Understanding Collective Reflection in Crowdsourcing for Innovation: A Semantic Network Approach

Yao Sun
College of Arts and Sciences
University of South Florida
yaosun@usf.edu

Ann Majchrzak
Marshall School of Business
University of Southern California
majchrza@usc.edu

Arvind Malhotra
Kenan-Flagler Business School
University of North Carolina at Chapel Hill
arvind_malhotra@kenan-flagler.unc.edu

Abstract

Empowered by the wisdom of crowds, innovation nowadays is increasingly relying on diverse individuals' knowledge collaboration. Research on crowdsourcing and open innovation has demonstrated that through deliberate understanding and reflective thinking, members of the online crowd collectively manage their knowledge to generate innovative ideas. However, the semantic patterns of how online crowd's collective reflection ultimately leads up to innovation remains unclear. Employing semantic network approach, this study analyzed a total of 1,116 posts contributed by online crowds responding to two organization-sponsored crowdsourcing open innovation challenges. Findings show that the semantic patterns of online crowds' knowledge collaboration evolve from one phase to another in accordance with crowd members' collective reflection on their diverse knowledge. Theoretical and practical implications are discussed.

1. Introduction

Crowdsourcing is a frequently used innovation strategy described in recent research [1-2]. Crowd-based open innovation refers to "the use of purposive inflows and outflows of knowledge to accelerate internal innovation, and to expand the markets for external use of innovation" [65, p.1], highlighting a sequential and recombinant innovation [66]. The information systems and innovation literature has viewed crowd-powered open innovation as an approach to sourcing distant knowledge and thus generating novel solutions [67-70].

As communication technology lends itself to large-scale interaction, crowd-based online knowledge exchange and sharing has been boosted [3]. On social media, in particular, knowledge collaboration is configured by dynamic sharing activities emerged from ongoing social interactions among contributors [4]. The proliferation of communication technology has facilitated the emergence and advancement of semantic knowledge network. Recognition of the value of word-to-word relationships can be traced back to the early research of Collins and Quillian [5], which suggested that cross-word connections can be viewed as a manifestation of collaborators' shared thoughts and common minds. According to their work, associated terms and words are stored hierarchically in individuals' minds, and therefore the meanings are constructed by the terms and words that refer to one another. Knowledge is stored in human memory in a similar manner. When individuals attempt to describe a thing, relevant hierarchically stored words will be activated, so that individuals can create sentences to build shared meanings [5-7].

Semantic network is similar to social network in that it exhibits the structure of a networked relationship; however, it differs from social network because it is based on communicators' textual connections rather than social linkages. In such a network, connections are formed by overlapping concepts instead of interactional instances [8,9]. The goal of semantic network analysis, therefore, is to allow the meanings to emerge and thus to be identified. Semantic network analysis does not employ a pre-defined scheme that is often seen in traditional content analysis; it utilizes natural language processing rather than human coding to decode the large-scale shared meanings, in order to inherently ensure reliability and validity [10]. As such, semantic network can be beneficial for studying knowledge contribution because this approach, to a great extent, allows valuable insights to surface during collaborative interaction.

Studying online collective action through the lens of semantic network, the crowdsourcing process configured by interconnected words can be unfolded. In the first place, texts are scanned in order to find the

URI: https://hdl.handle.net/10125/71245
978-0-9981331-4-0
(CC BY-NC-ND 4.0)

most frequently used words; then an adjacency matrix is built based on calculating the frequencies of meaningful words, such that word co-occurrence pairs can be identified. The matrix can be further analyzed so as to detect clusters or other types of inter-words correlations, for a better understanding of the interlinked meanings [11]. Semantic network approach allows researchers to capture the visible and quantifiable attributes based on which shared meanings are built and facilitates a deep understanding of individuals' common beliefs and values.

The present study examines online crowds' knowledge collaboration and collective reflection by scrutinizing various semantic attributes of the online crowd-generated content as well as exploring the development and evolution of the semantic networks emerged from online crowds' collaborative knowledge sharing.

2. Conceptual Development

2.1. Communicating Knowledge in Virtual Space

Social interaction is a natural human tendency [12]. In general, individuals engage in regular and frequent social interactions to fulfill the need to belong. As Baumeister and Leary [12] explained, individuals have "a need to form and maintain at least a minimum quantity of interpersonal relationships, [which] is innately present (and hence nearly universal) among human beings" (p. 499). Taking part in social interactions helps to build sustained connections, as well as establish and maintain a feeling of belonging.

