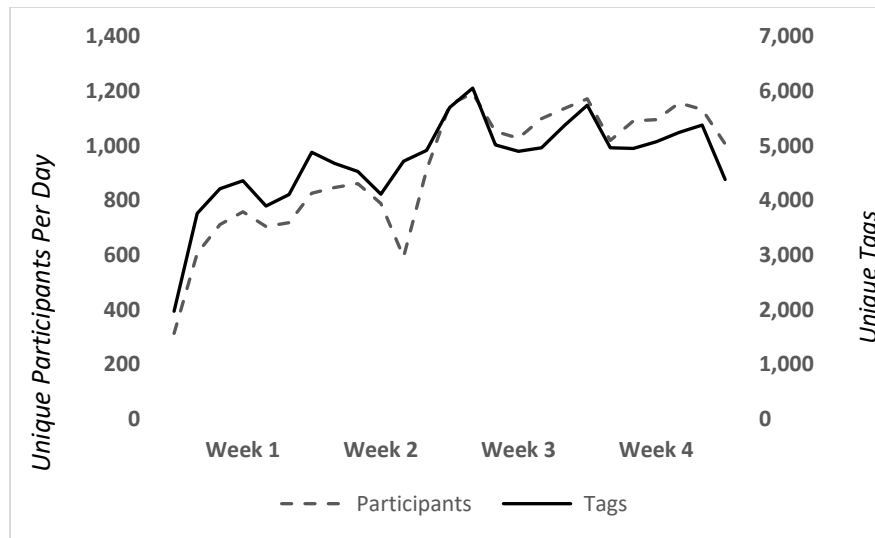


Figure 3.3 Second Crowdsourcing Run, October–November 2013

While the attention CS received helped enhance its legitimacy and impact in the institutional field, it also began attracting attacks from institutional actors who supported Regulation 1072/2009 as well as further deregulations. Such opponents often criticized the cognitive legitimacy of the data collection method by questioning the notions that isolated field observations could be used to map truck movements across a distance or that the CS algorithm could say anything about the environmental impact of East European trucks. Opponents also questioned the moral legitimacy of CS by repeatedly pointing to the xenophobic sentiments it was steering during the election season. Detractors opposed CS for various reasons. The former legal counsel for a major Forwarders' Association openly admitted opposing CS initially because it reduced his client's ability to be "flexible" with transportation labor costs. He also cited various other legal and ideological objections to CS.

In summary, the digital activism's success motivated attacks on its cognitive legitimacy from established interests (Table 3.4). However, early success also expanded shared accountability

across the institutional field, and other actors supplied additional data that helped to stabilize the cognitive legitimacy of the digital activism.

3.5.6 Avoiding Provider Rebellion (November 2014–December 2014)

When an East European trucker was arrested for driving with fake documents and released with no charges in the end of 2014, frustrations with the status-quo among Swedish truckers reached a boiling point. Some truckers began discussing roadside protests across the country, including calling for synchronized 15-minute traffic blockades on a newly created Facebook page (Figure 3.4). Many of the participants in this new protest movement—called *HÄR STANNAR SVERIGE!* (Sweden comes to a halt!)—were information providers in the crowdsourcing project who met through the CS Facebook page. One trucker urged his peers to engage in protests on the newly created Facebook page: “*Enough with tagging [data collection]; it is time for action.*”

Now that the initial objective—revealing the movement patterns of foreign trucks—was achieved, perceptions of CS’s overarching goal began to diverge. The producer’s team focused on further validating the findings and analyzing their implications. At the same time, an increasing proportion of information providers became impatient with the pace of change. After creating its own Facebook page, the new movement launched its own application to coordinate and trace roadside protests.



Figure 3.4 The “Sweden Comes to a Halt!” Movement Called on West European Truck Drivers to Stop for 15 Minutes, twice ad, for a Month to Protest Cabotage Laws

The producer was concerned that the aggressive tactics that disrupt traffic could backfire and turn public opinion against the truckers, subsequently reducing political support for change. He urged truckers to avoid roadside protest, warning them that: *“You are going to undermine everything we worked for and achieve nothing.”* Many truckers shared those concerns and were reluctant to join the rebellion. Although attempts to disrupt traffic did not gain critical mass support and essentially failed, the rift in the CS community suggested that many truckers were no longer interested in participating in the next scheduled crowdsourcing run. To save CS, the producer have initiated contact with the leaders of the new movement, many of whom were the same skeptical online opinion leaders he had previously engaged to secure support for the initial crowdsourcing run. Through a series of online communications and offline meetings, the producer convinced most of the detractors that antagonistic tactics would likely result in loss of the legitimacy they had painstakingly gained since the study began. The producer—by now perceived as “one of us” by the truckers—used his social capital to secure most detractors’ verbal commitment to not engage in road blockades, but instead to recommit to CS.

Coincidentally, as the protest movement was disintegrating, new stricter laws against cabotage were passed (see the “Institutional Change” section below for more details). Although the laws had little to do with the protests—and were initiated several months before they occurred—following validation of the crowdsourced data, the rebellion leaders took credit for the toughening laws, while also reasserting their reengagement in CS. The last crowdsourcing run took place in May 2015 and was even more successful than the previous two in terms of speed, number of participants and total observations (Figure 3.5).

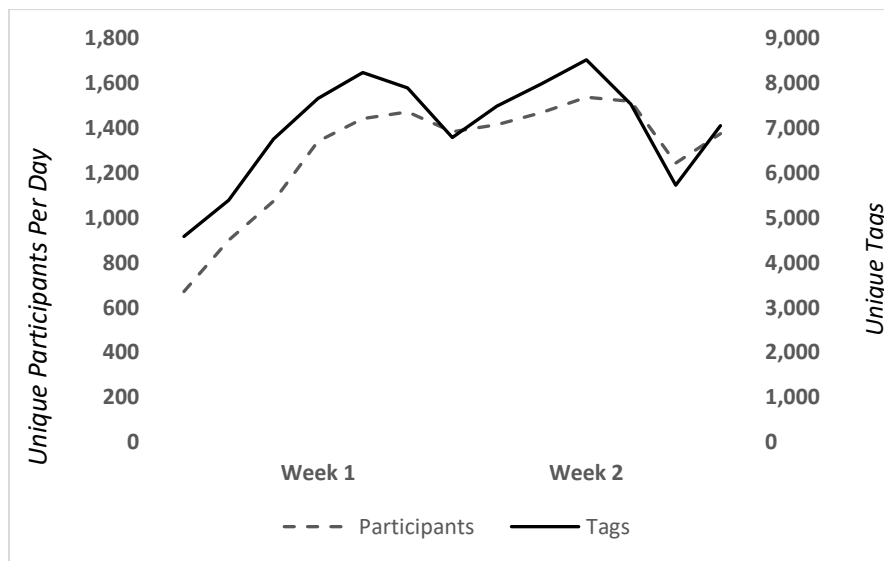


Figure 3.5 Third Data Crowdsourcing Run, May–June 2015

In summary, dissatisfied activists decide to move from tagging data to launching roadside protests—a move that had the potential to hurt the CS’s pragmatic and moral legitimacy (Table 3.4). Direct talks with and recognition of core providers reduced the rebellion’s intensity, and most of the opinion leaders in the trucker community recommitted to data crowdsourcing. This subsequently stabilized the pragmatic and moral legitimacy of the digital activism.

Table 3.4 Summary of the Case Study Analysis

Episode	Collective-Action Threats	Polycentric Governance Practices	Legitimate Digital Activism	Institutional Change
Instigating Institutional Change (Sept. 2012–March 2013)	<ul style="list-style-type: none"> • Rebellion: Truckers are highly skeptical about the crowdsourcing initiative even before it begins 	<ul style="list-style-type: none"> • Provider Recognition: The university researcher personally asks opinion leaders in the trucking community to help him, as principle producer, to address concerns about the livelihood of Swedish truck drivers 	<ul style="list-style-type: none"> • Pragmatic Legitimacy: The Swedish Transportation Union provided financial support for the research, which boosted its initial pragmatic legitimacy • Pragmatic Legitimacy: Opinion leaders in the trucking community commit themselves to CS 	<ul style="list-style-type: none"> • Institutional Practices: New research initiative to examine the impact of Policy 1072/2009
Emerging Collective Action (April 2013 - July 2013)	<ul style="list-style-type: none"> • Congestion: Study website repeatedly crashes due to high-volume information appropriation traffic • Pollution: Detection of nonsensical observations in data • Violation: Xenophobic sentiments expressed on the study's Facebook page 	<ul style="list-style-type: none"> • Incremental Adaptation: CS website was launched to give information appropriators access to interactive maps 	<ul style="list-style-type: none"> • Pragmatic Legitimacy: Increases as more people join data crowdsourcing • Moral Legitimacy: Decreases because of xenophobia • Cognitive Legitimacy: Decreases because of data pollution 	<ul style="list-style-type: none"> • Institutional Practices: Emergence of collective action involving truckers from opposing factions of the transport industry
Enacting Polycentric Governance (May 2013–September 2013)	N/A	<ul style="list-style-type: none"> • Incremental Adaptation: Expansion of the information commons' server capacity • Boundary Regulation: Creation of app registration and banning of polluters • Shared Accountability: Monitoring team for Facebook page and unofficial reprimanding of xenophobic comments • Provider Recognition: Ranking of top providers 	<ul style="list-style-type: none"> • Moral Legitimacy: Increases following the reduction of xenophobic posts • Cognitive Legitimacy: Increases as registration is required and polluters are banned 	<ul style="list-style-type: none"> • Institutional Practices: Trucker magazines and other supporters begin offering modest prizes to encourage information provision
Cascading Collective Action (May 2013–March 2014)	<ul style="list-style-type: none"> • Attacks: Supporters of Policy 1072/2009 question the study's cognitive and moral legitimacy 	<ul style="list-style-type: none"> • Shared Accountability: Additional institutional actors join the study and provide data; a large retailer in Sweden and three haulage firms share data 	<ul style="list-style-type: none"> • Cognitive Legitimacy: Increases as new data from industry actors enhances the study's factual accuracy; as a result, that accuracy is increasingly taken for granted, and its publicity and use in political discourse grows 	<ul style="list-style-type: none"> • Institutional Practices: Big retailers revamp their haulage sourcing practices • Institutional Regulations: The Swedish Parliament passes new laws to crackdown on illegal practices

Evading Provider Rebellion
(Nov 2014: Feb 2015)

- **Rebellion:** Dissatisfied truckers abandon the study and initiate road-side protests
- **Provider Recognition:** The principle producer personally engages protestors to convince them to abandon road-side protests in favor of data crowdsourcing
- **Pragmatic legitimacy-** Stabilizes as truckers recommit to data collection
- **Moral Legitimacy:** Stable, as roadside protests do not get out of hand
- **Institutional Practices:** A protest movement emerges

3.5.7 Institutional Change

More than three years after CS launch, we can see its institutional impact on both the regulatory and practice levels. Some of those changes occurred earlier on, while others are still unfolding. In response to the new revelations that CS brought to light, the Swedish Parliament passed new laws targeting violators of haulage regulations. The new provisions substantially increased the authority of Swedish traffic police to stop and inspect suspicious trucks, as well as to issue fines on the spot outside the regular legal due process. So, instead of issuing a ticket and setting a court date, the new provisions let police boot the trucks violating the law, immobilizing them on the roadside until the fines were paid. The fines also were increased: violations that were previously 3,000 SEK (approximately \$350 USD) were increased to 40,000 SEK (approximately \$5,000 USD). Haulers caught breaking the law also were fined heavily and potentially faced revocation of their business licenses. Similar fines and violator crackdowns were introduced in neighboring Denmark, Norway, and Finland. In our interviews with Union leaders and the high-ranking politician, we also learned that additional regulations are being initiated and discussed in the legislature; those regulations include better mechanisms for inspections, increases in booting times, and a provision making transport buyers legally liable for their subcontractors' offenses.

Prior to CS, Swedish political parties were split between those supporting further EU deregulations of cross-border transport and those opposing them. Following CS, however, all major Swedish parties vehemently opposed further EU cabotage deregulation. The Swedish representatives in the EU parliament built a coalition of West European countries to halt scheduled deregulations. Now, in 2018, previously scheduled deregulations are on hold, and the EU committee of cross-border commerce is currently engaging in a debate—with West European countries on one side, East European countries on the other—about increasing the regulations on working conditions and requiring safe-driving training of all truck drivers across the EU. The committee is tasked with redefining cabotage and suggesting a sensible way to regulate it.

The biggest practical impact of CS cited by all actors we interviewed was on public awareness. A little-known problem affecting an industry that few people think about became a central issue on the evening news and in the political arena on both national and international levels. CS brought to light the ugly unintended consequences of EU cross-border deregulations—rampant exploitation of East European labor (modern day “slaves” according to many of the people we interviewed), broken laws, safety issues associated with poorly trained truck drivers, an increase in CO2 emissions, and adverse effects on Swedish labor. The issue’s visibility caused many actors in the institutional field to alter their practices. Several retail companies and other transport buyers began openly questioning haulage companies about their practices to ensure that those companies were not breaking the law. For instance, ICA, Sweden’s largest retailer, and Elgiganten, its largest home electronics retailer, began requiring in their transport purchase agreements that all drivers be paid in accordance with the level of the Swedish Transport Union collective bargaining agreement. Other industry executives we spoke with advocated for improving working conditions for East European drivers, even if it meant increase their transportation costs. Reacting to pressure

from transport buyers, some haulage companies began cleaning house to reform their exploitative practices. The increase in enforcement and fines caused small East European haulage operators to leave Sweden and go elsewhere in the EU. Although many East European trucks continue operating in Sweden and Denmark, they typically belong to large haulage firms based in East Europe, which are less likely to break cabotage regulations than small-time independent truckers.

Finally—and somewhat unexpectedly—the attitude of Swedish drivers toward their East European counterparts has drastically shifted over the past few years. The initial resentment and calls for tougher punishments gradually turned to sympathy and calls for equal conditions. All truck drivers we spoke to expressed sympathy for the situation of East European drivers, many of whom they have come to know on a personal level.

3.6 Discussion and Theory Development

Using collective action as a shared perspective, we merged PIC theory (Mindel et al. 2018) and theory of institutional change (Hargrave and Van de Ven 2006; Seo and Creed 2002) to advance knowledge on the mutually constitutive duality (Feldman and Orlikowski 2011; Jones et al. 2004) between digital activism and the institutional context it aims to change. In addition, we drew on institutional theory on legitimacy (Suchman 1995; Zimmerman and Zeitz 2002) to propose that digital activism's success in institutional fields depends on its capacity to drive *legitimate* collective action. As the empirical foundation for our theorizing, we examined the CS case, in which actors harnessed social media and smartphone-based crowdsourcing to mobilize and drive collective action that successfully caused institutional change. The theoretical framing we adopted let us zoom in on the processes of and the relationships between establishing legitimacy, mobilizing collective action, and changing institutions through digital activism; it also let us focus

on how digital activism overcomes the degenerative collective-action threats that often arise. As Figure 6 shows, our work builds on PIC and institutional theories, the case study breakdown, and prior digital activism literature to develop a conceptual model and related propositions that explain how polycentric governance practices reduce collective-action threats to legitimate digital activism aimed at changing the institutional field.

Prior research asserts that the greater the legitimacy of an institutional order, the harder it is for a collective action to bring about change (Thomas et al. 1986; Walker et al. 1988). Our case study expands this theoretical insight by demonstrating how digitally enabled collective action undermined the legitimacy of an established institutional order. As our analysis revealed, this happened through a mutually constitutive process (Feldman and Orlikowski 2011; Jones et al. 2004) in which the digital activism established its own legitimacy, while at the same time eroding the legitimacy of the institutional order. Furthermore, the more the institutional order's legitimacy eroded, the more the digital activism's legitimacy grew. Hence, while CS was struggling to establish and maintain its legitimacy, once it secured that requisite pragmatic, moral, and cognitive legitimacy it was able to erode the legitimacy of the previous institutional order set by Regulation 1072/2009.

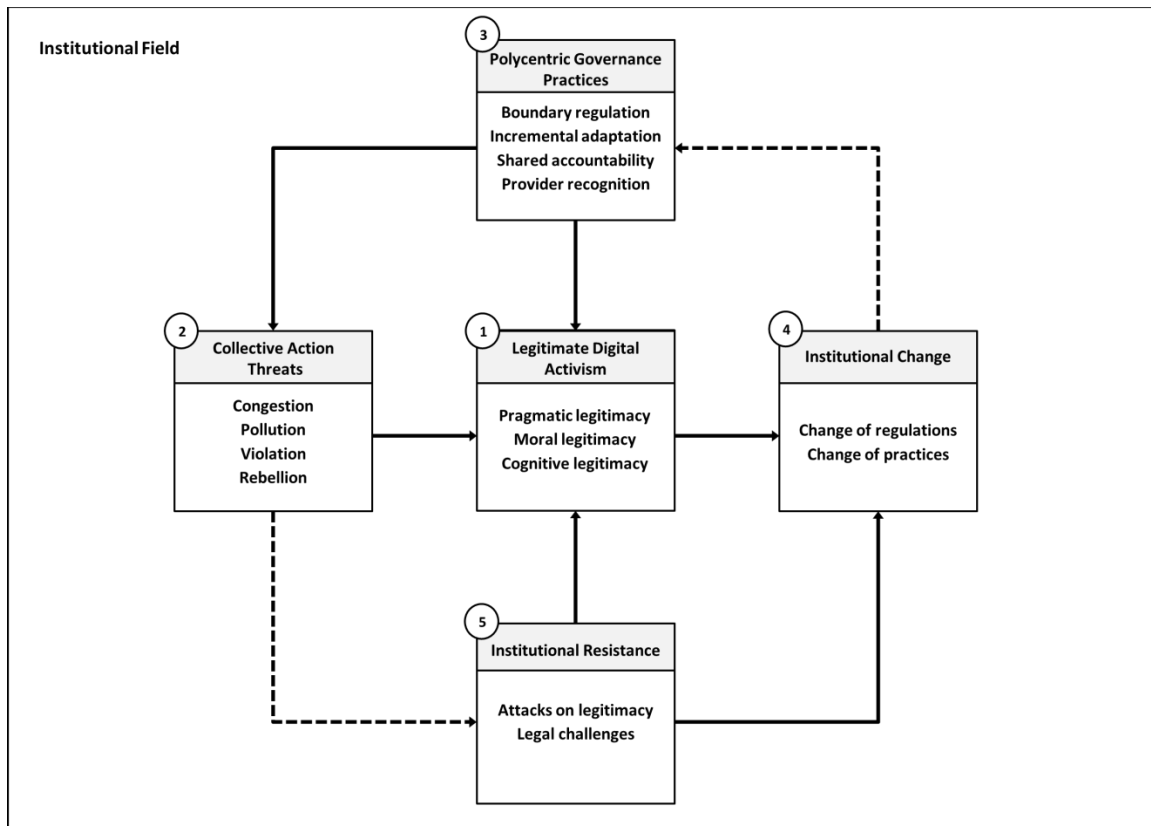


Figure 3.6 Conceptual Model

As the case demonstrates, it is difficult for digital activism—with its online and offline components—to establish and sustain legitimacy within an institutional field. The opportunity cost of participating in online activism is substantially lower in comparison. However, its pragmatic legitimacy often dissipates because it is hard for campaigns to remain continuously relevant in the fast-paced online environment, which has numerous competing causes and other attention-grabbing stimulations (Wasik 2009). At best, and however welcomed they might be, online campaigns are perceived as short-lived pushes to achieve a predefined goal—such as obtaining signatures or raising funds—rather than as long-term solutions (Kosinski 2014).

Where digital online activism is disadvantaged because of its focus on narrow goals, offline activism is disadvantaged because it often lacks focus (Indiviglio 2011). Digital offline activism typically aims to alter deep-rooted social, political, and economic institutions, and is therefore less

focused on specifics. Further, the fact that offline movements often lack a centralized voice—and various participants often demand different things—makes them seem naive or lacking a solid intellectual argument (Bhagwati 2004; Friedersdorf 2015; Indiviglio 2011). So, where online activism is often perceived as an exercise in clicktivism and therefore suffers from low moral legitimacy, offline activism suffers a similar fate because it trends toward echo-chamber effects and violent eruptions (White 2016).