Individuals communicate knowledge in virtual space for a variety of purposes. As a collective construction, knowledge sharing provides collaborators with an opportunity to build and maintain relationships as well as reach agreements. Research indicates that IT knowledge and experience positively affect the promotion of managers in different areas [71, 72]. As suggested by Nelson and Cooprider [13], communicating knowledge can help individuals satisfy the needs for interaction, build social connections, as well as develop mutual trust and shared visions. As the affordances of online knowledge collaboration community enable knowledge contributors to satisfy their self-presentation needs while co-creating knowledge artifacts [14], contributors often tend to play different roles in knowledge creation [15,16]. Moreover, visual anonymity and the freedom afforded by online communication offer alternative ways of managing their self-presentations in front of others [17], and members of online crowds enjoy the autonomy in customizing their own presentations and experimenting with alternative self-image constructions. As Wallace [18] noted, the Internet serves "an identity laboratory, overflowing with props, audiences, and players for our personal experiments" (p.48).

Communicating knowledge online is jointly motivated by an individual's cognition and the social context in which he or she is embedded. As computer-mediated communication enables the extension of existing offline social networks as well as the construction of new online social networks, individuals more and more rely on virtual communities to exchange their knowledge. According to social cognitive theory, activities in virtual space may be viewed as inherently "triadic, dynamic, and reciprocal interaction of personal factors, behavior, and the social network", as "virtual communities are online social networks in which people with common interests, goals, or practices interact to share information and knowledge, and engage in social interactions" [19]. In other words, virtual communities are essentially sustained by interrelated connections and networked interactions. According to Bandura [20], an individual's behavior not only depends on his or her own values or beliefs but is also influenced by the social structure in which the individual is embedded. When participating in online knowledge collaboration, members of the crowd are involved in virtual communities where their interactions are interrelated and networked, and thus knowledge can be exchanged along the network ties among them [21]. Knowledge sharing is often motivated by collaborators' expectations of reciprocal relationships [22,23]; satisfaction, sense of belonging, and collective identity can also promote collaborators' contributions to the shared knowledge repositories [24,25,26].

2.2. Collective Reflection in Knowledge Collaboration

Knowledge collaborators rely on reflective thinking to comprehensively understand the knowledge shared by others. Reflective thinking, or reflection, indicates the process in which individuals thoroughly consider previous performance in order to identify deficiencies and make improvements in future actions [27]. In the context of collaborative teamwork, reflection refers to the process that "team members overtly reflect upon the group's objectives, strategies, and processes and adapt them to current or anticipated endogenous or environmental circumstances" [28, p.559]. As suggested by West [29], reflection comprises behaviors such as "questioning, planning, exploratory learning, analysis, diversive exploration, making use of

knowledge explicitly, planfulness, learning at a meta-level, reviewing past events with self-awareness, and coming to terms over time with a new awareness" (p.4). Cognitively, reflection fosters the awareness of knowledge gaps and dissonance, so that individuals can be motivated to seek additional knowledge and make plans for future improvements. Through critical evaluations of strengths and weaknesses, individuals can come to a deeper realization of their current status, and hence can obtain a clearer vision of the future. Individuals often undertake self-reflection in attempting to make sense of new knowledge and recognize current dissonance. Self-reflection facilitates comprehensive analyzing and planning [30]. Through collective interactions, individuals have the opportunity to receive feedback from other discussants [31,32] and thus to reconsider and reevaluate their own performance.

When teams are reflective, they think thoroughly about long-term strategies and consequences, collective performance, as well as environmental factors in order to make advancements in future collaborative work [29]. Based on collective reflection, team collaborators make plans for adjusting next-step actions and goals, and such a process is usually intertwined with the execution of adaptations and the implementation of preplanned objectives [33]. With a blueprint agreed upon, collaborators can conduct goal-oriented actions to accomplish desired changes [29,34].

Reflection is more likely to occur among diverse collaborators than among homogeneous ones. It is important for collaborators to reflect on each other's diverse opinions when attempting to achieve agreements [35]. Collaborators with diverse knowledge domains and skillsets need to reflect on a shared goal in order to maintain a clear path towards this goal. Through reflective thinking, collaborators achieve a better and clearer understanding of each other's roles and contributions [36]; a mutual understanding among heterogenous collaborators is constructed, and thus positive collective outcomes such as collective effectiveness and creativity can be produced [35, 37, 38, 39]. Essentially, collaborators with heterogenous views are more capable of being attentive to various issues which further trigger a consideration of hidden facts and potential alternatives [40,41].