Against these schisms, the CS case illustrates how digital activism online can usefully mix with and support offline activism. Without having recruited core participants through Facebook in the days prior to its first crowdsourcing run, CS would likely have failed to obtain the needed pragmatic legitimacy. As the validity of the data collected by the crowds increased over time, public support for the cause of addressing the problems created by Regulation 1072/2009 grew, increasing the moral legitimacy of the digital activism offline carried out through data crowdsourcing. In turn, CS's increased public awareness and moral legitimacy—playing out in the media and the political arena—fueled participation online and offline, which stabilized the campaign's pragmatic legitimacy. During its two-plus years, CS had only three relatively short data crowdsourcing runs—of six, four, and two weeks, respectively—but the results were impactful enough to sustain adequate activity online. Thus, when the next offline push was needed, activists were ready. In a sense, the online activities were the continuous, low-burning flame that could be turned up for short bursts of intense offline activity. Likewise, the offline activism's success helped ensure the online flame continued burning. Hence, we propose:

Proposition 1a: *Digital activism will more likely achieve institutional change when it is pragmatically, morally, and cognitively legitimate.*

Proposition 1b: *Online and offline digital activism reinforce each other's pragmatic, moral, and cognitive legitimacy.*

Digital activism's openness is simultaneously a source of strength and of weakness. On one hand, the aggregate effect of collective action carried out by many dispersed, loosely coordinated individuals can generate remarkable and significant outcomes. On the other hand, within the parameters of the primary collective action, other forms of degenerative collective actions may emerge (Mindel et al. 2018). Seo and Creed (2002) address the importance of establishing legitimacy for change, but they do not detail the process through which that legitimacy is established. The CS case expands their theoretical proposition by illustrating digital activism's potential role in establishing legitimacy for mobilizing collective action, and by zeroing in how Suchman's three forms of legitimacy (1995) stabilize in the face of emerging collective-action threats caused by the open-access nature of digital activism.

Information pollution is increasingly recognized as characteristic of information commons online that primarily depend on anonymous amateurs for content and information (Mindel et al. 2018). When data pollution emerged in CS, its cognitive legitimacy was seriously jeopardized as supporters of Regulation 1072/2009 who opposed CS questioned the data's validity in an effort to undermine the study. However, information pollution was not the only problem arising that challenged the CS effort's legitimacy. Given the crowdsourcing initiative's early success, the study website suffered from congestion and repeatedly crashed due to high-volume information appropriation traffic. Later, when various violations such as xenophobic posts and harassments emerged online, CS's moral legitimacy was at risk, as supporters of Regulation 1072/2009 attempted to discredit the activist movement by questioning the motivations of study participants and other institutional actors associated with CS.

In-fighting can also weaken activism, and prior research finds that competition for resources among activist movements addressing similar issues hurts the overall cause (Zald and McCarthy 1979). In the CS case, many of the truckers decided to quit and form their own activist group because they were dissatisfied with the pace and magnitude of change; as a result, the study's foundational support was on shaky ground, undermining its pragmatic legitimacy. While CS's openness to anyone willing to participate helped it to achieve a remarkable feat—mapping the movement of mobile targets on an unprecedented geographical scale—this openness also created the conditions for sudden dropout when this trucker rebellion created a competing movement. Besides the risk that road blockades might alienate the public, producers of CS were concerned that there were not enough resources in terms of committed time and energy to support two separate movements. All in all, this case study shows that digital activism's open nature created the conditions for individual level actions that, when aggregated, could undermine the legitimacy of CS. Drawing on these insights, we propose:

Proposition 2: *Collective-action threats of congestion, pollution, violation, and rebellion will likely reduce the legitimacy of digital activism.*

Although the producer's team did not necessarily see it at the time, in retrospect, adopting polycentric governance principles (Mindel et al. 2018) was crucial for CS success. Research is conflicted on the question of whether setting boundaries on anonymity is beneficial; some studies find that registration requirements do not necessarily increase content quality and credibility (Fogg et al. 2001), while other research shows that it clearly helps to reduce vandalism (Van Oorschot and Stubblebine 2006). Other work asserts that the overarching goal of the information commons should determine the boundaries on anonymity (Ren et al. 2007). In this case, requiring users to register to participate in the crowdsourcing initiatives substantially reduced information pollution

and let the producers flag suspicious tags, which further helped calibrate the data analysis algorithm.

On the question of incremental versus rapid innovation and system design, the literature is split (Norman and Verganti 2014). We observe that incremental adaptation of the CS app and associated website, as well as a stepwise approach to adjusting community governance practices on the Facebook page, effectively balanced the need for improvements and the danger of alienating participants. Changes made to the app, such as adding a registration requirement, were introduced gradually and in response to information pollution. Even then, however, the producers did not immediately ban suspicious taggers; instead, they used the contaminated data to their advantage, improving the triangulation algorithm.

Shared accountability helped CS improve its moral legitimacy. In line with previous findings on the role of peer monitoring in information commons (Chua et al. 2007; Williams and Cothrel 2000; Wise et al. 2006), the emergence of shared accountability on the study's Facebook page helped to significantly reduce harassment and xenophobic posts. Peer drivers reprimanded posters who crossed the line, and volunteer page monitors quickly identified and promptly removed posts that might undermine the study's moral legitimacy.

Prior research finds that attention and recognition through direct feedback, "likes," followers, icons, avatars, and other status symbols often motivates engagement in information commons, whether in social media, crowdsourcing, or online communities (Huberman et al. 2009; Lampel and Bhalla 2007; Moon and Sproull 2008). Our interviews showed us how much the hundreds of encouraging posts, the ranking system of top taggers, and the personal messages from the producer team helped to keep the drivers motivated. Similarly, posts and messages online and offline

between information providers—mostly drivers—helped to maintain a robust core of participants who recruited others, giving CS the momentum it needed during the crowdsourcing phases. Provider recognition also served to create relationships and a sense of trust among providers toward the producer, who later used this social capital created overtime to convince drivers not to quit CS. Drawing on these insights, we propose:

Proposition 3: *Polycentric governance practices of boundary regulation, incremental adaptation, shared accountability, and provider recognition will likely lessen collective-action threats and increase the legitimacy of digital activism.*

Changes caused by introducing new technologies into an environment are likely to create feedback effects on the technology's trajectory (Orlikowski 2007) as expressions of a mutually constitutive duality (Feldman and Orlikowski 2011; Jones et al. 2004). Similarly, the CS case study shows that, as digital activism impacted change in the institutional field, it was also impacted by the events it helped to set in motion. For instance, the attention CS received from the media and politicians motivated information providers to keep tagging, which increased the study's pragmatic legitimacy; however, this increased attention simultaneously prompted an increase in the intensity of xenophobic sentiments expressed online, which reduced the study's moral legitimacy. The decrease in moral legitimacy made many supporters uncomfortable, which prompted the "we need to do something about it" discussions that caused the production team to adjust the polycentric practices of CS. On the other hand, the more the CS legitimacy stabilized, and institutional regulations and practices were set in motion, the less concerned the production team was about adjusting polycentric practices. For instance, when the political establishment embraced CS and began using it to promote the legislative agenda, producers became less concerned with adjusting the app and less occupied with the Facebook page, and instead shifted their focus to further disseminating the data. However, when the protest movement began brewing, the principle

producer acted quickly to secure the commitment of truckers for the next CS crowdsourcing run. Observing how developments in the institutional field informed the adjustments to the app and the Facebook page activities, we propose:

Proposition 4: *Polycentric governance practices will more likely be effective if they are adjusted in response to changes in institutional regulations and practices.*

Institutions are reluctant to change (DiMaggio and Powell 1983) and, when challenged, vested interests are likely to fight back to undermine the challenger and preserve their position (Seo and Creed 2002). Prior research on online activism finds that powerful institutions may filter content and use surveillance to combat online activism that they perceive as threatening to their position (Ghobadi and Clegg 2015). The CS study also demonstrates how institutional actors may subtly exploit vulnerabilities to undermine digital activism. We find that the emergence of collective-action threats due to digital activism's openness was used as ammunition by opposition actors. Opponents of CS openly questioned the integrity of the data and its analysis, both publicly and in closed political circles. In some instances, opponents even engaged in deliberate data pollution to further reduce CS's cognitive legitimacy. While information pollution does not necessarily hurt digital activism's sustainability, it likely reduces its legitimacy and its ability to cause institutional change. In our interviews with the principle CS producer and a CS opponent, both mentioned the importance of establishing the credibility of CS to cause institutional change.

Similarly, violations in the form of xenophobia, racism, and harassment on the CS Facebook page were used to discredit the digital activism's moral legitimacy, as well as to endanger its pragmatic legitimacy and make institutional actors—including politicians, media, and corporations—

disassociate themselves from CS. In our interviews with the Union leaders, industry actors, the high-ranking politician, and the CS producers, they all expressed those concerns.

Some institutional opponents continue to work to reverse the legislations caused by CS. These opponents both challenge the legal standing of the new Swedish regulations and lobby the EU to intervene in Swedish legislation. Evidently, an impactful digital activism within an institutional field is bound to step on the toes of certain established institutional actors; they, in turn, are likely to resist it and its impact. Exploiting information commons' vulnerabilities, which are inevitable due to their openness, is one tactic that opposing actors are likely to use to fight back when they feel threatened. Building on those insights, we propose:

Proposition 5: *Collective-action threats will likely strengthen institutional resistance toward digital activism through attacks on its legitimacy and related legal challenges.*

3.7 Conclusion

There is considerable interest in how online technologies enable collective action and change (Ghobadi and Clegg 2015; Kelly Garrett 2006; Van Laer and Van Aelst 2010), with a focus on how social network connectivity supports communication and lubricates mobilization (Bennett and Segerberg 2012; Diani and McAdam 2003; Obar et al. 2012; Segerberg and Bennett 2011). We extend this existing research by examining in detail the constitutive duality between the collective action occurring at the digitally enabled social network level and the collective action occurring at the broader institutional field level. Specifically, we focus on the process of establishing legitimacy for impactful digital activism in the face of collective-action threats, which arise due to digital activism's open nature and decrease its pragmatic, moral, and cognitive legitimacy. Building on PIC and institutional change theory, and informed by empirical

observations, we posit that incorporating polycentric governance practices reduces the prevalence of collective-action threats, which in turn leads to stronger digital activism legitimacy and a higher likelihood of impactful institutional change.

In addition to contributing to digital activism literature, our CS case study advances PIC theory (Mindel et al. 2018). At present, PIC theory is a general framework for examining the impact of collective-action threats on information commons' sustainability and how polycentric governance practices can reduce those threats. Mindel et al. acknowledge the need for greater nuance for studying different types of information commons with specific and varying objectives. By conceptualizing digital activism as a type of information commons, we add to PIC theory, examining how collective-action threats play out in this context and how polycentric governance practices help stabilize digital activism allowing it to reach its objective of shaking the status quo within an institutional field. The resulting hybrid model of PIC and institutional legitimacy theories (Table 3.3 and Figure 3.6) was instrumental in providing a unique perspective on the question of how digital activism could be harnessed for promoting institutional change.

The CS case study provides important practical lessons for those who wish to use digital activism to promote change. Our findings suggest that it is not enough to tap into digitally enabled social networks to support communication for collective action and mobilize accordingly. The key for successful digital activism, we assert, is to continuously establish its pragmatic, moral, and cognitive legitimacy within an institutional field. To improve the odds of establishing legitimacy, digital activism must combat congestion, pollution, violations, and user rebellion. Further, we find that polycentric online and offline digital activism reinforce each other's legitimacy when incorporated under an overarching institutional change goal. Finally, we find that institutional actors who oppose that change are likely to exploit vulnerabilities that arise from digital activism's

openness to actively undermine the effort's legitimacy. Although it by no means ensures success, incorporating polycentric governance practices can help stabilize digital activism's legitimacy and thereby increase the likelihood of impactful institutional change.

Although we aim for our theorizing to apply across different forms of digital activism and different types of institutional fields, generalizability from a single case study is limited and additional research is needed to reinforce and refine the theory. As such, we hope that the new perspective we introduce will spur additional research into this increasingly important issue.

3.8 References

- Ackland, R., and O’Neil, M. 2011. “Online Collective Identity: The Case of the Environmental Movement,” *Social Networks* (33:3), pp 177-190.
- Agarwal, R., Animesh, A., and Prasad, K. 2009. “Research Note—Social Interactions and the “Digital divide”: Explaining Variations in Internet Use,” *Information Systems Research* (20:2), pp 277-294.
- Agarwal, S. D., Barthel, M. L., Rost, C., Borning, A., Bennett, W. L., and Johnson, C. N. 2014. Grassroots Organizing in the Digital Age: Considering values and Technology in Tea Party and Occupy Wall Street,” *Information, Communication & Society* (17:3), pp 326-341.
- Ameripour, A., Nicholson, B., & Newman, M. (2010). “Conviviality of Internet social networks: An exploratory study of Internet campaigns in Iran,” *Journal of Information Technology*, 25(2), 244-257.
- Andriole, S. J. 2010. “Business Impact of Web 2.0 Technologies,” *Communications of the ACM* (53:12), pp 67-79.
- Anduiza, E., Cristancho, C., and Sabucedo, J. M. 2014. “Mobilization Through Online Social Networks: The Political Protest of the Indignados in Spain,” *Information, Communication & Society* (17:6), pp 750-764.
- Aouragh, M., and Alexander, A. 2011. “The Arab Spring| the Egyptian Experience: Sense and N Nonsense of the Internet Revolution,” *International Journal of communication* (15:5), pp 1344-1358.
- Benbasat, I., Goldstein, D. K., and Mead, M. 1987. “The Case Research Strategy in Studies of Information Systems,” *MIS Quarterly* (11:3), pp 369-386.
- Bennett, W. L., and Segerberg, A. 2011. “Digital Media and the Personalization of Collective Action: Social Technology and the Organization of Protests Against the Global Economic Crisis,” *Information, Communication & Society* (14:6), pp 770-799.
- Bennett, W. L., and Segerberg, A. 2012. “The Logic of Connective Action: Digital Media and the Personalization of Contentious Politics,” *Information, Communication & Society* (15:5), pp 739-768.
- Bhagwati, J. 2004. “Anti-Globalization: Why?” *Journal of Policy Modeling* (26:4), pp 439-463.
- Bimber, B., Flanagin, A. J., and Stohl, C. 2005. “Reconceptualizing Collective Action in the Contemporary Media Environment,” *Communication Theory* (15:4), pp 365-388.
- Braa, J., Monteiro, E., and Sahay, S. 2004. “Networks of Action: Sustainable Health Information Systems Across Developing Countries,” *MIS Quarterly* (28:3), pp 337-362.
- Branagan, M. 2013. *Global Warming, Militarism and Nonviolence: The Art of Active Resistance*. Springer.
- Butler, B. S., and Wang, X. 2012. “The Cross-Purposes of Cross-Posting: Boundary Reshaping Behavior in Online Discussion Communities,” *Information Systems Research* (23:3-part-2), pp 993-1010.
- Butler, M. 2011. *Clicktivism, Slacktivism, or “Real” Activism Cultural Codes of American Activism in the Internet Era*. University of Colorado at Boulder.
- Calhoun, C. 2013. “Occupy Wall Street in Perspective,” *The British Journal of Sociology* (64:1), pp 26-38.
- Carney, N. 2016. “All Lives Matter, But So Does Race: Black Lives Matter and the Evolving Role of Social Media,” *Humanity & Society* (40:2) pp 180-199.

- Christensen, C. 2011. "Twitter Revolutions? Addressing Social Media and Dissent," *The Communication Review* (14:3), pp 155-157.
- Chua, C. E. H., Wareham, J., and Robey, D. 2007. "The Role of Online Trading Communities in Managing Internet Auction Fraud." *MIS Quarterly*, pp 759-781.
- Clifford, C. Jan. 20, 2010. "EBay Sellers' Rebellion: The Aftermath," *CNN Money*. Retrieved from http://money.cnn.com/2010/01/20/smallbusiness/ebay_aftermath/
- Deephouse, D. L., and Suchman, M. 2008. *Legitimacy in Organizational Institutionalism*. The Sage Handbook of Organizational Institutionalism.
- Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, E., and Quattrociocchi, W. 2016. "The Spreading of Misinformation Online." *Proceedings of the National Academy of Sciences* (113:3), pp 554-559.
- DeLuca, K. M., Lawson, S., and Sun, Y. 2012. "Occupy Wall Street on the Public Screens of Social Media: The Many Framings of the Birth of a Protest Movement," *Communication, Culture & Critique* (5:4), pp 483-509.
- Di Tullio, D., and Staples, D. S. 2013. "The Governance and Control of Open Source Software Projects," *Journal of Management Information Systems* (30:3), pp 49-80.
- Diamond, D. 2014. "The ALS Ice Bucket Challenge Has Raised \$100 Million—And Counting," *Forbes.com*.
- Diani, M., & McAdam, D. (Eds.). (2003). *Social Movements and Networks: Relational Approaches to Collective Action*. Oxford University Press.
- DiMaggio, P., and Powell, W. W. 1983. "The Iron Cage Revisited: Collective Rationality and Institutional Isomorphism in Organizational Fields," *American Sociological Review* (48:2), pp 147-160.
- Doubek, J. 2017. "Thousands Protest at Airports Nationwide Against Trump's Immigration Order," *National Public Radio, npr.org*.
- Edwards, F., Howard, P. P. N., and Joyce, M. 2013. "Digital Activism and Non-Violent Conflict," Available at SSRN: <https://ssrn.com/abstract=2595115>
- Eisenhardt, K. M. 1989. "Building Theories from Case Study Research," *Academy of Management Review* (14:4), pp 532-550.
- Eisenhardt, K. M., and Graebner, M. E. 2007. "Theory Building from Cases: Opportunities and Challenges," *Academy of Management Journal* (50:1), pp 25-32.
- Elliot, S. 2011. "Transdisciplinary Perspectives on Environmental Sustainability: A Resource Base and Framework for IT-Enabled Business Transformation," *MIS Quarterly* (35:1), pp 197-236.
- Eltantawy, N., and Wiest, J. B. 2011. "The Arab Spring| Social Media in the Egyptian Revolution: Reconsidering Resource Mobilization Theory," *International Journal of Communication* (18:5), pp 1207-1224.
- Enjolras, B., Steen-Johnsen, K., and Wollebæk, D. 2013. "Social Media and Mobilization to Offline Demonstrations: Transcending Participatory Divides?," *New Media & Society* (15:6), pp 890-908.
- Feldman, M. S., and Orlikowski, W. J. 2011. "Theorizing Practice and Practicing Theory." *Organization Science* (22:5), pp 1240-1253.
- Fogg, B. J., Jonathan Marshall, Othman Laraki, Alex Osipovich, Chris Varma, Nicholas Fang, Jyoti Paul. 2001. "What makes Web Sites Credible?: A Report on a Large Quantitative Study." In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pp 61-68.

- Friedersdorf, C. 2015. "Will Black Lives Matter Be a Movement That Persuades?" *The Atlantic Online*.
- Ganley, D., and Lampe, C. 2009. "The Ties that Bind: Social Network Principles in Online Communities," *Decision Support Systems* (47:3), pp 266-274.
- Garrett, R. K. 2009. "Echo Chambers Online?: Politically Motivated Selective Exposure Among Internet News Users," *Journal of Computer-Mediated Communication* (14:2), pp 265-285.
- Gebelhoff, R. 2015. "Scientists are Crediting the ALS Ice Bucket Challenge for Breakthroughs in Research," *The Washington Post online*.
- Gerbaudo, P. P. 2018. *Tweets and the Streets: Social Media and Contemporary Activism*. Pluto Press.
- Gerbaudo, P. 2018. *Tweets and the Streets: Social Media and Contemporary Activism*. Pluto Press.
- Ghobadi, S., and Clegg, S. 2015. "These Days Will Never be Forgotten...": A Critical Mass Approach to Online Activism," *Information and Organization* (25:1), pp 52-71.
- Ghonim, W. 2012. *Revolution 2.0: The Power of the People is Greater than the People in Power: A Memoir*. Houghton Mifflin Harcourt.
- Gillan, K. 2009. "The UK Anti-War Movement Online: Uses and Limitations of Internet Technologies for Contemporary Activism," *Information, Communication & Society* (12:1), pp 25-43.
- Gladwell, M. 2010. Small Change: Why the Revolution Will not be Tweeted. *The New Yorker online*.
- Gleason, B. 2013. "# Occupy Wall Street: Exploring Informal Learning About a Social Movement on Twitter," *American Behavioral Scientist* (57:7), pp 966-982.
- Goh, J. M., Gao, G. G., and Agarwal, R. 2016. "The Creation of Social Value: Can an Online Health Community Reduce Rural-Urban Health Disparities?" *MISQ* (40:1), pp 247-263.
- Goodchild, M. F., and Glennon, J. A. 2010. "Crowdsourcing Geographic Information for Disaster Response: A Research Frontier," *International Journal of Digital Earth* (3:3), pp 231-241.
- Graham-Felsen, S. Sep. 7, 2006. "The Facebook Rebellion," *CBS News*. Retrieved from: <http://www.cbsnews.com/news/the-facebook-rebellion/>
- Gross, D. Dec. 19, 2012. "Instagram Backtracks After User Privacy Revolt," *CNN*. Retrieved from <http://www.cnn.com/2012/12/18/tech/social-media/instagram-terms-users/>
- Gu, B., Konana, P., Rajagopalan, B., and Chen, H.-W. M. 2007. "Competition among virtual communities and user valuation: The case of investing-related communities," *Information Systems Research* (18:1), pp 68-85.
- Gunitsky, S. (2015). Corrupting the Cyber-Commons: Social Media as a Tool of Autocratic Stability. *Perspectives on Politics* (13:1), pp 42-54.
- Halupka, M. 2014. "Clicktivism: A Systematic Heuristic," *Policy & Internet*, (6:2), pp115-132.
- Hargrave, T. J., and Van de Ven, A. H. 2006. "A Collective Action Model of Institutional Innovation," *Academy of Management Review* (31:4), pp 864-888.
- Harlow, S., and Harp, D. 2012. "Collective Action on the Web: A Cross-Cultural Study of Social Networking Sites and Online and Offline Activism in the United States and Latin America," *Information, Communication & Society* (15:2), pp 196-216.
- Heipke, C. 2010. Crowdsourcing Geospatial Data. *ISPRS Journal of Photogrammetry and Remote Sensing* (65:6), pp 550-557.

- Hertel, G., Niedner, S., and Herrmann, S. 2003. "Motivation of software developers in Open Source projects: an Internet-based survey of contributors to the Linux kernel," *Research Policy* (32:7), pp 1159-1177.
- Hilal, N. 2008. "Unintended Effects of Deregulation in the European Union: The Case of Road Freight Transport," *Sociologie du Travail* (50), pp 19-29.
- Hiltzik, M. 2014. "A Few (Impolite) Questions About the Ice Bucket Challenge," *LA Times* online.
- Howard, P. P. N., and Hussain, M. M. 2011. "The Role of Digital Media," *Journal of Democracy* (22:3), pp 35-48.
- Hsieh, J. P. A., Rai, A., and Keil, M. 2008. "Understanding Digital Inequality: Comparing Continued Use Behavioral Models of the Socio-economically Advantaged and Disadvantaged," *MIS Quarterly*, (32:1), pp. 97-126.
- Hu, N., Bose, I., Koh, N. S., and Liu, L. 2012. "Manipulation of Online Reviews: An Analysis of Ratings, Readability, and Sentiments," *Decision Support Systems* (52:3), pp 674-684.
- Huberman, B. A., Romero, D. M., and Wu, F. 2009. "Crowdsourcing, Attention and Productivity," *Journal of Information Science* (35:6), pp 758-765.
- Indiviglio, D. 2011. "5 Reasons Why 'Occupy Wall Street' Won't Work," *The Atlantic* online.
- Ives, B., and Olson, M. H. 1984. User Involvement and MIS Success: A Review of Research," *Management Science* (30:5), pp 586-603.
- Jarvenpaa, S. L., and Lang, K. R. 2011. "Boundary Management in Online Communities: Case Studies of the Nine Inch Nails and ccMixer Music Remix Sites," *Long Range Planning* (44:5), pp 440-457.
- Jones, M., Orlikowski, W., and Munir, K. 2004. "Structuration Theory and Information Systems: A Critical Reappraisal," *Social Theory and Philosophy for Information Systems*. pp 297-328.
- Juris, J. S. 2005. The new digital media and activist networking within anti-corporate globalization movements. *The Annals of the American Academy of Political and Social Science*, 597(1), 189-208.
- Karppf, D. 2010. "Online Political Mobilization from the Advocacy Group's Perspective: Looking Beyond Clicktivism," *Policy & Internet* (2:4), pp 7-41.
- Kata, A. 2012. "Anti-Vaccine Activists, Web 2.0, and the Postmodern Paradigm—An Overview of Tactics and Tropes Used Online by the Anti-Vaccination Movement," *Vaccine* (30:25), pp 3778-3789.
- Katz, J. 2017. "The Red Cross Won't Save Houston," *Slate.org*.
- Kelly Garrett, R. 2006. "Protest in an Information Society: A Review of Literature on Social Movements and New ICTs," *Information, Communication & Society* (9:02), pp 202-224.
- Ketter, W., Peters, M., Collins, J., and Gupta, A. 2016. "A Multiagent Competitive Gaming Platform to Address Societal Challenges," *MIS Quarterly* (40:2). pp447-460.
- Kittur, A., Suh, B., and Chi, E. H. 2008, November. "Can you Ever Trust a Wiki?" Impacting Perceived Trustworthiness in Wikipedia," In *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pp 477-480.
- Kosinski, B. 2014. "#IceBucketChallenge: Why You're Not Really Helping," *huffingtonpost.com*.
- Kristofferson, K., White, K., and Peloza, J. 2013. "The Nature of Slacktivism: How the Social Observability of an Initial Act of Token Support Affects Subsequent Prosocial Action," *Journal of Consumer Research* (40:6), pp 1149-1166.

- Kujala, S. 2003. "User Involvement: A Review of the Benefits and Challenges," *Behaviour & Information Technology* (22:1), pp 1-16.
- Kummer, S., Dieplinger, M., and Fürst, E. 2014. "Flagging out in Road Freight Transport: A Strategy to Reduce Corporate Costs in a Competitive Environment: Results from a Longitudinal Study in Austria," *Journal of Transport Geography* (36), pp 141-150.
- Lampel, J., and Bhalla, A. 2007. "The Role of Status Seeking in Online Communities: Giving the Gift of Experience," *Journal of Computer-Mediated Communication* (12:2), pp 434-455.
- Larkin, M. 2016. "'I can't subject myself any longer to the hate': Fifth Harmony's Normani Kordei quits Twitter over 'racist cyberbullying'," *Dailymail.com*.
- Lee, H. 2005. "Behavioral Strategies for Dealing with Flaming in an Online Forum," *The Sociological Quarterly* (46:2), pp385-403.
- Lee, A. S., and Baskerville, R. L. 2003. "Generalizing Generalizability in Information Systems Research," *Information Systems Research* (14:3), pp 221-243.
- Lee, Y. H., and Hsieh, G. 2013, April. "Does Slacktivism Hurt Activism?: The Effects of Moral Balancing and Consistency in Online Activism," In Proceedings of the *SIGCHI Conference on Human Factors in Computing Systems*, pp 811-820. ACM.
- Levina, N., and Arriaga, M. 2014. "Distinction and status production on user-generated content platforms: Using Bourdieu's theory of cultural production to understand social dynamics in online fields," *Information Systems Research* (25:3), pp 468-488.
- Li, T. B. Q. 2005. "Cyber-Harassment: A Study of a New Method for an Old Behavior," *Journal of Educational Computing Research*, (32:3), pp 265-277.
- Lim, M. 2013. "Many Clicks but Little Sticks: Social Media Activism in Indonesia," *Journal of Contemporary Asia* (43:4), pp 636-657.
- Lyytinen, K., and Newman, M. 2008. "Explaining Information Systems Change: A Punctuated Socio-Technical Change Model," *European Journal of Information Systems* (17:6), pp 589-613.
- Majchrzak, A., and Malhotra, A. 2013. "Towards an Information Systems Perspective and Research Agenda on Crowdsourcing for Innovation," *The Journal of Strategic Information Systems* (22:4), pp 257-268.
- Melville, N. 2010. "Information Systems Innovation for Environmental Sustainability," *MIS Quarterly*, (34:1), pp 1-21.
- Mignerat, M., and Rivard, S. 2009. "Positioning the Institutional Perspective in Information Systems Research," *Journal of Information Technology* (24:4), pp 369-391.
- Moon, J. Y., and Sproull, L. S. 2008. "The Role of Feedback in Managing the Internet-Based Volunteer Work Force," *Information Systems Research* (19:4), pp 494-515.
- Morris, A. D., and Staggenborg, S. 2004. "Leadership in Social Movements" *The Blackwell Companion to Social Movements*, pp 171-196.
- Myers, M. D. 1997. "Qualitative Research in Information Systems," *MIS Quarterly* (21:2), pp 241-242.
- Newman, M., and Robey, D. 1992. "A Social Process Model of User-Analyst Relationships," *MIS Quarterly* (16:2), pp 249-266.
- Norman, D. A., and Verganti, R. 2014. "Incremental and Radical Innovation: Design Research vs. Technology and Meaning Change," *Design Issues* (30:1), pp 78-96.

- Obar, J. A., Zube, P., and Lampe, C. 2012. "Advocacy 2.0: An Analysis of how Advocacy Groups in the United States Perceive and Use Social Media as Tools for Facilitating Civic Engagement and Collective Action," *Journal of Information Policy* (11:2), pp 1-25.
- Oliver, C. 1991. "Strategic Responses to Institutional Processes," *Academy of Management Review* (16:1), pp 145-179.
- Orlikowski, W. J. 2007. "Sociomaterial Practices: Exploring Technology at Work," *Organization Studies* (28:9), pp 1435-1448.
- Orlikowski, W. J., and Barley, S. R. 2001. "Technology and Institutions: What can Research on Information Technology and Research on Organizations Learn from Each Other?." *MIS Quarterly* (25:2), pp 145-165.
- Ostrom, E. 1990. *Governing the commons*, Cambridge university press.
- Ren, Y., Kraut, R., and Kiesler, S. 2007. "Applying common identity and bond theory to design of online communities," *Organization Studies* (28:3), pp 377-408.
- Richardson, J., and Mazey, S. (Eds.). 2015) *European Union: Power and Policy-Making*. Routledge.
- Saunders, C. 2007. "Editor's comments: Information Systems in Developing Countries," *MIS Quarterly*, (31:2), pp iii-vi.
- Schimmelfennig, F. 2001. "The Community Trap: Liberal Norms, Rhetorical Action, and the Eastern Enlargement of the European Union," *International Organization* (55:1), pp 47-80.
- Seegerberg, A., and Bennett, W. L. 2011. "Social Media and the Organization of Collective Action: Using Twitter to Explore the Ecologies of two Climate Change Protests," *The Communication Review* (14:3), pp 197-215.
- Selander, L., and Jarvenpaa, S. L. 2016. "Digital Action Repertoires and Transforming a Social Movement Organization," *MIS Quarterly* (40:2), pp 331-352.
- Seo, M. G., and Creed, W. D. 2002. "Institutional Contradictions, Praxis, and Institutional Change: A Dialectical Perspective," *Academy of Management Review* (27:2), pp 222-247.
- Shulman, S. W. 2009. "The Case Against Mass e-mails: Perverse Incentives and Low Quality Public Participation in US Federal Rulemaking," *Policy & Internet*, (1:1), pp 23-53.
- Silverman, C., and Singer-Vine, J. 2016. "Most Americans who see Fake News Believe it, New Survey Says," *BuzzFeed News*.
- Sternberg, H., Filipiak, M., Hofmann, E., and Hellström, D. 2015. "CABOTAGESTUDIEN-A study on trucking deregulation and cabotage in Scandinavia and beyond,"
- Stevens, M., and D'Hondt, E. 2010, September. "Crowdsourcing of Pollution Data Using Smartphones," In Workshop on Ubiquitous Crowdsourcing.
- Streeck, W., and Thelen, K. A. 2009. *Institutional Change in Advanced Political Economies. In Debating Varieties of Capitalism* (pp. 95-131). Oxford University Press.
- Suchman, M. C. 1995. "Managing Legitimacy: Strategic and Institutional Approaches," *Academy of Management Review* (20:3), pp. 571-610.
- Sullivan, L. 2015. "In Search of the Red Cross' \$500 Million in Haiti Relief," *National Public Radio*. NPR.org.
- Tankersley, J. 2011. "Occupy Wall Street's Marketing Crisis: What Would an OWS Brand Look Like?" *The Atlantic* Online.
- Tapscott, D., and Williams, A. D. 2008. *Wikinomics: How Mass Collaboration Changes Everything*. Penguin.

- Theocharis, Y., Lowe, W., Van Deth, J. W., and García-Albacete, G. 2015. "Using Twitter to Mobilize Protest Action: Online Mobilization Patterns and Action Repertoires in the Occupy Wall Street, Indignados, and Aganaktismenoi Movements," *Information, Communication & Society* (18:2), pp 202-220.
- Thomas, G. M., Walker, H. A., and Zelditch Jr, M. 1986. "Legitimacy and Collective Action," *Social Forces* (65:2), pp 378-404.
- Tremayne, M. 2014. Anatomy of Protest in the Digital Era: A Network Analysis of Twitter and Occupy Wall Street," *Social Movement Studies* (13:1), pp 110-126.
- Tufekci, Z., and Wilson, C. 2012. "Social Media and the Decision to Participate in Political Protest: Observations from Tahrir Square," *Journal of Communication* (62:2), pp 363-379.
- Van Laer, J., and Van Aelst, P. P. 2010. "Internet and Social Movement Action Repertoires: Opportunities and Limitations," *Information, Communication & Society* (13:8), pp 1146-1171.
- Van Oorschot, P. P. C., and Stubblebine, S. 2006. "On Countering Online Dictionary Attacks with Login Histories and Humans-in-the-Loop," *ACM Transactions on Information and System Security (TISSEC)* (9:3), pp 235-258.
- Venkatesh, V., and Sykes, T. A. 2013. "Digital Divide Initiative Success in Developing Countries: A Longitudinal Field Study in a Village in India," *Information Systems Research*, (24:2), pp 239-260.
- Vie, S. 2014. "In defense of "slacktivism": The Human Rights Campaign Facebook Logo as Digital Activism," *First Monday*, (19:4).
- Von Krogh, G., Haefliger, S., Spaeth, S., and Wallin, M. W. 2012. "Carrots and Rainbows: Motivation and Social Practice in Open Source Software Development," *MIS Quarterly* (36:2), pp 649-676.
- Walker, H. A., Rogers, L., and Zelditch Jr, M. 1988. "Legitimacy and Collective Action: A Research Note," *Social Forces* (67:1), pp 216-228.
- Walsham, G., Robey, D., and Sahay, S. 2007. "Foreword: Special Issue on Information Systems in Developing Countries," *MIS Quarterly*, (31:2), pp 317-326.
- Warzel, C. (Aug 2016). "A Honeypot for Assholes: Inside Twitter's 10-Year Failure to Stop Harassment," *Buzzfeed.com*
- Wasik, B. 2009. *And then there's this: How Stories Live and Die in Viral Culture*. Penguin.
- Wattal, S., Schuff, D., Mandviwalla, M., and Williams, C. B. 2010. "Web 2.0 and Politics: The 2008 US Presidential Election and an E-Politics Research Agenda," *MIS Quarterly* (34:4), pp 669-688.
- White, M. 2016. *The End of Protest: A New Playbook for Revolution*. Knopf Canada.
- White, M. 2017. "Occupy and Black Lives Matter Failed. We can either Win Wars or Win Elections," *The Guardian online*.
- Williams, R. L., and Cothrel, J. 2000. "Four Smart Ways to Run Online Communities," *Sloan Management Review* (41:4), pp 81-92.
- Wise, K., Hamman, B., and Thorson, K. 2006. "Moderation, Response Rate, and Message Interactivity: Features of Online Communities and their Effects on Intent to Participate," *Journal of Computer-Mediated Communication* (12:1), pp 24-41.
- Xu, S. X., and Xiaoquan (Michael) Zhang. 2013. "Impact of Wikipedia on Market Information Environment: Evidence on Management Disclosure and Investor Reaction," *MISQ* (37:4), pp 1043-1068.

- Yan, L., and Tan, Y. 2014. "Feeling Blue? Go Online: An Empirical Study of Social Support Among Patients," *Information Systems Research* (25:4), pp 690-709.
- Yates, D., and Paquette, S. 2011. "Emergency Knowledge Management and Social Media Technologies: A Case Study of the 2010 Haitian Earthquake," *International Journal of Information Management* (31:1), pp 6-13.
- Yin, R. K. 2008. *Case Study Research and Applications: Design and Methods*. Sage Publications.
- Zald, M. N., and McCarthy, J. D. 1979. "Social movement industries: Competition and Cooperation Among Movement Organizations,"
- Zimmerman, M. A., and Zeitz, G. J. 2002. "Beyond Survival: Achieving New Venture Growth by Building Legitimacy," *Academy of Management Review* (27:3), pp 414-431.

CHAPTER 4: POLYCENTRIC RESOLUTION OF INFORMATION POLLUTION IN OPEN COLLABORATION SYSTEMS

ABSTRACT

The impact of open collaboration (OC) systems online is widespread, touching many individuals and institutions in society. At the same time, empirical and anecdotal evidence suggests that OC systems can be contaminated with low-quality, biased, unintentionally misleading, and even deliberately manipulated content. While prior research has examined how different organizational mechanisms relate to the quality of final or mature OC outputs, we know little about how OC systems address information pollution problems as they arise during output development. Against this backdrop, we draw on polycentric information commons (PIC) theory to offer a detailed analysis of how information pollution is resolved in Wikipedia articles. The result is two contributions to theory. First, we advance knowledge on information pollution as an inherent threat to OC systems and on how these systems can successfully address pollution through polycentric practices. Second, we offer empirical validation of PIC theory and extend it with new insights into how polycentric principles may be incorporated on the operational levels of governance.

Key Words: Open collaboration, polycentric information commons, information pollution, quality management, Wikipedia

4.1 Information Pollution in Open Collaboration Systems Online

Open-collaboration (OC) systems online—such as Wikis, open source software (OSS) development, crowdsourced innovation, and other online sociotechnical arrangements that rely mostly on self-selecting participants for creating value (Levine and Prietula 2013)—represent a novel production paradigm distinctively different from traditional business organizations (Benkler 2006; Baldwin and Von Hippel 2011; Von Hippel and Von Krogh 2003). Prior organization research on OC systems has examined their emerging structures (Baldwin and Von Hippel 201; Kittur et al. 2009), dynamics (Kane and Ransbotham 2016; Ransbotham and Kane 2011), procedures (Faraj et al. 2011; Levine and Prietula 2013; Viégas et al. 2007), and governance (Aaltonen and Lanzara 2015; Forte et al. 2009; Markus 2007; Shah 2006), and how these different mechanisms relate to the quality of final or mature outputs¹⁵ (Aberdour 2007; Wilkinson and Huberman 2007). However, our understanding of how OC systems address quality issues as they arise remains limited. Drawing on polycentric information commons (PIC) theory (Mindel et al. 2018), we conceptualize arising quality issues as “information pollution” events to examine how polycentric governance helps resolve pollution in OC systems.

It is hardly surprising that open-access platforms—including social media, media-sharing websites, blogs, peer-to-peer networks, and online-review platforms—can become inundated with polluted content (Clauson et al. 2008; Gyimothy et al. 2005; Mayzlin et al. 2014). The pollution spectrum ranges from unintentionally misleading or deliberately manipulated information to disturbingly gruesome content. For instance, reports indicate that Facebook and Twitter are flooded with false rumors, fake news stories, and other polluted content (Allcott and Gentzkow

¹⁵ Output of OC systems, whether code or information, is constantly updated and refined. Prior research has typically relied on the quality scores assigned by the contributor community or outside experts after the output is mature and relatively stable.

2017). Daily, an army of moderators in these companies work “to soak up the worst of humanity in order to protect the rest of us” (Chen 2014), taking down racist, masochistic, violence-inciting posts, as well as shocking images and videos of animal cruelty, beheadings, and child pornography (Kanter 2018). Facebook and Twitter concede that they are overwhelmed, despite constantly increasing the number of monitors (Kanter 2018). By comparison, most open-access platforms have far fewer resources to combat the problem.