Reflection can be found effective in crowd-based open innovation challenges, whereby a common vision is often set for participants collaborating on solving a broadly defined problem through the exchange of their unique knowledge. Collaborators having a shared vision are inclined to engage in collective reflection, as the shared vision fosters a commitment to the task and encourages risk taking and exploration [42,43]. When collaborators work towards a common goal, their collective reflection can assist shared understanding and enhance the common ground for future communication [35]. Besides, collaborators' communicative skills which help improving the cohesiveness of a team or a community also play an important role because a proactive communication approach can shape collective reflection in a positive way [44, 45].

Reflective interaction spurs innovation [45]. In collective idea generation, individuals conduct reflective thinking to help each other work effectively and creatively [46]. As reflection essentially builds upon a deep processing and critical evaluation of the shared knowledge, creative ideas usually emerge when collaborators are able to reflect on their performance rather than working without reflection [39,47-50]. Therefore, when leveraging reflexivity to crowdsourcing innovation challenges in which diverse crowd members constantly make sense of existing knowledge, actively produce new knowledge, share a common vision, as well as cooperate on finding solutions, the dynamic patterns of crowd members' reflection are worthy of close examination.

2.3. Socio-Semantic Network and Knowledge Sharing

As a manifestation of online collective knowledge sharing and collaborative knowledge management, semantic network has increasingly received scholarly attention. Identifying and motivating expertise contribution and experience sharing, online communication sites have introduced an advanced way of social participation. These sites employ a communicative structure similar to the blogosphere [51, 76], and in such a sphere, communication and collaboration are embodied by participants' non-verbal referencing to each other. For example, online sites for knowledge collaboration such as crowdsourcing innovation challenge platforms are often constructed as a blog-based network in which participants share, exchange and produce knowledge through initiating or commenting on blog-like posts. The blogosphere-type of online community is essentially a socio-semantic network in which each blog can be recognized by both semantic and relational attributes. Relational attribute means that individuals' positions within the network can be configured by their back-forth interactions, whereas semantic attribute refers to the cognitive embodiments displayed in each post [52].

Accordingly, the blogosphere community offers a unique avenue for observing knowledge sharing and knowledge flow. Compared to social networks, topic-oriented blog-based semantic networks enable a closer examination of the influential nodes, emerging topics,

as well as the ever-evolving link structures. Research on community structures of blogosphere, for instance, has presented that the community is maintained by inter-post linkages among blog contributors [53].

The patterns of knowledge sharing are grounded in network structural features of the blog-based virtual community. For example, research has investigated the trend of topic evolution [54], the crowd's sentiment underlying their opinions [55], as well as the co-existents and the cyclic pattern of chatters and spikes in online conversations [56-58]. Viewing blog-based virtual community as a socio-semantic web in which the members co-produce visible symbolic artifacts, crowd-based knowledge collaboration can be investigated by observing these interconnected cognitive artifacts and configurations [59]. In addition, socio-semantic web highlights the human interaction underlying the creation of blog posts, as well as maintains a community in which participants collectively elicit and contribute knowledge to improve their collaborative work [60, 75].

Adopting the perspective of socio-semantic network, the semantic patterns of knowledge collaboration in crowdsourcing innovation challenges can be examined specifically through mapping the evolving knowledge configurations that crowd members' networked interaction displays. In view of these foundational theoretical frameworks, this study proposes:

RQ1: From the perspective of semantic network, what is the pattern of crowd members' collective reflection when they collaboratively share knowledge in online crowdsourcing innovation challenges?