An examination of prominent OSS development platforms finds a high prevalence of defunct, dirty data (Howison and Crowston 2004). Moreover, OSS (like all software) may contain bugs (Gyimothy et al. 2005; Stamelos et al. 2002) and is vulnerable to misuse by actors that deliberately insert malicious code (Ransbotham 2010). Similarly, Wikis may contain incomplete, biased, manipulated, and erroneous information (Holman Rector 2008; Kupferberg and Protus 2011; Lavasa et al. 2011). Given that virtually everyone, from students (Haigh 2011; Lim 2009) to professionals (Brokowski and Sheehan 2009; Miller and Murray 2010; Peoples 2009), turn to Wikis for information, frequent and persisting low-quality information is likely to spillover and impact society as whole.¹⁶ In addition, frequent and persisting low-quality information is likely to reduce the trust we put in OC systems. Hence, the question of how these systems address pollution quickly before it spills over and causes mistrust is crucial to the lasting viability of the novel and innovative OC mode. Although extant research provides important empirical findings on some factors associated with OC system quality (Arazy et al. 2011; Arazy and Nov 2010; Kittur and Kraut 2008), it provides little theoretical insight on which OC governance practices best achieve timely resolution of quality issues as they arise.

¹⁶ For instance, false information on vaccinations adversely impacts the public health, while false information used in litigation may create bad precedence.

PIC theory builds on the tragedy of the commons (Hardin 1968) and collective-action paradigms (Ostrom 1990, 2000) to conceptualize online systems characterized by low barriers to participation as *information commons*. Defined as *highly accessible, self-rising information systems in which stakeholders share an overarching goal* (Mindel et al. 2018, pp. 609), information commons, more so than traditional centralized systems, are susceptible to information pollution and other collective-action threats because of their openness (Mindel et al. 2018). PIC theory asserts that one way to address these collective-action threats is by incorporating polycentric principles into the system's governance practices and technical features (Mindel et al. 2018).

To advance our understanding of how OC systems can effectively resolve pollution, we draw on PIC theory to examine how polycentric practices—shared accountability through work distribution, boundary regulation on content, incremental adaptation of the information, and recognition of peer contributions—impact resolution of pollution events in Wikipedia. We identify the posting of a cleanup template and collective actions toward its removal as an information pollution event on an article. Because pollution in the physical world as well as in the information world may spread outside of its immediate boundaries, it is important to resolve it effectively in an expedient manner. Using pollution events as the unit of analysis, we therefore zoom-in on the relationship between polycentric practices during the event and the time it takes to resolve the event as an indicator of the effectiveness of polycentric governance. As a result, our research advances knowledge on information pollution and how using polycentric practices in OC systems may successfully address it. In addition, our research contributes to PIC theory by empirically examining its propositions and by offering novel insights into how polycentric principles may be incorporated in governance at the operational level.

4.2 Addressing Information Pollution in Open Collaboration

OC systems online, initially spearheaded by OSS development and wiki technologies, represent a novel model of organizing production that is distinctly different from traditional managerial settings (Aaltonen and Lanzara 2015; Benkler 2006; Forte and Lampe 2013; Levine and Prietula 2014). OC systems are defined as an “*online environment that supports the collective production of an artifact through a technologically mediated collaboration platform that presents a low barrier to entry and exit and supports the emergence of persistent but malleable social structures*” (Forte and Lampe 2013, p. 57). These systems generate value by broadly inviting people to participate in the production of a common good and allowing them to build relatively freely on each other’s work (Kumar et al. 2011; Levine and Prietula 2014).

Research on OC systems provides various insights into tensions underlying their operation. Although low barriers to participation create conditions for fast growth, they also may spur rapid decline when participants suddenly disengage (Ransbotham and Kane 2011). This tension between freedom of entry and exit in OC systems introduces a source of volatility that is largely non-existent in traditional organizations. In fact, research has found that most participants contribute only once and never return (Anthony et al. 2009; Panciera et al. 2009; Shah 2006). Most of the work is subsequently shouldered by a relatively small group of dedicated core participants (Panciera et al. 2009; Shah 2006) who may also exit at any time. Thus, it is often essential in OC systems to continuously attract new contributors to replace those who leave (Ransbotham and Kane 2011) and to inject renewed energy, perspectives, and ideas to the system (Morgan et al. 2013). Revitalization of content and participants is essential for the long-term prospects of OC systems.

The need to attract and retain new participants gives rise to another tension—that between long-time and new participants. While most organizations experience the challenge of incorporating new members (Allen and Meyer 1990), the challenge can be even more pronounced in OC systems that lack dedicated human resource managers, new employee orientation programs, or clear incentive structures. Further, research shows that veteran participants can be impatient with newcomers who are not familiar with task requirements or the system’s established operational norms (Halfaker et al. 2013; Suh et al. 2009). Research has also found that participants in OC systems are often driven by a combination of needs for both personal and social fulfillment (Hertel et al. 2003; Morgan et al. 2013), and that the rejection of newcomers by veterans reduces motivation, causing many to leave (Halfaker et al. 2013). It is therefore critical for OC systems to resolve such tensions between veteran and new providers in a constructive manner.

Work in OC systems is typically self-selected rather than assigned (Crowston et al. 2005; Kittur and Kraut 2008; Stvilia et al. 2007). This freedom of choice is a major catalyst of creativity (Benkler 2006), but it also may cause bottlenecks in production, inefficiencies, and conflict (Kittur and Kraut 2010; Mishra et al. 2002; Yasseri et al. 2012). Specifically, the increasing empirical and anecdotal evidence of low-quality and manipulated content in OC systems (Gyimothy et al. 2005; Hafner 2007; Holman Rector 2008; Kupferberg and Protus 2011) suggests that the low barriers to participation amplify the systems’ vulnerability to misuse in comparison to traditional closed systems. To manage quality, OC systems typically adopt various governance mechanisms that seek to strike a balance between necessary controls and inherent openness, typically by including elements of bureaucracy, democratic procedures, and mechanisms that facilitate independent participation. Different OC systems often include some combination of communication mechanisms (Gutwin et al. 2004, Viegas et al. 2007), formal and informal rules (Butler et al. 2008;

Scacchi 2007), and soft hierarchies (Crowston and Howison 2006; Cabunducan et al. 2011) to improve coordination and overall performance.

To understand how current research addresses the relation between these different organizational mechanisms and the quality of outputs, Table 4.1 offers an overview of OC system quality research. It is important to note that the intangible nature of outputs—that is, content, code, and ideas—and the fact that outputs constantly change through updates and edits, make substantive quality a moving target. Many researchers therefore rely on small-scale qualitative assessments (Arazy and Nov 2010; Chesney 2006; Giles 2005) to determine quality. Other researchers turn to internal rating schemes created in OC systems to examine which factors are associated with quality (Blumenstock 2008; Kane 2010; Kittur and Kraut 2008; Wöhner and Peters 2009) or use proxy metrics such as contribution retention rate in relation to output lifetime as a quality indicator (Anthony et al. 2007). Other approaches include using automated tools to detect errors in code (Stamelos et al. 2002) and using quality indicators in Wikipedia articles (Dalip et al. 2009). Although each of these studies provides important insights, the difference between methods and units of analysis makes it difficult to compare and consolidate findings. Moreover, current research focuses on factors associated with the quality of final or mature outputs rather than on the mechanisms used to resolve quality issues as they arise during development of these outputs.

Study	Theoretical perspective	Method and data	Unit of analysis	Findings and insights
Anderka et al. 2011	Inductive machine-learning-based prediction model of quality problems; no specific hypotheses	Machine-learning density estimation with class probability estimation	10,000 Wikipedia articles with one or more out of 10 types of cleanup templates	<ol style="list-style-type: none"> 1. Method for mining and extracting cleanup templates 2. The algorithm successfully predicted quality issues associated with 4 out of 10 templates
Anthony et al. 2007	Deductive hypothesis development	Examination of the effect of contributor anonymity and edit	Quality of contributions measured as the	<ol style="list-style-type: none"> 1. Registered users' quality increases with more contributions

	partially based on previous empirical findings	history on quality of contributions in Wikipedia	rate of each contributor's content retained in the current version	2. Large number of high-quality contributions come from anonymous contributors who contribute only once
Arazy and Nov 2010	Deductive hypothesis development partially based on previous empirical findings	Qualitative assessment of the accuracy and completeness of 50 articles by "experts" Quantitative measures of editing activities on those articles	Wikipedia article and editors	1. Coordination positively impacts article quality 2. Unequal contribution among editors on the article level is not associated with quality 3. Unequal contribution among the same editors across Wikipedia is positively associated with quality 4. Number of editors is not directly associated with quality
Blumenstock 2008	None (no hypotheses)	Comparison of length (word count) in Wikipedia's featured articles and a random sample of other articles	Wikipedia article word count	1. Positive association between word count and designation of featured article status
Chesney 2006	None (deductive hypothesis development)	Survey-based assessment of the credibility of 30 articles by experts, 24 random-assignment to non-experts and Wikipedia as a whole	Perception of the credibility of the Wikipedia article and editor	1. Experts found articles to be more credible than non-experts 2. 13% of articles were reported for mistakes
De la Calzada and Dekhtyar 2010	None (exploratory hypothesis development)	Comparison of two models of quality estimation of 50 stable and 29 controversial Wikipedia articles	Article quality as assessed by students	1. Quality of articles of different type should be computed using different means 2. Methodological insights on using multiple amateur quality evaluators as opposed to individual experts
Giles 2005	None (no hypotheses)	Blind comparison by experts of 42 science articles in Wikipedia and Encyclopedia Britannica	Wikipedia articles	1. Articles from both sources are found to contain a similar average number of errors
Halfaker et al. 2008	Collective effort model (no hypotheses)	Examination of the dropout rate of editors in two sets of 200K edits, reverted and non-reverted	Wikipedia edits; edits that remain untouched longer used as a proxy for quality of edit	1. Reverts discourage contributions, especially reverts made by veteran providers of new providers 2. Editors who continue to do work in Wikipedia after being reverted increase the quality of their work
Hu et al. 2007	None (mathematical model development)	Comparing the performance of measurement models through a series of experiments using 242 Wikipedia articles about various countries	Article quality as indicated by its links	1. Authors with more authority produce higher quality edits 2. Length of articles associated with their quality
Javanmardi and Lopes 2010	None (no hypotheses; builds on previous research)	Comparing the time featured and non-featured articles are in a high-quality state Conceptualizing revisions that last longer as high quality	Quality of featured and non-featured articles	1. The average article quality increases with more edits 2. Featured articles are of high quality 86% of the time, while non-featured are high quality 74% of the time

Kane 2010	None (deductive hypothesis development partially based on case study findings)	Qualitative case study of editing a featured article Quantitative analysis of factors associated with the promotion of 188 articles nominated to receive featured article status	Wikipedia article promoted to receive featured article status Virginia Tech shooting article (editing case study)	<ol style="list-style-type: none"> 1. The number of edits and editors is not associated with article promotion 2. The number of anonymous edits and breadth of top contributor experience negatively associated with article promotion 3. Revisions and knowledge-depth of top contributors positively associated with article promotion
Kittur and Kraut 2008	Theories of group coordination (no hypotheses)	Quantitative analysis of variables associated with 23,619 articles with changing quality scores	Wikipedia article quality score	<ol style="list-style-type: none"> 1. High concentration of the main work, with more contributors in supporting roles associated with change in quality score 2. Equal work distribution and coordination on the talk page is not associated with quality scores
Stvilia et al. 2007	None (grounded approach)	Case study with qualitative and quantitative techniques	Wikipedia editing activities across articles, talk, policy, and user pages	<ol style="list-style-type: none"> 1. When editors modify each other's contributions, it implicitly or explicitly involves an evaluation of the quality of those contributions 2. Self-selection allows significant savings in terms of selection and coordination costs
Wilkinson and Huberman 2007	None (no hypotheses)	Comparing editing activities in featured and non-featured Wikipedia articles	Wikipedia featured article status as a proxy of quality	<ol style="list-style-type: none"> 1. Correlation found between the number of edits and distinct editors and quality
Wöhner and Peters 2009	None (no hypotheses)	Algorithm-based comparison of the life cycles of 100 featured and good articles and 100 articles nominated for deletion	Wikipedia articles	<ol style="list-style-type: none"> 1. Identified 11 life-cycle metrics associated with quality used for the algorithm 2. Length of articles and high intensity in contribution activity is most associated with quality score

The overview in Table 4.1 also reveals a second shortcoming—namely, that current research on organizational mechanisms and quality in OC systems is not based on an integrative theoretical framework. With the exception of a few studies (Arazy et al. 2011; Arazy and Nov 2010; Kittur and Kraut 2008), almost all research examining quality in OC systems is exploratory in nature, rather than anchored in theory or based on theoretically informed hypothesis testing (Anderka et al. 2011; Blumenstock 2008; De la Calzada and Dekhtyar 2010; Halfaker et al. 2011; Hu et al. 2007; Javanmardi and Lopes 2010; Wilkinson and Huberman 2007; Wöhner and Peters 2009). This lack of theoretical anchoring makes it difficult to know how mechanisms relate to outcomes and how to consolidate findings from different studies. As a result, prominent researchers in the field call for studies that are more theory-informed to build new theoretical knowledge on OC

system governance beyond what we can know through empirical work or research applying theories established long before OC systems emerged (Arazy et al. 2011; Majchrzak 2009; Von Krogh et al. 2012).

To advance our understanding of quality management in OC systems, we draw on PIC theory (Mindel et al. 2018) to focus on resolution of content quality issues related to information pollution. The lens of PIC theory allows us to consolidate previous findings in the OC system quality literature under a single overarching theoretical framework. Moreover, zooming in on the information pollution issue promotes broader theoretical understanding of how systems characterized by low barriers to participation work to reduce errors, improve quality, and address biased as well as manipulated content.

4.3 Information Pollution in Polycentric Information Commons

Three decades ago, Orman (1984) identified information pollution as a major problem in the information age. Defined as *the contamination of information supply with incomplete, inconsistent, or irrelevant information* (pp. 64), Orman argued that information pollution will negatively impact organizations. This notion received initially little attention, yet there is a renewed theoretical interest in the phenomenon as it pertains to open-access platforms characterized by low barriers to participation such as social media, online communities and Wikis Drawing on the tragedy of the commons (Hardin 1968) and collective-action paradigms related to the governance of common access resource systems in the physical world (Ostrom 1990), Mindel et al. (2018) conceptualize open-access platforms online as “information commons”, and theorize that their long-term sustainability is threatened by information pollution.

PIC theory posits that when an unlimited number of often-anonymous participants are free to upload content online, it is almost inevitable that some contributions will be polluted (Mindel et al. 2018). Previously researched phenomena exemplifying information pollution include manipulated online reviews (Mayzlin et al. 2014), fake and heavily biased news circulating on social media (Allcott and Gentzkow 2017), buggy code (Gyimothy et al. 2015), and erroneous information uploaded to online forums and Wikis (Clauson et al. 2008; Kata 2012). Information pollution can be viewed as any type of content that is misaligned with the goal of the information commons (Mindel et al. 2018), including uploading factually accurate information about a certain topic to an information commons dedicated to an entirely different issue. Unchecked persistent pollution is likely to reduce the value of the information commons to those who appropriate its content; it can also discourage those who provide content from making further contributions. Because the long-term sustainability of information commons depends on the balance between continuous engagement and renewal of participants¹⁷ (Butler 2011; Ransbotham and Kane 2011), continuing pollution can potentially set in motion a downward spiral (Mindel et al. 2018).

Typically, pollution of open-access systems—such as a contaminated lake or a littered park—is addressed by paid workers who are employed by the government or subcontracted through a private firm. Similarly, government and private resources are increasingly being deployed to monitor information pollution online (Chen 2014; Kanter 2018). Still, the sheer scale of the information uploaded online daily makes it impossible for content monitors to address it all in an effective manner. Information commons¹⁸ such as Facebook, Twitter, and YouTube employ

¹⁷ PIC theory distinguishes between three broad categories of participant stakeholders: producers of the systems, providers of content, and appropriators of content.

¹⁸ According to PIC theory, any system online that relies on self-selecting content contributors who are free to participate or exit at any time may be viewed as an information commons, regardless of its ownership structure.

thousands of dedicated content monitors, but even these resource-rich systems find it difficult to keep up with information pollution (Chen 2014; Kanter 2018). Inspired by research on bottom-up, polycentric governance of physical local resource systems (Ostrom 1990), PIC theory suggests providing the community of users with the power and the tools to monitor pollution and address it themselves on the local level of the information commons.¹⁹

4.3.1 Operational Level of Governance

PIC theory points out that polycentric principles manifest at three different but closely related levels of governance: constitutional, collective-choice, and operational (Table 4.2, Figure 4.1). Although the tiers are distinct, they build on each other and influence each other's evolution through feedback loops. The constitutional level of governance defines the *overarching goal*²⁰ of the system through basic guidelines, terms-of-use rules, and the technical features put in place. Constitutional rules are set by the system developers—that is, the producers of the information commons. Collective-choice is the level of governance where participants—that is, content providers—influence the adoption and development of rules and features. The degree of collective-choice effectiveness depends on the system producers' inclusivity and receptiveness to feedback. In most information commons, providers may petition producers to evolve and change rules and features. In some PICs, such as Wikipedia, providers may even devise certain rules and features. The rules and features determined on the constitutional and collective-choice levels directly impact how information providers operationally engage with the system day to day. In turn, problems and phenomenon encountered on the operational level may further inform action on the collective-

¹⁹ We address the downsides of allowing participants more power to monitor each other in the “Discussion” section.

²⁰ Information pollution, according to PIC theory, is information that departs or conflicts with the information commons' overarching goal. For example, uploading fiction to an information commons dedicated to creating encyclopedic information would be considered pollution.

choice level. In some cases, collective-choice may influence the adjustment of core constitutional policies and features (Mindel et al. 2018).

<i>Level</i>	<i>Definition</i>	<i>Examples from Wikipedia</i>
Constitutional	The core policies and technical features set by the information commons producers that create the conditions for polycentricity	<ul style="list-style-type: none"> • The constant availability of live-editing function to all users • The five pillars (Wikipedia's constitutional rules)
Collective-choice	The policies, rules, and technical features created or influenced by the information commons providers	<ul style="list-style-type: none"> • Tens of key policies and guidelines set by Wikipedia's providers²¹ • Thousands of templates created to tag problems in articles • Arbitration mechanisms to resolve disputes
Operational	The day-to-day information provision activities based on the rules and features set by producers on the constitutional level and by peer providers on the collective-choice level	<ul style="list-style-type: none"> • Hundreds of millions of edits to articles • Hundreds of thousands of postings of template tags to indicate a problem with an article • Hundreds of thousands of postings of links to guidelines and policies in comments on edits

In Wikipedia, the five fundamental pillar rules²² represent the constitutional level of polycentric governance granting the community of participants the freedom of collective-choice in determining the rules that govern most operational details. The collective-choice rules created by the community, in turn, directly and indirectly guide the everyday editing of articles and associated talk pages. An example of the nesting of polycentric governance is the constitutional rule, “Wikipedia is written from a neutral point of view,” which inspires the collective action responsible for creating templates for flagging biased content that, when posted, inspire operational actions such as deletion of biased content by self-directed providers.

²¹ https://en.wikipedia.org/wiki/Wikipedia:List_of_policies_and_guidelines

²² (1) Wikipedia is an encyclopedia. (2) Wikipedia is written from a neutral point of view. (3) Wikipedia is free content that anyone can use, edit, and distribute. (4) Wikipedia's editors should treat each other with respect and civility. (5) Wikipedia has no firm rules. (see https://en.wikipedia.org/wiki/Wikipedia:Five_pillars)

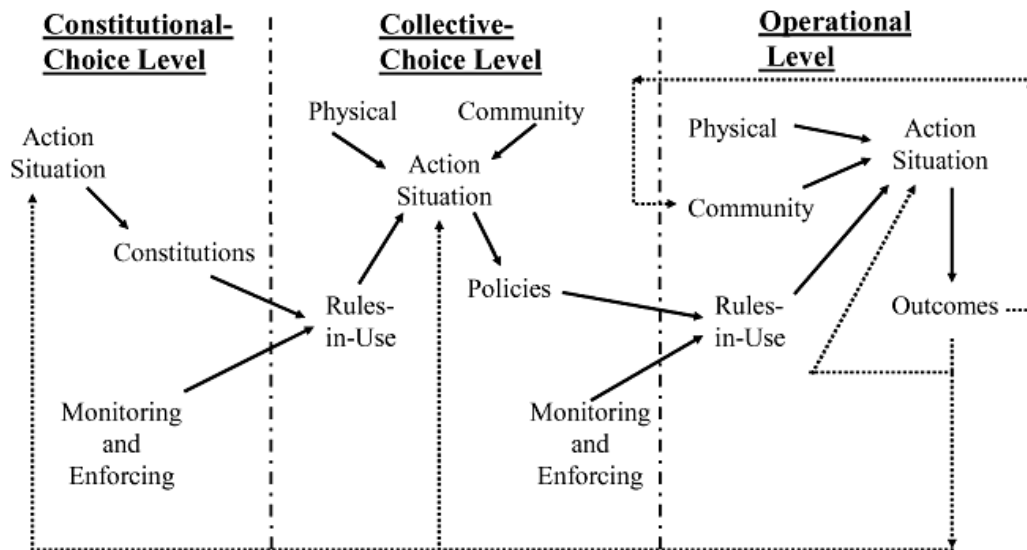


Figure 4.1 Ostrom's Levels of Governance²³

At present, PIC theory addresses the constitutional and collective-choice levels of polycentric governance of information commons by focusing on greater inclusivity in decision-making on rules and technical features for boundary setting, peer monitoring, modifications, and peer recognition. In our research, we draw on PIC theory and expand it to include the operational level of governance by examining how polycentricity impacts pollution resolution when it arises during content development.²⁴

4.3.2 Polycentric Principles

Polycentric governance evolves incrementally through collective-choice processes; this contrasts starkly with traditional organization governance, which is deployed top-down on subjects at lower levels of the hierarchy. A high degree of collective-choice—the notion that the community sets the rules through consensus—is found to be associated with the long-term success of local arrangements governing shared resource systems such as grazing grounds, fisheries, and forests

²³ Adapted from Ostrom et al (1994, p. 47).

²⁴ “Action Situation” on the operational level.

(Ostrom 1990). Correspondingly, PIC theory asserts that a multitude of independent centers of decision-making operating within the frame of an overarching goal is the key to the long-term sustainability of information commons (Mindel et al. 2018). The theory also asserts that open content provision is a double-edged sword—potentially generative and degenerative at the same time—and that an information commons should ideally incorporate four polycentric principles in its governance to reduce the prevalence and impact of threats arising from the system’s open nature.

First, an information commons should not solely rely on paid outsiders for monitoring, but rather should allow the community to police itself as much as possible. Defined as “*the extent to which rules and features afford peer monitoring and gradual sanctioning to support appropriate behavior and dispute resolution in an information commons*” (Mindel et al. 2018, p. 619), the principle of shared accountability aims to create a sense of ownership toward the information commons among participants. The principle of shared accountability suggests that the more the community actively monitors against those who break the information commons boundaries, the more it will be able to serve its overarching goal. Focusing on resolving information pollution on the *operational level*, shared accountability can manifest in different ways, including flagging of polluted content or disruptive users, direct removal of content, and community-based arbitration mechanisms to resolve disputes. Shared accountability works to reduce the cost of monitoring against information pollution and is predicated on the idea that, when participants in the information commons share the task of maintaining its integrity, they are more likely to resolve pollution effectively.

Second, an information commons should allow participants a high degree of freedom in determining the rules on content and conduct boundaries. Defined as “*the extent to which rules*

and technical infrastructure features afford information provision and appropriation consistent with the information commons' overarching goal" (Mindel et al. 2018, p. 619), the principle of boundary regulation helps address collective-action threats by increasing the community's sense of the content and conduct that should characterize the information commons. Every information commons has a different overarching goal, which necessitates context-specific boundaries (Ren et al. 2007). The principle of boundary regulation suggests that a high degree of inclusion of participants in setting rules will lead to higher consensus and a greater likelihood that the boundaries will be accepted and followed. Resolving pollution on the *operational level* requires a certain level of awareness of the boundaries—that is, the rules on content and acceptable conduct—by participants. In most cases, boundary regulation occurs when participants simply communicate the rules to those who seem unaware of them; in Wikipedia, for instance, participants might ask someone not to post content about dogs in an article dedicated to cats, or cite the policy against original research in article editing when that issue arises. In other instances, core participants with administrative authority may issue more-official warnings and may even ban those who insist on ignoring the information commons' boundaries.

Third, an information commons should avoid sudden top-down dictated changes that might upset participants and instead incrementally adapt in a bottom-up fashion. Defined as "*the extent to which changes in infrastructure and rules are gradually introduced and providers and appropriators are actively involved in shaping them*" (Mindel et al. 2018, p. 619), the principle of incremental adaptation aims to strike a balance between the need to adapt and sensitivity to participants' needs and opinions. The principle of incremental adaptation suggests that those who are most closely involved with the everyday operational dynamics of the information commons are also more likely to provide constructive feedback for its adaptation. Moreover, it suggests that

smaller changes are more likely to be positively received, as opposed to sudden extensive changes that may alienate many participants. Incremental adaptation also reduces the cost of experimenting with changes. If a certain change is only gradually different from the previous version and is well received, it can be quickly incorporated and further adapted. On the other hand, if it is not well received, it is easy to replace it with another incremental adaptation. On the *operational level*, incremental rather than extensive modification of OSS code or wiki content is more likely to lead to output that is more balanced, contains fewer errors, and gains more consensus.

Lastly, because information commons depend on volunteers for content, they should create a positive environment for those actively participating. Defined as “*the extent to which providers are acknowledged by peers, appropriators, and producers*” (Mindel et al. 2018, p. 619), the principle of *provider recognition* aims to enhance the personal and social experience for participants. The provider recognition principle suggests that a volunteer workforce must be motivated to continuously engage with the information commons. On the *operational level*, provider recognition may be manifested in the form of direct recognition for high-quality work, as well as in the awarding of virtual status and popularity symbols such as avatars, icons, likes, smiley faces, and any other signal of positive acknowledgment.

4.4 Study Design and Hypotheses

In our research, we focus on pollution events in one of the most influential OC systems online: the English Wikipedia²⁵ (Fallis 2008; Tapscott and Williams 2008). Wikipedia is an archetype of polycentricity as its governance has almost entirely been devised by its community members (Aaltonen and Lanzara 2015; Butler et al. 2008; Viegas et al. 2007). The website is a dynamic

²⁵ Appendix A provides details on data.

information repository that is constantly updated and improved through the work of numerous volunteers operating independently outside of a centralized organizational order. Prior research shows that although Wikipedia contains inaccurate, vandalized, biased, and incomplete information (Holman Rector 2008; Kupferberg and Protus 2011; Lavasa et al. 2011), it is for the most part considered to be a high-quality source of encyclopedic information (Brown 2011; Chesney 2006; Giles 2005). As such, it may provide important lessons regarding pollution resolution during content development for other information commons characterized by low barriers to participant entry and exit.

4.4.1 Dependent Variable

Wikipedia’s overarching goal is to provide encyclopedic information. The articles should be accurate, complete, relatively concise, and stick to reporting facts—any information that fails to align with these criteria and overall encyclopedic aims can be viewed as polluted.

To govern article development, Wikipedia’s users have devised, through a process of collective-choice, a multitude of templates that can be placed on article pages to signal to readers and the community of information providers that something is not right (Anderka et al. 2011). At present, Wikipedia has 195 such major templates and hundreds of subtemplates; one of the more familiar templates is the general cleanup template, which signals that an article is in suboptimal shape and needs intensive attention to resolve issues with its content (Figure 4.2). When an article is tagged for cleanup, a message box appears on the top of the article page stating that the article “may require cleanup to meet Wikipedia’s quality standards.”



Figure 4.2 Wikipedia’s Cleanup Tag

The guidelines for cleanup template usage encourage those who post it to specify the parts of the article requiring attention, but often the posted templates are simply general notices that caution readers and serve as a call for action aimed at prompting potential providers to improve the state of the article. The posting of the cleanup template also automatically adds the article to a list on the “Wikiproject cleanup” page, where information providers interested in cleanup tasks can see it. Data shows that a cleanup template can remain on an article page for months and even years before an information provider makes the evaluation that it is no longer needed and removes it. Based on the premise that timely pollution resolution is important for stabilizing article quality while reducing its adverse impact on trust among appropriators, in our analysis, we focus on the period between the posting and removal of the general cleanup template. Hence, the unit of analysis is *pollution event* and the outcome variable of interest is the *time to resolve the event* as an indication of effective resolution of information pollution.

4.4.2 Independent Variables

PIC theory asserts that polycentric governance helps reduce pollution on average; the challenge is to identify how polycentricity manifests in the context of an individual information commons and how different polycentric governance principles work in detail (Mindel et al. 2018). Wikipedia is a behemoth encompassing multiple tiers of governance (Forte et al. 2009; Morgan and Zachry 2010). However, because our objective is to identify polycentric governance on the operational level of article editing, we focus on identifying polycentric practices and examining their impact on resolution time during pollution events.

The posting of a cleanup template represents an act of shared accountability in which an information commons member cares enough about the site’s overarching goal to take the extra

step to warn readers and spur action by the community. While in most information commons the act of shared accountability ends with flagging (after which an algorithm or an employed content monitor will address the issue), in Wikipedia, the community of participants determines the state of the content. Hence, shared accountability continues after the posting of a cleanup template through shared efforts to edit the article until the cleanup signal is deemed unnecessary. As such, we consider the work distribution during the pollution event to measure the extent to which the community shares the responsibility of resolving it. Prior research finds that relatively few providers are responsible for the bulk of the work in OC systems (Kittur and Kraut 2008; Von Krogh et al. 2013), and PIC theory suggest that too many participants active at the same time can create congestion (Mindel et al. 2018). Building on prior research and PIC theory, we therefore assert that the extent to which work is proportionally distributed among a few productive providers determines the speed with which pollution resolution is achieved. Hence, we hypothesize:

H1: *Shared accountability in terms of proportionate work distribution in relation to number of providers reduces the time it takes to resolve pollution events.*

Being an encyclopedia, each Wikipedia article is naturally bounded by its topic and the site's overarching goal of producing accurate, complete, well written, relatively concise, and neutral information. Over the years, the Wikipedia community has devised multiple editorial policies to guide the boundary of articles (Butler et al. 2008; Morgan and Zachry 2010). Examples of these policies include guidelines calling for neutrality, verifiability, and avoiding the use of copyrighted material and original research. Prior examination of the prevalence of policy citations calls for more research on the relationship between evocation of policies and consequential collective actions in OC systems (Beschastnikh et al. 2008). As such, we draw on PIC theory to assert that policy citation in article editing represents an act of boundary regulation.

In the same way that evoking rules and policies in the physical world often represents an act of boundary regulation in a shared resource system (for instance, the highway patrol officer cites the law when enforcing the boundaries of acceptable conduct on the road), evoking policies in information commons represents an act of boundary regulation. In Wikipedia, we therefore consider the number of policies cited in edits during the pollution event to measure the level of boundary regulating practices. We expect increasing evocation of policies to initially slow down pollution resolution before speeding it up. We further explain why. Policies in Wikipedia are developed by self-selecting policymakers and determined through consensus on the collective-choice level. Most Wikipedia editors do not take part in this process and are likely unaware of the policies guiding content boundaries. An initial increase in citation of rules during pollution events thus indicates misalignment between editing activities and the set standards and boundaries on content. However, as policies are continuously cited, editors—that is, information providers—increasingly get on the “same page” which leads to pollution resolution.²⁶ Drawing on PIC theory on the role of boundary regulation during pollution resolution, we predict that, after reaching a threshold, policy citation will help stabilize boundaries which, in turn, will help reach pollution resolution.

H2: *Boundary regulation through policy citations increases pollution resolution time up to a point and reduces it thereafter.*

Wiki technology predicated on the idea that unconstrained individual users can make numerous modifications conveniently through their web browsers (Leuf and Cunningham 2001). This opportunity to implement incremental changes gives rise to a constantly growing body of digital artifacts. As a result, every aspect of Wikipedia is modified in an incremental manner, including

²⁶ This is a little like getting teenagers to clean after themselves through continuous reminders. At first, they might not listen and keep making a mess, but with enough nagging they are more likely to clean.

its rules and policies pages, templates, projects, and, of course, the article pages themselves. Each discrete edit is an adaptation that slightly modifies the article from its previous form. In our data, we measure the number of article revisions during the pollution event. Our expectation is that more revisions will lead to faster pollution resolution. Drawing on PIC theory, we hypothesize:

H3: *Incremental adaptation through revisions reduces the time it takes to resolve pollution events.*

Much has been written on what motivates Wikipedia participants to spend their time and energy editing articles for no pay. Researchers have identified diverse motivations that can be loosely grouped into reasons of personal and social fulfillment that are partially intrinsic and partially driven by peer recognition (Kuznetsov 2006; Nov 2007; Yang and Lai 2010). Wikipedia has a few mechanisms to signal peer recognition for work on the collective-choice level, including “barn stars”—virtual medals awarded by peers (Kriplean et al. 2008)—and assignment of administrator privileges (Burke and Kraut 2008). On the operational level of article editing, however, Wikipedia does not have a peer-recognition mechanism; that is, individual edits do not receive “likes” or “smiley faces.” On the other hand, there is a peer-rejection mechanism in place.

As part of its effort to combat vandalism of articles, Wikipedia has a revert mechanism to restore the article to its previous, pre-vandalized form. While the revert mechanism has proven instrumental for quick resolution of vandalism (Priedhorsky et al. 2007), it is also known to be abused at times (Sumi and Yasseri 2011). Instead of reverting to the previous version to combat vandalism, some providers revert articles to a previous version simply because they like it better. By doing that, providers essentially disregard the effort of their peers. Thus, while we do not have a valid way to operationalize provider recognition, we suggest that reverts of multiple revisions represent a form of *provider rejection*. PIC theory asserts that provider recognition is positively

associated with pollution resolution. Using a reversed logic, we theorize that provider rejection is likely to have the opposite effect. Because the revert mechanism is intended to reverse vandalism and low-quality edits, its use is likely to be associated with faster pollution resolution at first. However, drawing on the logic of PIC theory, we hypothesize that persistent rejection of contributions is ultimately likely to lead to a slowdown in pollution resolution.

H4: *Provider rejection through multiple revision reverts reduces pollution resolution time up to a point and increases it thereafter.*

To recap, PIC theory predicts that the more that independent participants share the responsibility of incrementally adapting an article—while abiding by the community-set rules and avoiding rejecting each other’s efforts—the faster pollution will be addressed. Table 4.3 summarizes the main constructs and variable operationalization.

Construct	Definition	Variable	Operationalization	Unit
Pollution	The extent to which information fails to align with the overarching goal of the information commons	Pollution resolution time (DV)	The period between the posting of a cleanup template and its removal	<i>Days</i>
Shared Accountability	The extent to which rules and features afford peer monitoring and gradual sanctioning to support appropriate behavior and dispute resolution	Work distribution in relation to information providers (IV1)	Work distribution equality between the posting of a cleanup template and its removal	$\frac{1 - GINI^{27}}{\# \text{ of editors}}$
Boundary Regulation	The extent to which rules and technical infrastructure features afford information provision and appropriation consistent with the information commons’ overarching goal	Policy citations (IV2)	Number of policies cited in edit comments between the posting of a cleanup template and its removal	<i>Links to policies</i>

²⁷ Measure of equality ranging from 0 to 1 where 0 indicates complete equal work distribution while 1 indicates a complete unequal work distribution.

Incremental Adaptation	The extent to which changes in infrastructure and rules are gradually introduced and providers and appropriators are actively involved in shaping them	Revisions (IV3)	Number of edits made to the article between the posting of a cleanup template and its removal	<i>Discrete edits</i>
Provider Rejection	The extent to which providers feel rejected by peers	Reverted revisions (IV4)	Number of reverts of revisions between the posting of a cleanup template and its removal	<i>Reverted revisions</i>

4.4.3 Control Variables

Previous research assessing projects in OC systems typically controls for a project's level of popularity and completeness (Ransbotham and Kane 2011). Mature projects or projects that naturally draw more participants are likely to exhibit different patterns and dynamics in comparison to other, less developed and more peripheral projects. In our analysis, we use five control variables indicative of article completeness and popularity among providers prior to the pollution event. We control for article maturity by accounting for its length in characters and its age prior to the posting of the cleanup template. We control for the level of the article popularity and activity by accounting for the number of unique providers making edits to it prior to the posting of the cleanup template. We control for the number of references an article has at the time when pollution event begins because it partially indicates the article maturity and quality state prior to the posting of the cleanup template. We control for the article history of explicit coordination (Kittur and Kraut 2008) by accounting for the number of unique participants on the talk page prior to the posting of the cleanup template.

4.5 Dataset Construction

We used a freely available English Wikipedia database dump from 1 January 2017. It would be computationally very demanding to sample pollution events directly from the dump; instead, we first took a large random sample of 340,000 articles and identified 5,487 pollution events in them

using regular expression patterns. The identified events appeared on 4,679 different articles; i.e. some articles contained more than one pollution event (see Table 4.4).

Number of pollution event	<i>N</i> (pages)	Mean (days)	SD (days)	Min (days)	Max (days)
<i>1</i>	4,020	520.66	718.98	0 ²⁸	3,896.59
<i>2</i>	519	356.14	507.36	0	2,841.33
<i>3 or more</i>	119	188.8	361.23	0.001	2,770.88
Pollution event order					
<i>First</i>	4,679	478.12	689.85	0	3,896.59
<i>Second</i>	639	421.9	563.4	0	2,841.33
<i>Third or more</i>	169	274.37	457.47	0.001	2,770.88

Because an event appears in exactly one article, every event has an equal probability of being sampled and our base sampling strategy is unbiased. There are, however, three potential sources of selection bias that emerged from the way we further processed the data.

First, the current version of our Python tools could not process three very large articles (more than ~2GB data); the article topics were “Barack Obama” (page id 534366), “Syrian Civil War” (page id 30741795), and “Adolf Hitler” (page id 2731583). Given the topics of the three omitted articles, it is not surprising that they had attracted many edits. Although these articles may also have contained pollution events, we expect the potential bias introduced by the omission of these articles to be negligible. If each omitted article contained 10 pollution events (vs. the maximum of 8 events per article in the remaining sample), we still would have omitted only 0.5 % of the randomly sampled events.

²⁸ Some articles receive the cleanup template right at their birth to prompt providers to improve them.

Second, to reduce heterogeneity, we limited our analysis to 4,325 pollution events ranging in duration from 24 hours to 1,500 days.²⁹ Cleanup templates that appear for less than 24 hours on an article either have a high likelihood of being posted by mistake or did not reach consensus for being required. At the other extreme, we find that exceptionally long pollution events lasting more than 1,500 days are mostly on outlier articles covering obscure topics. Such articles typically have low levels of editorial activity.

Third, some articles contained more than one pollution event, which raised a question as to whether the two (or more) events on the same article were independent. To address this, we ran a series of robustness analyses in which we tested our model on articles with a varying number of pollution events and their varying order.

We also retrieved data for 172,576 talk pages corresponding to article pages in the sample (though not all articles have a talk page) to examine communications and coordination activities during pollution event. However, due to the relatively low intensity of activity on talk pages during pollution events (addressed in the “Descriptive Statistics” section), we decided not to pursue this angle further. Appendix A offers more details about the data processing.

4.6 Analysis Results

4.6.1 Descriptive Statistics

It takes a relatively long time to resolve an information pollution event in Wikipedia; the mean pollution event lasted 324 days (SD 370.2 days). Although some articles in Wikipedia are “classic” encyclopedic topics, most address topics that are less familiar to the general population, and thus

²⁹We follow the logic of past researchers of Wikipedia article quality, who often bound the analysis to avoid comparing pages significantly different in their stage of development; some, for instance, removed “stub” pages (Wilkinson and Huberman 2007).

require the attention of specific editors knowledgeable about the subject. We believe that this disparity in topic familiarity and the number of potential contributors knowledgeable of it is behind the skewed distribution and the high standard deviation in pollution resolution time.

Variable	Mean	S.D.	Min	Max
Pollution event length (days)	324	370.2	0 ³⁰	1,499.2
Article length at start of pollution event (characters)	10,049.6	16,283.6	0	207762
Article age when pollution event starts (days)	845.2	996	0	5,353.5
Number of providers active before pollution event	43.1	108.4	3	2,160
Number of providers active before pollution: talk page	5.8	17.8	0	640
Number of providers active during the pollution event	13.3	19.4	1	301
References before pollution event	4.9	18.4	0	485
Article length change (characters)	2,662.3	6,424.7	0	151,191
Policy citations during the pollution event	5.3	13.1	0	301
Revisions	43.5	94.2	0	1,854
Work distribution (GINI)	0.63	0.18	0	0.97
Reverted revisions	6.7	25.5	0	656
Edit distance (net of all added/subtracted characters)	8,290.7	20,582	2	456,359

An average of 13.3 providers made 43.5 revisions on average during a pollution event. The providers added and deleted a net average of 8,290.7 characters, which ultimately resulted in a change of 2,662.3 characters before the event was resolved. An average of 5.3 revisions during the event contained a reference to Wikipedia policy. Consistent with previous findings (Kittur and Kraut 2008) work distribution is fairly concentrated with a GINI coefficient of 0.63. We calculated work concentration as each article's proportion of edit-distance (total characters added or removed) by each unique editor during the pollution event. The high GINI score indicates that a disproportionate amount of work was carried by a small number of editors.