3. Method

This study seeks to examine the semantic representation of knowledge collaboration in online crowdsourcing innovation challenges. Collaborating with two companies that have employed 10-day crowdsourcing tournaments to generate innovative solutions to the companies' strategic problems, this study has harvested a total of 1,116 unique crowd-generated posts. The innovation tournaments were held by a third-party platform provider. Incentives were provided to participants who generated most creative ideas determined by the companies. Both companies were selected because of their top positions in the industry and their pioneering work in co-creating value with customers. The crowdsourcing challenges were open to the public, and there was no limitation on the amount of contribution that each participant could make. The difference in the number of contributions between the challenges, therefore, was naturally determined by the level of participants' activeness. To ensure anonymity, participants from various social and knowledge backgrounds were allowed to create their usernames, so that they could contribute without revealing their identities. As such, this study did not record participants' individual characteristics. In particular, Organization A is a finance company based in the United States, and it crowdsourced for new ideas on the design of its business model; it successfully collected a total of 368 posts through the crowdsourcing innovation challenge. In the ten-day tournament, fifty-one participants contributed to the innovation challenge. This innovation challenge particularly sought ways to reach new customers, establish new supply chains, provide novel customer solutions, find new revenue streams, etc. Organization B is a government-funded non-commercial environmental management organization based in New Zealand, which sought for open strategies; this organization has obtained a total of 748 posts. In the ten-day challenge, a total of ninety-nine individuals made their contributions. Specifically, the innovation challenge attempted to find new environmental management approaches and develop novel strategies to improve environmental services. Previous research showed that the crowd performs differently across various stages of idea generation. For example, the first period of time after an innovation challenge launches may often be used for a cognitive warm-up, whereas towards the end of a time-bound challenge, the potential for producing creative ideas may decline due to exhaustion of sharing and thinking [73, 74]. Therefore, in order to capture the evolution of semantic representations in crowdsourcing, all posts were clustered into three phases based on their unique timestamps assigned by the system according to the unique sequential position of each post in the 10-day discussion, using k-means clustering analysis with MATLAB. In particular, the three phases were termed the early, middle, and late phase of the discussion.

3.1. Data Analysis

For each innovation challenge, a semantic network analysis using text analytics tool Leximancer (https://info.leximancer.com/) was conducted regarding each of the three discussion phases separately. In pre-processing, typical stop words were filtered out so that the remaining words that concretely contributed to the collective meanings can be analyzed. In general, stop words removed from this study include articles, conjunctions, prepositions, and transitive verbs, such as "an", "as", "between", "just", "then", "you", etc.

Using natural language processing techniques, Leximancer first analyzes the occurrence and

frequency of each word and generates several *concepts* that represent a collection of interrelated words. In alignment with Doerfel [8], such concepts refer to a constellation of words appearing together. For instance, the concept "pleased" may contain words like "happy", "glad", and "delighted". Then words and terms are weighted by analyzing the frequency that they appear in a sentence together with the concepts. After generating a list of concepts, this analysis produces a co-occurrence network matrix of all concepts, in which the value in each cell refer to the frequency that two concepts occur together in a single sentence. On the basis of this co-occurrence network, clusters of connected concepts are developed and visualized, with each cluster characterized by a unique theme that reflects the major interest of the crowd [61].

In this study, a list of frequency counts of mostly used concepts was generated, followed by a semantic network analysis demonstrating the emergence and evolution of the semantic attributes of online crowds' knowledge collaboration.

4. Results

The top 10 concepts generated by online crowd are presented in Table 1 and Table 2. Semantic network maps are then presented for visualizing connections among terms as well as concepts. The size of each concept node in the maps indicates the count number of co-occurrence, meaning that compared to small nodes, larger nodes are connected with more concepts and are thus taking a more central position in crowd-generated semantic network of knowledge collaboration.

4.1. Crowdsourcing Innovation Challenge on Finance and Banking

In the innovation challenge on finance and banking, concepts like "financial", "services", "banks", "idea" consistently ranked the top of the list across three phases, naturally because of the business of the company as well as the theme of the company's crowdsourcing innovation challenge (Table 1 & Figure 1). When comparing the concepts across three phases, the early phase was characterized by concepts related to financial literacy as well as basic knowledge of banking.

Table 1. The top 10 most frequently-occurring concepts in Crowdsourcing Innovation Challenge on Finance and Banking

Phase 1		Phase 2		Phase 3	
Concepts	*Relevance*	*Concepts*	*Relevance*	*Concepts*	*Relevance*
financial	100%	financial	100%	financial	100%
services	59%	services	72%	money	89%
money	57%	companies	69%	banks	89%
idea	52%	idea	62%	idea	80%
banks	46%	banks	56%	people	67%
savings	43%	money	47%	credit	59%
literacy	41%	credit	44%	app	50%
account	35%	people	41%	need	50%
users	33%	generation	38%	students	43%
game	33%	media	31%	future	43%