Past studies of work dynamics in OC systems are conflicted on the importance of explicit coordination; researchers find evidence that explicit coordination helps to improve outcomes

³⁰ Some articles receive the cleanup template right at birth to prompt providers to improve them.

(Arazy and Nov 2010), but also that a disproportionate amount of work is carried out independently without any back-and-forth communication among participants (Kittur and Kraut 2008). In our data, we find that almost all editing activities during pollution events are largely devoid of explicit coordination via Wikipedia’s talk pages. During the considered pollution events, an average of only 1.8 providers participated on the talk page, making 4.7 revisions to it and citing on average only 0.06 policies.

Examination of pairwise correlations (Table 4.6) shows a close association between three of the four variables of polycentric practices: policy citations and revisions ($r = .83$), policy citations and reverted revisions ($r = .87$), and revisions and reverted revisions ($r = .87$). PIC theory asserts that, in an ideal scenario, polycentric governance practices complement each other (Mindel et al. 2018), and we are not surprised to see the strong correlation.

Table 4.6 Correlations

	1	2	3	4	5	6	7	8	9	10	11	12
1 Pollution event length												
2 Policy citations	.30**											
3 Policy citations (squared)	.10**	.82**										
4 Revisions	.22**	.83**	.63**									
5 Reverted revisions	.13**	.87**	.76**	.87**								
6 Work distribution (GINI)	.29**	.32**	.12**	.41**	.26**							
7 Edit distance	.12**	.45**	.28**	.63**	.42**	.37**						
8 Article age	.15**	.26**	.10**	.20**	.18**	.17**	.15**					
9 Length of article	.30**	.30**	.16**	.36**	.27**	.26**	.46**	.42**				
10 Providers during pollution	.30**	.83**	.56**	.92**	.81**	.45**	.55**	.26**	.33**			
11 Providers before pollution	.06**	.35**	.20**	.35**	.35**	.19**	.25**	.50**	.61**	.38**		
12 Providers before: talk page	-.05	.27**	.16**	.28**	.27**	.15**	.24**	.39**	.56**	.30**	.87**	
13 References before pollution	-.02	.19**	.10**	.17**	.14**	.11**	.31**	.29**	.65**	.17**	.47**	.45**

* $p < .05$; ** $p < .001$; *** $p < .0001$

4.6.2 Model Specification and Analysis Procedures

We specify our baseline model as follows:

$$PEL = \beta_0 + \beta_1 Controls + \beta_3 SHA + \beta_4 BR + \beta_5 BR^2 + \beta_6 INA + \beta_7 PRJ + \beta_8 PRJ^2 + \varepsilon \text{ (Equation 1)}$$

In the model, *PEL* is the pollution event length, *SHA* is shared accountability, *BR* is boundary regulation, *INA* is incremental adaptation, and *PRJ* is provider rejection. Equation 1 includes the three polycentric practices (*SHA*, *BR*, and *INA*). However, we had to revise Equation 1's specification because of the high colinearity between reverted revisions, policy citations, and revisions. Specifically, we follow a two-step process to estimate the unique portion of *PRJ* (*U_PRJ*) that is not overlapped with *BR* and *INA*. We first use Equation 2 to estimate the portion of *PRJ* that is overlapped with *BR* and *INA* (*O_PRJ*). We then use Equation 3 to compute the unique portion of *PRJ* (*U_PRJ*) that is not overlapped with *BR* and *INA*:

$$O_PRJ = \beta_0 + \beta_1 BR + \beta_2 INA \text{ (Equation 2)}$$

$$U_PRJ = PRJ - O_PRJ \text{ (Equation 3)}$$

Next, we replace *PRJ* with *U_PRJ* in Equation 1 to get the following model:

$$PEL = \beta_0 + \beta_1 Controls + \beta_3 SHA + \beta_4 BR + \beta_5 BR^2 + \beta_6 INA + \beta_7 U_PRJ + \beta_8 U_PRJ^2 + \varepsilon \text{ (Equation 4)}$$

We use hierarchical OLS regression analysis to estimate the model, adding variables in the following order: (1) the control variables, (2) shared accountability, (3) boundary regulation, (4) squared value of boundary regulation (*BR*²) to test for the hypothesized curvilinearity, (5) incremental adaptation, (6) standardized residual of provider rejection, and (7) standardized residual of provider rejection squared (*PRJ*²).

4.6.3 Results

Equation 1's model explains 40.5% of the variance in pollution event length (where 11.9% is explained by the control variables). Each individually added polycentricity variable uniquely

contributes to the variance explained (Table 4.7). The variance inflation factor (VIF) values are less than 5, suggesting that multicollinearity is not a major issue.

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Article age	.314***	.303***	.283***	.279***	.262***	.238***	.215***
Length of article	.065***	.007	-.088***	-.068***	-.052***	-.058***	-.065***
Providers before pollution event	-.521***	-.550***	-.541***	-.650***	-.654***	-.600***	-.545***
Providers before: talk page	.126***	.125***	.019	.100***	.109***	.109***	.029*
References before pollution	-.025**	.005	.075***	.054***	.043***	.030***	.055***
Work distribution / providers (SHA)		-.274***	-.206***	-.161***	-.172***	-.187***	-.135***
Policies cited in revisions (BR)			.375***	.853***	.972***	.964***	1.632***
Policies cited in revisions ² (BR ²)				-.493***	-.500***	-.490***	-1.17***
Revisions (INA)					.262***	-.145***	-.148***
Reverted Revisions (U_PRJ)						-.241***	-.447***
Reverted Revisions ² (U_PRJ ²)							.347***
Model F	774***	1,115***	1,744***	1,719***	1,545***	1,779***	1,775***
F change		2,487***	4,477***	1,082***	107***	2,615***	1,073***
Adj. R ²	0.119	0.189	0.298	0.324	0.334	0.382	0.405
ΔR ²		0.07	0.109	0.025	0.01	0.056	0.022

* $p < .05$; ** $p < .001$; *** $p < .0001$

H1, which predicted that shared accountability (proportionate work distribution in relation to number of providers) will reduce the length of pollution events, is supported ($\beta = -.135$, $p < .0001$). H2, which predicted that boundary regulation (policy citations) will be associated with a longer pollution event to a point and a quicker pollution resolution thereafter, is supported. Policy citations are associated with an increase in the length of pollution events ($\beta = 1.632$, $p < .0001$) and policy citations squared with a decrease in their length ($\beta = -1.1705$, $p < .0001$). H3, which predicts that incremental adaption (number of revisions) will decrease the length of pollution events, is supported ($\beta = -.148$, $p < .0001$). H4, which predicted that provider rejection (reverted revisions) will be associated with a shorter pollution event length to a point and a longer pollution event resolution thereafter, is supported.

We find all control variables to be significantly associated with pollution event length. We find that age of articles is associated with increased pollution event length ($\beta = .215, p < .0001$), and article length is associated with a decrease in pollution event length ($\beta = -.065, p < .0001$). The number of providers involved with an article prior to the posting of the cleanup template is associated with decreased pollution event length ($\beta = -.545, p < .0001$). The number of providers involved on the article talk page prior to the posting of the cleanup template is associated with increased pollution event length ($\beta = .029, p < .05$). The number of references prior to the posting of the pollution tag is associated with an increase in pollution event length ($\beta = .055, p < .0001$). These results suggest that more-mature articles with a history of high provider involvement are more likely to resolve any arising pollution faster.

4.6.4. Robustness and Additional Analyses

We conducted several robustness tests and additional analysis related to the impact of five key factors: the variance in the number of pollution events across articles, the variance in the order of pollution events, provider rejection, talk page activity, and veteran vs, new providers.

Number of Pollution Events

To test for the consistency of our hypothesized model across articles with a varying number of pollution events, we conducted three analyses. First, we tested the model with articles that had only a single pollution event ($N=4,020$). As Table B1 in Appendix B shows, the model results are consistent with our main analysis of pollution events from articles with varying numbers of pollution events. Second, we tested the model with articles that had two pollution events (519 articles, $N=1,038$). As the model results (Table B2, Appendix B) show, incremental adaptation (revision) is no longer statistically significant in relation to pollution resolution time. Finally, we tested the model on articles containing three or more pollution events (119 articles, $N=408$). As

Table B3 in Appendix B shows, incremental adaption (revision) is positively associated with pollution event length, suggesting that it *slows down* pollution resolution. The results of these analyses suggest that incremental adaptation (revision) helps resolve pollution faster in articles with a single pollution event, but that this association reverses itself in articles with three or more pollution events.

Order of Pollution Events

To test for our hypothesized model's consistency across pollution events based on their order, we conducted three analyses. First, we tested the model across all first events in our sample ($N=4,679$). As Table B4 in Appendix B shows, the direct relationship between boundary regulation and pollution resolution time is statistically insignificant. In all other respects, the model is consistent with the main model. Second, we tested the model across all second events ($N=639$). The model results (Table B5, Appendix B) are consistent with the main model, in which we do not distinguish between pollution event order. Finally, we tested the model across the third or more pollution event ($N=169$). As Table B6 in Appendix B shows, these results are inconsistent with our main model: (1) We detect no significant relationship between incremental adaptation and pollution resolution time. (2) We found no curvilinear relationship between provider rejection and pollution resolution time. This is slightly different from our research model findings. However, it is important to note that 97 percent of pollution events in the random sample has two or less pollution events; having a third or more pollution event is rather an atypical occurrence.

Provider Rejection

In the main specification, we incorporated U_PRJ and its squared term in the model specification (Equation 1), as PRJ is highly correlated with boundary regulation (policy citations) and

incremental adaptation (revision). By way of robustness, we evaluated a model with controls and only *PRJ* and *SHA* as per the following equation:

$$PEL = \beta_0 + \beta_1 Controls + \beta_2 PRJ + \beta_3 PRJ^2 + \beta_4 SHA + \varepsilon$$

We used a three-step hierarchical OLS regression analysis to estimate the equation, adding variables in the following order: 1) control variables, (2) provider rejection, and (3) squared value of provider rejection to test for the hypothesized curvilinearity.

The model explains 24.3% of the variance in pollution event length (where 11.9% is explained by the control variables). The addition of provider rejection and provider rejection squared uniquely contributes to variance explained (Table 4.8). The VIF values are less than five, suggesting that multicollinearity is not a major issue.

Variable	Model 1	Model 2	Model 3	Model 4
Article age	.314***	.329***	.337***	.335***
Length of article	.065***	-.023**	-.020*	-.109***
Providers before pollution event	-.521***	-.536***	-.609***	-.549***
Providers before: talk page	.126***	.046**	.114***	.057***
References before pollution	-.025**	.046***	.029***	.061***
Reverted revisions (<i>PRJ</i>)		.274***	.528***	.263***
Reverted revisions ² (<i>PRJ</i> ²)			-.274***	-.072***
Work distribution / providers (<i>SHA</i>)				-.276***
Model <i>F</i>	773.6	1,044.7	954.2	1,156.1
<i>F</i> change		2,115.1	337.9	2,085.2
Adj. <i>R</i> ²	.119	.179	.19.8	.243
ΔR^2		.06	.019	.045

* $p < .05$; ** $p < .001$; *** $p < .0001$

Our results also support H4, which predicted that provider rejection (reverted revisions) will reduce the length of pollution events to a point and increase it thereafter. Provider rejection is associated with an increase in pollution event length ($\beta = .263$, $p < .0001$), and provider rejection

squared is associated with a decrease in pollution event length ($\beta = -.072, p < .0001$). Apart from references before a pollution event, the control variables are statistically associated with pollution event length.

Talk Pages

Previous research shows mixed results regarding the importance of explicit coordination; researchers have found evidence that explicit coordination helps to improve outcomes (Arazy and Nov 2010; Viegas et al. 2007), but also that a disproportionate amount of work is carried out in relative isolation, without communication between contributors working on the same article (Kittur and Kraut 2008). In our data, we find that almost all editing activity during pollution events is devoid of explicit coordination through the talk page. An average of only 1.9 providers (SD 3.5) participated on the talk page during an event, making an average of four revisions (SD 16.9) to it and citing on average only 0.05 policies (SD 0.6). We performed a hierarchical OLS hierarchical analysis adding variables in the following order: (1) control variables, (2) polycentric governance variables, and (3) talk page variables (see Table B7, Appendix B). We find that the addition of talk page variables adds only modestly to variance explained (adjusted R^2 changes only 1.7 %) and that our hypotheses still hold when we include the talk page variables.

Veteran and New Providers

We further examined how work distribution between veteran and new providers impacts pollution resolution. Consistent with previous research (Pancieria et al. 2009; Shah 2006), we find that greater involvement of veterans is associated with faster pollution resolution. Interestingly, however, we find that most of the work during pollution events is shouldered by providers new to the article rather than by previous contributors to it. Approximately 85% of providers active during the pollution event were not active on the article prior to the cleanup template posting; further,

new providers were responsible for approximately three-fifths of the total aggregate edit distance during the pollution event. This indicates that the cleanup template acts as a call for action and is more effective at injecting new blood into the article than in rousing existing providers to improve their work. We again performed a hierarchical OLS analysis adding variables in the following order: (1) control variables, (2) polycentric governance variables, and (3) variable on veteran providers. The veteran provider variables are: (a) the ratio of editors who remain active (*veteran* editors of the article) in relation to all editors who were active before the pollution event; (b) the number of veterans; (c) the ratio of veterans to all editors active during the pollution event; and (d) the proportion, in characters, of the length change that veterans made. As Table B8 in Appendix B shows, we find that the involvement of veteran providers reduces the length of pollution events, but including those variables adds only modestly to variance explained (adjusted R^2 changes 2.6%). Again, our hypotheses hold when considering veteran's activities.

4.7 Discussion

Our empirical analysis supports the general assertion that polycentric governance practices influence pollution resolution in Wikipedia. Specifically, we find that shared accountability and incremental adaptation are linearly associated with faster pollution resolution, while boundary regulation and provider rejection have more complicated relationships. Based on these findings, we first bring together theoretical contributions to the literature on addressing quality in OC systems (Arazy and Nov 2010; Chesney 2006; Giles 2005; Kane 2010; Kittur and Kraut 2008; Wöhner and Peters 2009). We then move to discuss the findings in light of PIC theory (Mindel et al. 2018) and the implications of our findings for resolving pollution in OC systems and other online platforms with low barriers to participation.

4.7.1 Resolving Information Pollution in Open Collaboration Systems

Our analysis shows how shared accountability, content boundary regulation, incremental adaptation of the output, and rejection of the work of peers affect the duration of pollution events. We thereby expand previous work on quality management (Halfaker et al. 2011) and other research on the importance of continuous, even mundane, contributions in OC systems (Beck et al. 2015). Our dependent variable, pollution event resolution time—as opposed to various article quality metrics that have been used in the past—provides new insights on quality management in OC systems such as Wikipedia when a local quality “crisis” arises. These insights concern the distribution and number of edits during the pollution event, as well as the use of policy citations as a governance mechanism at the operational level.

Previous research on the connection between work distribution and quality in Wikipedia finds that unequal work distribution, where a limited number of participants do most of the work, is associated with higher quality of mature outputs (Kittur and Kraut 2008). Our findings add further insight by providing evidence on how work distribution across a limited number of participants impacts pollution resolution time. Consistent with previous work, we find that the limited number of participants is indeed positively associated with outcomes; however, we also find that the more equal the work distribution is across those participants, the faster the pollution resolution occurs. Future research should examine the optimal balance between number of participants and their work distribution in addressing quality problems as they arise.

Previous research is mixed regarding the association between the number of edits and the final quality score of Wikipedia articles. Some studies find no relationship (Arazy and Nov 2010; Kane 2010), while others detect a positive relationship (Javanmardi and Lopes 2010; Wilkinson and Huberman 2007). We find that, on the pollution event level, the number of edits is associated with

faster cleanup in articles with two or less pollution events, and that it becomes associated with slower pollution resolution in articles with three or more pollution events. This suggests that we need more research on how governance differences impact isolated and persistent pollution.

Prior research provides insights into policy citations as a governance mechanism in Wikipedia (Beschastnikh et al. 2008), but it does not specify how these citations relate to quality management. Our findings suggest that policy citation—a mechanism for regulating article boundaries—at first slows down the cleanup but, after reaching a threshold, is associated with faster cleanup. We speculate that boundary regulation is initially likely to indicate confusion or disagreement regarding an article’s trajectory. However, consistent boundary regulation helps better align contributors on the article, which can stabilize its quality faster.

Finally, prior research on the revert mechanism in Wikipedia finds it to be useful for combating vandalism (Priedhorsky et al. 2007) as a form of information pollution. However, researchers also find that the revert mechanism sometimes creates edit wars between providers with conflicting views on an article’s preferred trajectory (Yasseri et al. 2011). Other research finds that persistent rejection of contributions from new participants cause them to drop out, which is counterproductive for Wikipedia’s long-term sustainability (Halfaker et al. 2011,2013). We build on these insights and find that reverts of revisions between participating contributors (as opposed to individual reverts aimed at combatting attempts to presumably vandalize an article) are initially associated with faster pollution resolution, while persistent rejection of peer work slows it down. This suggests that the revert mechanism is effective initially for removing low-quality contributions but, when continuously used, it discourages participants from continuing to help clean up the article. This is consistent with past findings, yet the conceptualization of revision

reverts as indicating peer rejection gives us better theoretical insight into their negative association with quality management.

4.7.2 Contribution to PIC Theory

PIC outlines a multifaceted theoretical argument on how collective-action threats negatively impact the sustainability of online systems characterized by low barriers to participation. The theory argues that, to address collective-action threats and improve the odds of sustainability, these systems must create the conditions for free, unstructured participation in determining the system's trajectory by incorporating polycentric governance practices in their design of rules and technical features. In other words, PIC theory asserts that users should be given more freedoms and face less restrictions in setting the boundaries, monitoring against rule violators, and gradually impacting the adjustment of system rules and features. The argument is based on Ostrom's (1990) observations that polycentric governance is more effective than centralized control at sustaining shared resource systems in the physical world. However, adapting polycentric governance from the physical to the virtual realm is not straightforward and to date remains purely theoretical. By examining a specific part of the theory—the relationship between the four polycentric governance principles and the resolution of information pollution as one of five collective-action threats³¹—we confirm the relationship between polycentric governance and pollution resolution and provide some important new insights currently not addressed in PIC theory.

While we find shared accountability and incremental adaption to be directly associated with pollution resolution as predicted in PIC theory, we find the association between boundary regulation and pollution to be curvilinear—initially associated with a slowdown in resolution, and

³¹ The others are: free-riding, congestion, violations, and rebellion (Mindel et al. 2018).

only after reaching a threshold associated with more expedient pollution resolution. Incremental adaptation was found to speed-up resolution only in articles with a single pollution event; in articles with three or more pollution events, incremental adaptation was found to be negatively associated with resolution time. We speculate that repeated pollution is indicative of a difficulty in reaching a consensus over an article's form and content. In our random sample, only about 3% of articles have three or more pollution events. Nonetheless, we identify here an opportunity for future research to examine in more detail polycentric governance in the face of repeated pollution.

PIC theory did not predict the initial worsening in pollution resolution because of boundary regulation, yet it makes intuitive sense after observing this empirically. Citing rules during a collaborative effort to maintain its boundaries indicates that the participants are not aligned, but after the rules have been expressed enough times, the workflow becomes smoother and faster. This is what we believe happens during resolution of information pollution in Wikipedia—the initial increase in policy citations indicates a state of flux, yet persistent reminders about the article's boundary help to get all participants aligned.

Finally, PIC theory asserts that to keep the community members engaged, they should receive recognition for their efforts. Although Wikipedia has a couple of noteworthy provider recognition mechanisms, those mechanisms are expressed only on the collective-choice level in the background of article editing; no provider recognition mechanisms exist on the operational level, where the actual provision work takes place. For that reason, we could not directly examine how provider recognition positively relates to pollution resolution, but instead had to rely on the reverse logic that “provider rejection”—persistent rejection of peer contributions—is likely to negatively relate to pollution resolution. Although Wikipedia's revert mechanism was adapted to combat vandalism and has proven effective for that, it also has the unintended consequence of becoming

a tool for edit wars and is used by editors with diverging views on an article's trajectory to undermine each other's work. Though we have no way to know for sure, we speculate that the negative impact of persistent rejection on pollution resolution is the combination of two factors: multiple rejections of contributions slows down the progress of cleaning; it also discourages those editors whose contributions were rejected from meaningfully participating in the cleaning of pollution.

Our research also contributes to PIC theory by extending it to the operational level of polycentric governance. In Ostrom's (1990) seminal work, she identifies three interrelated levels of governance: (1) constitutional, (2) collective-choice, and (3) operational. As Figure 4.2 shows, the three levels of governance, though distinct, are very much interrelated through permeable links of rules, policies, and feedback loops. In our research, we focus on the operational level, where policies and rules determined on the levels of constitutional and collective-choice inform the operational, day-to-day activities of contributors in the face of pollution.

At present, PIC theory, while mentioning the three levels, does not directly address the distinction between levels. Better distinction between the levels of governance in the analysis of rules and technical features, however, provides more nuance as to how polycentric practices are incorporated in a system. In our research, we find that, while article governance hinges on the rules and technical features determined on the constitutional and collective-choice levels, conceptualizing polycentricity on the operational level is far from straight forward.

The principle of shared accountability refers to the general desirability of having the community of providers participate together in governing the information commons (Mindel et al. 2018). However, logic and past research tell us that it is impractical and counterproductive to have

everyone participate (Kittur and Kraut 2008). Polycentricity is not about equal participation, but rather is about the spontaneous emergence and disappearance of independent centers of influence. We find that, in relation to pollution resolution on the operational level, it is desirable to have a limited number of providers, yet the more equally the workload is spread among them, the better the outcome. Thus, shared accountability on the operational level is not about how many people show up to help, but more about how the few that self-select to show up share the workload.

Incremental adaptation is another abstract principle that appertains to the general desirability of having the information commons gradually adapt so as to avoid “rocking the boat” too much and also preserve the benefit of having multiple perspectives that build on previous foundations. As such, all of Wikipedia’s guidelines and templates are continuously generated in a gradual manner through continuous collective-choice processes. On the operational level, each single edit represents a pattern of incremental adaptation. In our research, we examined how such adaptation affects resolution of pollution events and find that there is a limit to how effective it can be. That is, we find that incremental adaptation on the operational level, though helpful in most cases (which have a single pollution event), is also associated with slower pollution resolution in articles with persistent pollution (three or more instances).

The rules and policies, as well as the revert mechanism, which were developed in Wikipedia through collective-choice processes over time to improve article quality and contribute to the overarching goal of producing encyclopedic information, do not directly alleviate pollution when employed on the operational level. Here, citation of policies determined on the collective-choice level for regulating article boundaries slowdown pollution resolution before alleviating it, suggesting a lag effect between the enforcement of boundaries and the redirection of an article’s trajectory.

Because some information commons do not have an explicit provider recognition mechanism on the operational level, researchers must be creative in examining other factors that impact provider motivation. Consistent with previous research (Halfaker et al. 2011,2013), we find that persistent provider rejection through reverts of revisions slows down pollution resolution. This finding is theoretically consistent with the notion introduced in PIC theory on the positive impact of provider recognition, yet it also differs in two key respects: (1) Rejection has an initial positive impact on pollution resolution time, which suggests that it—at least initially—acts as a boundary regulation mechanism. (2) We can only speculate as to whether persistent rejection reduces the effectiveness of pollution resolution because it slows down the cleaning process or because it causes providers trying to help to quit. Previous research finds that rejection causes demotivation, but we suspect that contributors might quit simply because they are frustrated with having to start over. There is an opportunity for future research to zoom in on this question.

4.7.3 Practical Implications

Our research has implications for the management of OC systems and other online platforms with low barriers to participation. The openness of various OC systems makes them a fertile ground for information pollution, ranging from low-quality and erroneous information to incidental and deliberate misinformation to uploads of disturbingly gruesome content (Chen 2014; Holman Rector 2008; Kanter 2018; Kupferberg and Protus 2011; Lavasa et al. 2011). Further, the increase in information pollution reduces public trust in the systems, and there have been calls to impose more traditional regulations on them (Barrett 2018). In short, information pollution can jeopardize the long-term viability and independence of OC systems characterized by low barriers to participation.

In response to such threats, systems are spending an increasing portion of their resources to employ professional moderators to identify and remove polluted content, but even so they often concede that it is difficult to keep up with the flood of content (Chen 2014; Kanter 2018). Moreover, many systems do not have the means to employ content moderators and are even more vulnerable to degeneration because of pollution.

Our findings suggest that online systems characterized by low barriers to participation should adapt mechanisms and rules that facilitate participation in the governance of those systems. This may often seem managerially counterintuitive, as systems under threat tend to close themselves and seek control rather than further opening for participation. Facebook's recent decisions to invest greater resources into controlling and monitoring content illustrate this instinct (Kanter 2018). By contrast, building on our insights and PIC theory, we assert that systems with low barriers to participation may need to allow more of their users greater freedom at setting boundaries and monitoring and adjusting the system. For instance, most open-access systems already incorporate various flagging mechanisms that let users identify bad content; we assert that the producers of these systems should consider going further and give some editorial privileges to trusted users. As our findings indicate, shared accountability is not about having everyone participating, but more about getting the right people to share the burden.

Instead of relying only on professional monitors, systems such as Facebook should find inspiration in the polycentric Wikipedia model and experiment with ways in which the community of users can discuss and agree on article boundaries; monitor, identify, and help remove polluted content; and suggest new mechanisms for maintaining the system's information integrity. There is no general recipe for how open-access systems online should do it, but if those systems find ways to tap into their user base—not just for participation in content provision and exchange, but also for

monitoring and removing pollution—they will likely resolve pollution events in a more efficient and cost-effective manner.

4.8 References

- Aaltonen, A., and Lanzara, G. F. 2015. “Building governance capability in online social production: insights from Wikipedia,” *Organization Studies* (36:12), pp. 1649–1673.
- Aberdour, M. 2007. “Achieving quality in open-source software,” *IEEE Software* (24:1), pp 58-64.
- Allcott, H., and Gentzkow, M. 2017. “Social media and fake news in the 2016 election,” *National Bureau of Economic Research*. (No. w23089).

- Allen, N. J., and Meyer, J. P. 1990. "Organizational socialization tactics: A longitudinal analysis of links to newcomers' commitment and role orientation," *Academy of Management Journal* (33:4), pp 847-858.
- Anderka, M., Stein, B., and Lipka, N. 2012, August. "Predicting quality flaws in user-generated content: the case of Wikipedia," In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*, pp 981-990. ACM.
- Anthony, D., Smith, S.W. and Williamson, T., 2005. "Explaining quality in internet collective goods: Zealots and good Samaritans in the case of Wikipedia," Hanover: *Dartmouth College*.
- Arazy, O., and Nov, O. 2010, February. "Determinants of Wikipedia Quality: The Roles of Global and Local Contribution Inequality," In *Proceedings of the 2010 ACM conference on Computer Supported Cooperative Work*, pp 233-236. ACM.
- Arazy, O., Nov, O., Patterson, R., and Yeo, L. 2011. "Information Quality in Wikipedia: The Effects of Group Composition and Task Conflict," *Journal of Management Information Systems* (27:4), pp 71-98.
- Baldwin, C., and von Hippel, E. 2011. "Modeling a Paradigm Shift: From Producer Innovation to User and Open Collaborative Innovation," *Organization Science* (22:6), pp 1399-1417.
- Barrett, B. 2018. "How Would Regulating Facebook Look Like?" *Wired*.
- Beck, R., Rai, A., Fischbach, K. and Keil, M., 2015. "Untangling knowledge creation and knowledge integration in enterprise wikis," *Journal of Business Economics* (85:4), pp 389-420.
- Benkler, Y. 2006. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*, Yale University Press.
- Benkler, Y., and Nissenbaum, H. 2006. "Commons-based peer production and virtue*," *Journal of Political Philosophy* (14:4), pp 394-419.
- Beschastnikh, I., Kriplean, T. and McDonald, D.W., 2008, March. "Wikipedian Self-Governance in Action: Motivating the Policy Lens," In *ICWSM*.
- Blumenstock, J. E. (2008). *Automatically Assessing the Quality of Wikipedia Articles*. UC Berkeley: School of Information.
- Brokowski, L., and Sheehan, A. H. 2009. "Evaluation of pharmacist use and perception of Wikipedia as a drug information resource," *Annals of Pharmacotherapy* (43:11), pp 1912-1913.
- Brown, A.R., 2011. "Wikipedia as a data source for political scientists: Accuracy and completeness of coverage," *PS: Political Science & Politics* (44:2), pp 339-343.
- Burke, M., and Kraut, R. 2008, November. "Mopping up: modeling Wikipedia promotion decisions," In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, pp 27-36.
- Butler, B. S., Bateman, P. J., Gray, P. H., and Diamant, E. I. 2014. "An attraction-selection-attrition theory of online community size and resilience," *Mis Quarterly* (38:3), pp 699-728.
- Butler, B., Joyce, E., and Pike, J. 2008, April. "Don't Look Now, but We've Created a Bureaucracy: The Nature and Roles of Policies and Rules in Wikipedia," In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp 1101-1110.
- Butler, B., Joyce, E., and Pike, J. 2008, April. "Don't look now, but we've created a bureaucracy: the nature and roles of policies and rules in Wikipedia," In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp 1101-1110. ACM.

- Butler, B.S., 2001. "Membership size, communication activity, and sustainability: A resource-based model of online social structures," *Information systems Research* (12:4), pp346-362.
- Cabunducan, G., Castillo, R., and Lee, J. B. 2011, July. "Voting behavior analysis in the election of Wikipedia admins," In *IEEE, International Conference on Advances in Social Networks Analysis and Mining*, pp 545-547.
- Chen, A. 2014. "The Laborers Who Keep Dick Picks and Beheadings out of your Facebook Feed," *Wired*.
- Chesney, T. 2006. "An empirical examination of Wikipedia's credibility," *First Monday* (11:11).
- Clauson, K. A., Polen, H. H., Boulos, M. N. K., and Dzenowagis, J. H. 2008. "Scope, completeness, and accuracy of drug information in Wikipedia," *Annals of Pharmacotherapy* (42:12), pp 1814-1821.
- Cunningham, W. and Leuf, B., 2001. "The Wiki way," In *Quick Collaboration on the Web*. Addison-Wesley Boston, MA.
- Dalip, D. H., Gonçalves, M. A., Cristo, M., and Calado, P. 2011. "Automatic assessment of document quality in web collaborative digital libraries," *Journal of Data and Information Quality (JDIQ)*, (2:3), 14.
- De la Calzada, G., and Dekhtyar, A. 2010, April. "On measuring the quality of Wikipedia articles." In *Proceedings of the 4th workshop on Information credibility*, pp. 11-18. ACM.
- Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., and Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences* (113:3), pp 554-559.
- Fallis, D. 2008. "Toward an epistemology of Wikipedia," *Journal of the Association for Information Science and Technology* (59:10), pp 1662-1674.
- Faraj, S., Jarvenpaa, S. L., and Majchrzak, A. 2011. "Knowledge Collaboration in Online Communities," *Organization Science*, (22:5) pp. 1224-1239.
- Forte, A., and Lampe, C. 2013. "Defining, understanding, and supporting open collaboration: Lessons from the literature," *American Behavioral Scientist* (57:5), pp 535-547.
- Forte, A., Larco, V. and Bruckman, A., 2009. "Decentralization in Wikipedia Governance," *Journal of Management Information Systems*, (26:1), pp.49-72.
- Giles, J. 2005. "Internet Encyclopedias go head to head," *Nature* (438:7070), pp 900-902.
- Greenstein, S., and Zhu, F. 2012. "Is Wikipedia Biased?" *American Economic Review* (102:3), pp 343-48.
- Gutwin, C., Penner, R., and Schneider, K. 2004, November. "Group awareness in distributed software development," In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, pp 72-81. ACM.
- Gyimothy, Tibor, Rudolf Ferenc, and Istvan Siket 2005. "Empirical validation of object-oriented metrics on open source software for fault prediction." *IEEE Transactions on Software Engineering* (31:10), pp 897-910.
- Haigh, C.A., 2011. "Wikipedia as an evidence source for nursing and healthcare students," *Nurse Education Today* (31:2), pp 135-139.
- Halfaker, A., Geiger, R.S., Morgan, J.T. and Riedl, J., 2013. "The rise and decline of an open collaboration system: How Wikipedia's reaction to popularity is causing its decline," *American Behavioral Scientist* (57:5), pp.664-688.

- Halfaker, A., Kittur, A., and Riedl, J. 2011, October. "Don't bite the newbies: how reverts affect the quantity and quality of Wikipedia work," In *Proceedings of the 7th international symposium on wikis and open collaboration*, pp 163-172. ACM.
- Hardin, G. 1968. "The tragedy of the commons," *Science* (162:3859), pp 1243-1248.
- Hertel, G., Niedner, S., and Herrmann, S. 2003. "Motivation of software developers in Open Source projects: an Internet-based survey of contributors to the Linux kernel," *Research Policy* (32:7), pp. 1159-1177.
- Hippel, E. V., and Krogh, G. V. 2003. "Open Source Software and the "Private-Collective" Innovation Model: Issues for Organization Science," *Organization Science* (14:2), pp 209-223.
- Holman Rector, L. 2008. "Comparison of Wikipedia and other Encyclopedias for Accuracy, Breadth, and Depth in Historical Articles," *Reference Services Review* (36:1), pp 7-22.
- Howison, J., and Crowston, K. 2004, May. "The perils and pitfalls of mining SourceForge," In *Proceedings of the International Workshop on Mining Software Repositories*, pp. 7-11.
- Hu, M., Lim, E. P., Sun, A., Lauw, H. W., and Vuong, B. Q. 2007, November. "Measuring article quality in Wikipedia: models and evaluation," In *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*, pp 243-252. ACM.
- Javanmardi, S., and Lopes, C. 2010, July. "Statistical measure of quality in Wikipedia," In *Proceedings of the First Workshop on Social Media Analytics*, pp 132-138. ACM.
- Kane, G. C. 2011. "A multimethod study of information quality in wiki collaboration," *ACM Transactions on Management Information Systems*, (2:1).
- Kane, G. C., and Ransbotham, S. 2016. "Content as Community Regulator: The Recursive Relationship Between Consumption and Contribution in Open Collaboration Communities," *Organization Science*, (27:5) pp. 1258-1274.
- Kanter, J. 2018. "A former Facebook moderator says she took down beheadings, child pornography, and animal abuse every day — but was 'treated like nothing'," *Business Insider*.
- Kata, A. 2012. "Anti-vaccine activists, Web 2.0, and the postmodern paradigm—An overview of tactics and tropes used online by the anti-vaccination movement," *Vaccine* (30:25), pp 3778-3789.
- Kittur, A. and Kraut, R.E., 2008, November. "Harnessing the wisdom of crowds in wikipedia: quality through coordination," In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, pp. 37-46. ACM.
- Kittur, A. and Kraut, R.E., 2010, February. "Beyond Wikipedia: coordination and conflict in online production groups," In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, pp. 215-224. ACM.
- Kriplean, T., Beschastnikh, I., and McDonald, D. W. 2008, November. "Articulations of wikiwork: uncovering valued work in Wikipedia through barnstars," In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, pp 47-56.
- Kupferberg, N., and Protus, B. M. 2011. "Accuracy and completeness of drug information in Wikipedia: an assessment," *Journal of the Medical Library Association* (99:4), pp 310-313.
- Kuznetsov, S. 2006. "Motivations of contributors to Wikipedia," *ACM SIGCAS computers and society* (36:2), 1.

- Lavsa, S.M., Corman, S.L., Culley, C.M. and Pummer, T.L., 2011. "Reliability of Wikipedia as a medication information source for pharmacy students," *Currents in Pharmacy Teaching and Learning* (3:2), pp 154-158.
- Levine, S. S., and Prietula, M. J. 2014. "Open Collaboration for Innovation: Principles and Performance," *Organization Science* (25:5), pp 1414-1433.
- Lim, S. and Simon, C., 2011. "Credibility judgment and verification behavior of college students concerning Wikipedia," *First Monday* (16:4).
- Majchrzak, A. 2009. "Comment: Where is the Theory in Wikis?" *MIS Quarterly* (33:1), pp 18-20.
- Mayzlin, D, Yaniv D, and Judith C, 2014. "Promotional reviews: An empirical investigation of online review manipulation." *American Economic Review* (104:8), pp 2421-55.
- Mettler, K. 2017. "Rip currents swept away a Florida family. Then beachgoers formed a human chain," *The Washington Post* online.
- Miller, J. C., and Murray, H. B. 2010. "Wikipedia in court: When and how citing Wikipedia and other consensus websites is appropriate," *St. John's Law Review* (84), pp 633-656.
- Mindel, V., Mathiassen, L. and Rai, A., 2018. "The Sustainability of Polycentric Information Commons," *Management Information Systems Quarterly* (42:2), pp 607-631.
- Mishra, B., Prasad, A., and Raghunathan, S. 2002. "Quality and profits under open source versus closed source," *ICIS 2002 Proceedings*, 32.
- Morgan, J.T. and Zachry, M., 2010, November. "Negotiating with angry mastodons: the wikipedia policy environment as genre ecology," In *Proceedings of the 16th ACM international conference on Supporting group work*, pp 165-168.
- Morgan, J.T., Gilbert, M., McDonald, D.W. and Zachry, M., 2013. "August. Project talk: Coordination work and group membership in WikiProjects," In *Proceedings of the 9th International Symposium on Open Collaboration* (p. 3). ACM.
- Nazaryan, A. 2016. Kitty Genovese's 50-year-old Murder Story Still haunts New York City," *Newsweek* online.
- Oreg, S., and Nov, O. 2008. "Exploring motivations for contributing to open source initiatives: The roles of contribution context and personal values" *Computers in human behavior*, (24:5), pp 2055-2073.
- Orman, L. 1984. "Fighting information pollution with decision support systems," *Journal of management information systems* (1:2), pp 64-71.
- Ostrom, E. 1990. *Governing the commons*, Cambridge university press.
- Ostrom, E. 2009. "A general framework for analyzing sustainability of social-ecological systems," *Science* (325:5939), pp 419-422.
- Ostrom, E., Gardner, R., Walker, J. and Walker, J., 1994. *Rules, games, and common-pool resources*. University of Michigan Press.
- Pancier, K., Halfaker, A., Terveen, L. 2009, May. "Wikipedians are born, not made: a study of power editors on Wikipedia," In *Proceedings of the ACM 2009 international conference on Supporting group work*, pp 51-60. ACM.
- Peoples, L.F., 2009. "The Citation of Wikipedia in Judicial Opinions," *Yale Journal of Law & Technology* (12), pp 1-51.
- Priedhorsky, R., Chen, J., Lam, S. T. K., Pancier, K., Terveen, L., and Riedl, J. 2007, November. "Creating, destroying, and restoring value in Wikipedia," In *Proceedings of the 2007 international ACM conference on Supporting group work*, pp 259-268.