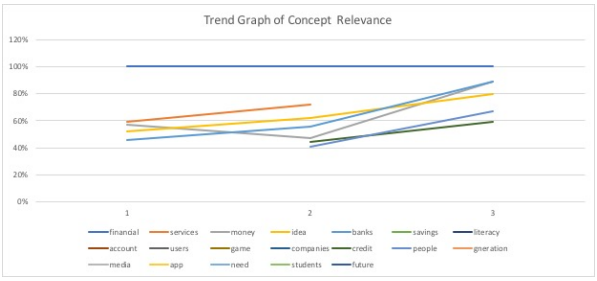

Figure 1. Trend Graph of Concept Relevance

Semantic network analysis was conducted to examine the interconnections among concepts as well as the emergence of major themes (see Figures 2-4). In these network maps, the bubbles represent unique themes and the individual dots inside the bubbles represent major concepts that emerged from crowd's knowledge sharing in the innovation challenge. For each phase, different but overlapping themes were identified. When comparing the knowledge shared across these three phases, it can be found that along with the unfolding of the discussion, members of the crowd develop their thoughts from focusing on basic financial and banking activities (such as "savings") to highlighting the role of technology (such as "apps") in improving financial and banking business models.

Specifically, in the early phase (Figure 2), eight major themes emerged that indicated: 1) ways to facilitate savings, such as developing smartphone applications or help individuals make long-term goals (the bubble on the bottom left); 2) financial and banking services and products (the bubble on the bottom center); 3) usage of financial services (the bubble at the very bottom); 4) the ideas for improving the public's financial literacy such as using games (the bubble on the bottom right); 5) people's money that can be taken care of by financial programs and services (the bubble in the very center); 6) time invested to the management of money (the bubble on the top left); 7) young adults' involvements in financial activities (the bubble on the top center); 8) general information about financial market (the bubble on the top right). An

example of the posts could be "I feel like we need a smarter saving service" or "We should have more diverse types of financial products".

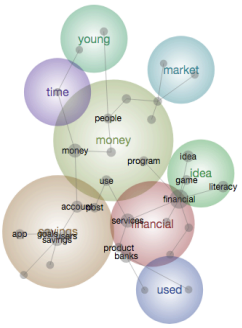

Figure 2. Semantic Network from Financial and Banking Service Crowdsourcing Challenge, Early Phase

Likewise, eight major themes have been identified in the middle phase (Figure 3), demonstrating the content different from that in the early phase: 1) details about financial services such as free service (the bubble at the very bottom); 2) general information about current users of financial services (the bubble on the bottom left); 3) generation-related financial and banking issues (the bubble on the bottom center); 4) information regarding people's financial life (the bubble on the bottom right); 5) banking activities related to credit or debit cards (the bubble on the top left); 6) financial and banking services based on websites or e-platforms (the bubble in the very center); 7) ideas to improve financial services (the bubble on the top center), 8) individuals' financing activities (the bubble on the top right). Although the first theme "financial" has appeared in both the early and the middle phase, the concepts it included has been enriched in the middle phase. Different from the posts in previous phase, posts in the middle phase focused more on technology. For example, participants mentioned "I think banking apps compatibility should be improved" or "Safety should be the top priority when developing banking apps".

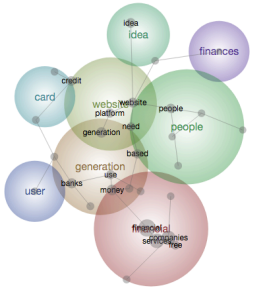

Figure 3. Semantic Network from Financial and Banking Service Crowdsourcing Challenge, Middle Phase

Finally, in the late phase, the theme "financial" (the bubble at the very center) has been further enriched by including new concepts such as "students", "time", "young", and "college". In addition, crowd members in this phase generated seven other themes that demonstrate: 1) need-based ideas for improving financial services (the bubble at the very bottom); 2) the similarity in terms of characteristics of target customers (the bubble on the bottom left); 3) possible smartphone applications that could be developed to provide better banking services (the bubble on the bottom right); 4) banking-related issues such as credit and savings (the bubble on the center right); 5) proposed programs that can facilitate people's daily financial activities (the bubble on the top left); 6) the uses of financial services (the bubble on the top center); 7) millennials who are the major customers of future financial services. In this phase, the focus of discussion switched from service and technology to the future of banking and financial service. For example, participants posted ideas like "I think we should design better service models for Generation Z" or "Younger people are fans of customization".