- Ransbotham, S. 2010, June. "An Empirical Analysis of Exploitation Attempts Based on Vulnerabilities in Open Source Software," *In WEIS*.
- Ransbotham, S. and Kane, K.C., 2011. "Membership Turnover and Collaboration Success in Online Communities: Explaining Rises." *Management Information Systems Quarterly* (35:3), pp 613-627.
- Ren, Y., Harper, F. M., Drenner, S., Terveen, L. G., Kiesler, S. B., Riedl, J., and Kraut, R. E. 2012. "Building Member Attachment in Online Communities: Applying Theories of Group Identity and Interpersonal Bonds," *MIS Quarterly* (36:3), pp 841-864.
- Scacchi, W. 2002. "Understanding the Requirements for Developing Open Source Software Systems," *IEEE Proceedings-Software* (149:1), pp 24-39
- Shah, S.K., 2006. "Motivation, governance, and the viability of hybrid forms in open source software development," *Management Science* (52:7), pp 1000-1014.
- Stamelos, I., Angelis, L., Oikonomou, A., and Bleris, G. L. 2002. "Code quality analysis in open source software development," *Information Systems Journal* (12:1), pp 43-60.
- Stvilia, B., Twidale, M.B., Smith, L.C. and Gasser, L., 2008. "Information quality work organization in Wikipedia," *Journal of the American society for information science and technology* (59:6), pp 983-1001.
- Suh, B., Convertino, G., Chi, E.H. and Pirolli, P., 2009, October. "The singularity is not near: slowing growth of Wikipedia," *In Proceedings of the 5th International Symposium on Wikis and Open Collaboration*, Article Number 8, ACM.
- Sumi, R. and Yasserli, T., 2011, October. "Edit wars in Wikipedia," *In Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom)*, pp 724-727.
- Tapscott, D., and Williams, A. D. 2008. *Wikinomics: How mass collaboration changes everything*. Penguin.
- Viegas, F. B., Wattenberg, M., Kriss, J., and Van Ham, F. 2007, January. "Talk Before You Type: Coordination in Wikipedia," *IEEE System Sciences, HICSS 2007. 40th Annual Hawaii International Conference on*, pp. 78-78.
- von Krogh, G., Haefliger, S., Spaeth, S., and Wallin, M. W. 2012. "Carrots and Rainbows: Motivation and Social Practice in Open Source Software Development," *MIS Quarterly* (36:2), pp 649-676.
- Wilkinson, D. M., and Huberman, B. A. 2007, October. "Cooperation and quality in Wikipedia," *In Proceedings of the 2007 international symposium on Wikis*, pp 157-164.
- Wöhner, T. and Peters, R., 2009, October. "Assessing the quality of Wikipedia articles with lifecycle based metrics," *In Proceedings of the 5th International Symposium on Wikis and Open Collaboration* (16). ACM.
- Yang, H. L., and Lai, C. Y. 2010. "Motivations of Wikipedia content contributors," *Computers in Human Behavior* (26:6), pp 1377-1383.
- Yasserli, T., Sumi, R., Rung, A., Kornai, A., and Kertész, J. 2012. "Dynamics of conflicts in Wikipedia," *PloS One* (7:6).
- Zentneo, R. 2016. "Good Samaritans rescue 9-year-old from flipped SUV," *CNN* online.

Appendix A: Dataset Construction

Wikimedia Foundation makes data underpinning different Wikipedia sites freely available at <https://dumps.wikimedia.org>. We downloaded a data dump for the English Wikipedia on 1 January 2017 and transformed it into a tabulated dataset following a similar approach applied in Aaltonen and Lanzara (2015) and Aaltonen and Seiler (2015).

According to ‘enwiki-20170101-site_stats.sql’ file in the data dump, the raw data contains the full text of 867,033,963 article revisions to 41,099,459 articles since January 2001. 5,321,706 of these articles represent encyclopedia articles, while the rest of the articles are lists, redirects, portal pages and other means of organizing the content, user profiles, help pages and pages that support administrative and editorial functions such as Wikipedia policies and guidelines, categories, templates, and more. Each article may be accompanied by a talk page that allow users to discuss its content.

The downloaded data contains XML-formatted page records that encapsulate revision records that is, individual edits to the page. The page record contains a header section with the page title, namespace, id and some optional tags such as “redirect” in the example below. The header section is followed by revision records. Each revision record contains an id, timestamp, contributor record, the full text of the revision and some optional tags such as “comment” that is a summary of the edit (not to be confused with talk page comments). Below is an example of page record. Note that we have truncated the record to show only the first revision of the page and slightly altered indentation to make it easy to perceive the XML structure.

```
<page>
  <title>AccessibleComputing</title>
  <ns>0</ns>
  <id>10</id>
  <redirect title="Computer accessibility" />
```

```

<revision>
<id>233192</id>
<timestamp>2001-01-21T02:12:21Z</timestamp>
<contributor>
<username>RoseParks</username>
<id>99</id>
</contributor>
<comment>*</comment>
<model>wikitext</model>
<format>text/x-wiki</format>
<text xml:space="preserve">This subject covers

* AssistiveTechnology

* AccessibleSoftware

* AccessibleWeb

* LegalIssuesInAccessibleComputing

</text>
<sha1>8kul9t1wjm9oxgvqzbwuegt9b2830vw</sha1>
</revision>
.
.
.
<page>

```

An exact graphical depiction of the pipeline that was used to transform the can be found in Figure A1, together with a brief explanation of each step in Table A1. The following is a narrative overview of the process.

We take a random sample of 340,000 articles and associated 172,576 talk pages (not all articles have a talk page) and transform their XML records into an initial tabulated dataset using a purpose-built Python script (Step 5, 6, 7, 8). Due to a technical limitation of our processing environment we have dropped 5 pages from further processing due to their large size (over ~2 GB). The transformation process stores page title and id, revision id, timestamp, user id and name, and lots of other information for each sampled page revision. Each revision is represented by a row in the initial dataset (output from Step 8). During the initial transformation, we also calculate a number of further metrics such as the Levenshtein edit distance between consecutive page revisions to

estimate the amount of content change or effort put into each revision. We also analyze the content of each revision by matching a number of regular expression patterns against the revision content, creating a variable for each matched pattern.

We then identified 4,924 articles with the presence of a cleanup template at some point in their history (Step 9, 10). We further removed pages with intractable revert patterns during pollution events to arrive at a sample of 4,679 articles and 5,487 pollution events as some articles contain more than one event, that is, a cleanup template appears on them more than once.

Table A1 Data Pipeline Steps

Step	Label	Description
1	enwiki-20170101	Data dump download from https://dumps.wikimedia.org
2	20170216-enwiki-20170101-filenames	Creation of a list of filenames in the data dump
3	20170303-one-id	Creation of a dummy file required by the processing script
4	20170303-enwiki-20170101-page-ids	Creation of a list of page ids
5	20170703-random-sample-ids	Random sampling of article pages and their talk pages (ids)
6	20180422-enwiki-user-groups	Identification of users that belong to a particular user group
7	20180422-enwiki-regexp-analyses	Creations of regular expression patterns to analyze page and edit summary content
8	20180501-article-sample-with-talk-pages	Transformation of XML records for the sampled pages into a tabulated dataset
9	20180505-revision-dataset-per-polluted page	Creation of a separate datasets that combines both page and talk page revisions for each page with the presence of Cleanup or Cleanup rewrite template.
10	20180507-pollution-events-dataset	Creation of an aggregated dataset with one row per observed pollution event.

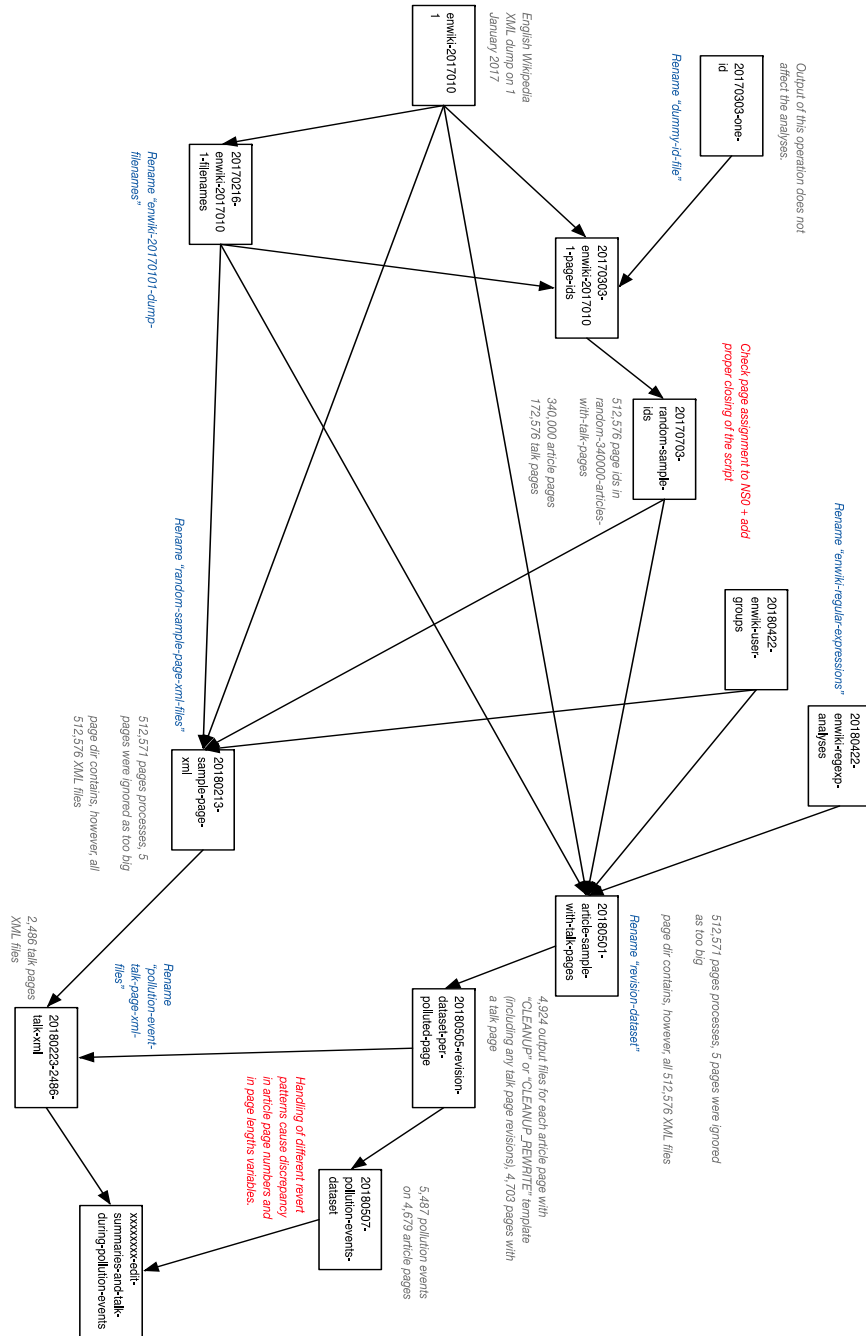


Figure A1. Data Pipeline Graph

Appendix B: Robustness Analysis

Table B1. Articles with Single Pollution Event Results, DV: Event Length Time (N=4,020)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Article age	.144***	.106***	.097***	.076***	.057***	.059***	.060***
Length of article	.015	-.033	-.068**	-.095***	-.061**	-.067**	-.069**
Providers before pollution event	-.116**	-.167***	-.317***	-.393***	-.346***	-.319***	-.307***
Providers before: talk page	.066	.103**	.142***	.225***	.201***	.174***	.173***
References before pollution	-.025	-.014	-.023	-.020	-.043*	-.058**	-.058**
Work distribution / providers (SHA)		-.458***	-.357***	-.273***	-.289***	-.261***	-.259***
Policies cited in revisions (BR)			.392***	.796***	.964***	1.051***	1.080***
Policies cited in revisions ² (BR ²)				-.466***	-.425***	-.350***	-.429***
Revisions (INA)					-.256***	-.235***	-.207***
Reverted Revisions (U_PRJ)						-.395***	-.502***
Reverted Revisions ² (U_PRJ ²)							.134**
Model F	10.1***	129.3***	200.3***	241***	229.3***	236.9***	216.9***
F change		712.2***	487.8***	347.6***	79.7***	174.1***	9.7**
Adj. R ²	0.016	0.22	0.338	0.413	0.429	0.463	0.465
ΔR ²		0.203	0.118	0.075	0.017	0.034	0.002

* $p < .05$; ** $p < .001$; *** $p < .0001$

Table B2. Articles with Two Pollution Events Results, DV: Event Length Time (N=1038)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Article age	.314***	.246***	.210***	.174***	.153***	.149***	.132***
Length of article	-.028	-.082	-.090*	-.102*	-.065	-.097*	-.111*
Providers before pollution event	-.423***	-.467***	-.639***	-.619***	-.615***	-.540***	-.458***
Providers before: talk page	.232**	.261***	.374***	.373***	.375***	.342***	.301***
References before pollution	-.017	.035	.049	.042	.020	.015	.012
Work distribution / providers (SHA)		-.370***	-.270***	-.205***	-.220***	-.208***	-.208***
Policies cited in revisions (BR)			.376***	.796***	.914***	1.003***	1.235***
Policies cited in revisions ² (BR ²)				-.441***	-.440***	-.352***	-.618***
Revisions (INA)					-.155**	.048	.098
Reverted Revisions (U_PRJ)						-.395***	-.830***
Reverted Revisions ² (U_PRJ ²)							.447***
Model F	16.6***	42.7***	66.3***	68.2***	62.6***	63.7***	63.6***
F change		159.4***	165.6***	55.6***	12.1**	47.3.1**	38.6**
Adj. R ²	0.077	0.201	0.315	0.350	0.358	0.386	0.409
ΔR ²		0.128	0.114	0.036	0.008	0.029	0.023

* $p < .05$; ** $p < .001$; *** $p < .0001$

Table B3. Articles with Three or More Pollution Events Results, DV: Event Length Time (N=408)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Article age	.352***	.316***	.281***	.255***	.240***	.216***	.187***
Length of article	.015	-.044	-.067	-.048	-.030	-.083	-.091
Providers before pollution event	-.329**	-.321**	-.336**	-.392***	-.403***	-.305**	-.266**
Providers before: talk page	.020	.020	.002	.029	.035	.027	.013
References before pollution	.056	.078	.064	.045	.040	.028	.034
Work distribution / providers (SHA)		-.314***	-.252***	-.187***	-.205***	-.185***	-.183***
Policies cited in revisions (BR)			.327***	.905***	1.124***	1.274***	1.886***
Policies cited in revisions ² (BR ²)				-.605***	.240***	-.541***	-1.22***
Revisions (INA)					-.030	.245	.365**
Reverted Revisions (U_PRJ)						-.712***	-1.51***
Reverted Revisions ² (U_PRJ ²)							.771***
Model F	8.1***	14.8***	21.4***	24.1***	22***	24.6***	26.3***
F change		43.8***	49.4***	31.3***	3.8	31.4***	26.4**
Adj. R ²	0.085	0.178	0.272	0.326	0.332	0.382	0.421
ΔR ²		0.094	0.094	0.055	0.007	0.051	0.040

* p < .05; ** p < .001; *** p < .0001

Table B4. First Pollution Event Only, DV: Event Length Time (N=4,679)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Article age	.170***	.094***	.094***	.095***	.099***	.101***	.095***
Length of article	-.018	-.073***	-.073***	-.072***	-.124***	-.135***	-.130***
Providers before pollution event	-.171***	-.162***	-.162***	-.176***	-.209***	-.184***	-.189***
Providers before: talk page	.055	.050	.050	.049	.047	.036	.026
References before pollution	.017	.040*	.040*	.037*	.064*	.061***	.059***
Work distribution / providers (SHA)		-.428***	-.428***	-.424***	-.372***	-.364***	-.360***
Policies cited in revisions (BR)			.094	-.005	-.009	-.009	-.006
Policies cited in revisions ² (BR ²)				.094***	.077**	.057**	.226***
Revisions (INA)					-.223***	-.366***	-.301***
Reverted Revisions (U_PRJ)						-.215***	-.148***
Reverted Revisions ² (U_PRJ ²)							.273***
Model F	21.4***	184.8***	158.4***	146.2***	152.5***	144***	135***
F change		978.5***	0.023	48.6***	162***	52***	34.4**
Adj. R ²	0.023	0.196	0.196	0.204	0.232	0.24	0.246
ΔR ²		0.174	0.000	0.009	0.028	0.009	0.006

* p < .05; ** p < .001; *** p < .0001

Table B5. Seconds Pollution Event Only, DV: Event Length Time (N=639)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Article age	.226***	.198***	.182***	.150***	.121**	.100*	.088*
Length of article	-.081	-.135*	-.130*	-.141**	-.100	-.123*	-.128*
Providers before pollution event	-.381***	-.388***	-.527***	-.526***	-.511***	-.412***	-.399***
Providers before: talk page	.192*	.182*	.258***	.271***	.268***	.236**	.227**
References before pollution	.007	.070	.091	.083	.057	.030	.041
Work distribution / providers (SHA)		-.332***	-.247***	-.185***	-.206***	-.181***	-.174***
Policies cited in revisions (BR)			.296***	.706***	.902***	1.060***	1.320***
Policies cited in revisions ² (BR ²)				-.432***	-.396***	-.363***	-.695***
Revisions (INA)					-.275***	-.237***	-.221***
Reverted Revisions (U_PRJ)						-.527***	-1.13***
Reverted Revisions ² (U_PRJ ²)							.536***
Model F	9.2***	21.8***	28.8***	30***	29***	31***	32.3***
F change		78.7***	58.8***	28.6***	15.3***	34.8***	29.9***
Adj. R ²	0.063	0.169	0.242	0.275	0.292	0.329	0.360
ΔR ²		0.107	0.073	0.034	0.018	0.038	0.031

* $p < .05$; ** $p < .001$; *** $p < .0001$

Table B6. Third or More Pollution Events, DV: Event Length Time (N=169)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Article age	.283**	.251**	.226**	.198**	.191*	.180*	.184*
Length of article	-.066	-.064	-.008	.058	.058	.078	.074
Providers before pollution event	-.320**	-.346**	-.377***	-.414***	-.413***	-.427***	-.423***
Providers before: talk page	.261**	.188*	-.025	-.143	-.127	-.129	-.141
References before pollution	.040	.025	.014	-.038	-.036	-.034	-.037
Work distribution / providers (SHA)		-.298***	-.289***	-.229***	-.234***	-.224***	-.222**
Policies cited in revisions (BR)			.301**	1.147***	1.217***	1.486***	1.385***
Policies cited in revisions ² (BR ²)				-.803***	-.804***	-.875***	-.773**
Revisions (INA)					-.087	-.062	.024
Reverted Revisions (U_PRJ)						-.267*	-.255*
Reverted Revisions ² (U_PRJ ²)							-.094
Model F	5.6***	7.7***	8.4***	10***	8.9***	8.7***	7.8***
F change		15.5***	10.2***	15.5***	.19	4.8***	.2*
Adj. R ²	0.125	0.201	0.246	0.312	0.318	0.335	0.352
ΔR ²		0.078	0.049	0.067	0.006	0.018	0.017

* $p < .05$; ** $p < .001$; *** $p < .0001$

Variable	Model 1	Model 2	Model 3
Article age	.314***	.223***	.206***
Length of article	.065**	-.083***	-.124***
Providers before pollution event	-.521***	-.534***	-.447***
Providers before: talk page	.126***	.024*	-.044***
References before pollution	-.025**	.053***	.074***
Work distribution / providers (<i>SHA</i>)		-.162***	-.167***
Policies cited in revisions (<i>BR</i>)		1.824***	1.897***
Policies cited in revisions ² (<i>BR</i> ²)		-1.127***	-1.118***
Revisions (<i>INA</i>)		-.307***	-.128***
Reverted Revisions (<i>PRJ</i>)		-1.420***	-1.338***
Reverted Revisions ² (<i>PRJ</i> ²)		.738***	.719***
Number of editors on talk page			.189***
Policies cited in talk pages			.047***
Policies cited in talk pages ²			-.002***
Talk page revisions			.005***
Reverted revision talk page			-.267***
Reverted revision talk page ²			.129
Model <i>F</i>	773.6	1749.2	1213.6
<i>F</i> change		2258.2	139.1
Adj. <i>R</i> ²	.119	.401	.418
ΔR^2		.283	.017

* $p < .05$; ** $p < .001$; *** $p < .0001$

Variable	Model 1	Model 2	Model 3
Article age	.152***	.022	.061***
Length of article	-.028	-.103***	-.047*
Providers before pollution event	-.230***	-.236***	-.191***
Providers before: talk page	.087*	.115***	.114***
References before pollution	.019	-.010	-.006***
Work distribution / providers (<i>SHA</i>)		-.226***	-.201***
Policies cited in revisions (<i>BR</i>)		1.335***	1.272***
Policies cited in revisions ² (<i>BR</i> ²)		-.699***	-.661***
Revisions (<i>INA</i>)		.041	-.121***
Reverted Revisions (<i>PRJ</i>)		-.890***	-.839***
Reverted Revisions ² (<i>PRJ</i> ²)		.429***	.383***
Ratio of remaining veterans to previously engaged editors			.042**
Number of veterans			-.106***
Ratio of veteran during pollution event			-.132***
Proportion of length change by veterans			.014
Veteran's edit distance during pollution event			-.070***
Proportion of Veteran's edit distance during pollution event			-.006
Model <i>F</i>	21	254.4	182.9
<i>F</i> change		437.5	31
Adj. <i>R</i> ²	.024	.408	.434
ΔR^2		.385	.026

* $p < .05$; ** $p < .001$; *** $p < .0001$