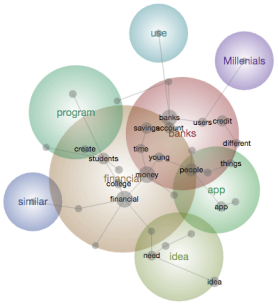

Figure 4. Semantic Network from Financial and Banking Service Crowdsourcing Challenge, Late Phase

To compare the semantic networks generated throughout three phases of crowdsourcing, the analysis demonstrates that concepts like "game" and "workshops" only appeared in the first phase, because both these two can serve as educational tools to improve individuals' financial literacy. In the middle phase, concepts related to "credit", "generation", "website", "score" and "media" emerged, suggesting that as discussion went on, crowd members started to think about in-depth topics that are more relevant to the core theme of innovation challenge. Finally, the late phase was characterized by 1) the production of solutions to current problems, such as "app", "savings", "students", "college", and 2) the emergence of the concept like "future" that demonstrated a collective attention given to the long-term development of the company.

4.2. Crowdsourcing Innovation Challenge on Environmental Management

In the innovation challenge on environmental management, the evolution of the content indicates that crowd members are reflectively involved in the discussion (see Figures 6-8). The most frequently-occurring concepts in this open innovation challenge were presented in Table 2 and Figure 5.

Table 2. The top 10 most frequently-occurring concepts in Crowdsourcing Innovation Challenge on Environmental Management

Phase 1		Phase 2		Phase 3	
Concepts	*Relevance*	*Concepts*	*Relevance*	*Concepts*	*Relevance*
pest	100%	pest	100%	pest	100%
need	52%	control	68%	traps	53%
species	48%	possums	66%	need	53%
areas	47%	traps	50%	rats	52%
control	46%	need	50%	land	48%
possums	43%	species	46%	control	45%
eradication	39%	areas	45%	use	42%
cats	39%	rats	43%	possums	38%
native	37%	cats	42%	species	29%
people	35%	research	31%	work	29%

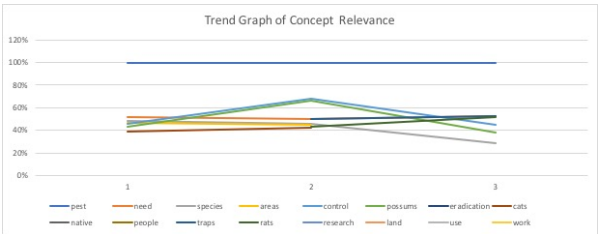

Figure 5. Trend Graph of Concept Relevance

Throughout the early, middle and late phases, concepts such as "pest", "control", "need", "species", and relevant animals like "possums", "cats", "birds" and "rats" remained frequently used. The top-ranked theme "eradication" in the early phase had fallen in the middle phase, and finally disappeared in the late phase. The concept "traps" which had not been mentioned very frequently in the early phase, became popular in both middle and late phases. More interestingly, the concept "research" only occurred in the middle phase rather than in the early or late phases. Such a cross-phase evolution of concepts indicates that crowd members learn one another's thoughts during crowdsourcing and make attempts to refine their own ideas. For example, in the early phase, eradication was considered as a useful pest management approach by most members of the crowd, however, after a period of thought-provoking discussion, the crowd started to realize that eradication was difficult and that using traps might be a more efficient approach. Along with the discussion, crowd members also realized the need of more research because a satisfactory environmental management relies on scientific tools.

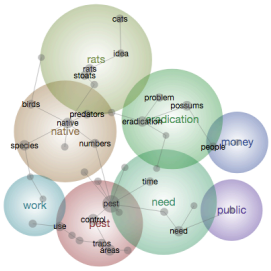

Figure 6. Semantic Network from Environmental Management Crowdsourcing Challenge, Early Phase

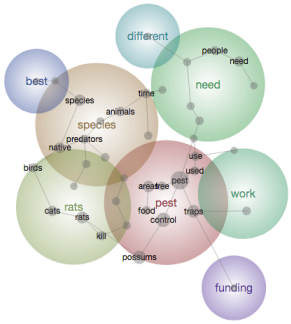

Figure 7. Semantic Network from Environmental Management Crowdsourcing Challenge, Middle Phase

When comparing the semantic networks generated throughout early, middle and late phase of crowdsourcing, several themes remained consistent whereas others evolved. In the early phase, unique themes were identified as 1) the general concern or life status of the public (the bubble on the bottom right); 2) management of relevant native species and animals (the bubble on the top left); 3) the possibility of eradication of pests (the bubble at the top right). In the middle phase, several new themes emerged such as: 1) the focus of people's need (the bubble on the top right); 2) the balance among various requests raised by different people (the bubble at the very top); 3) the best approach to manage a variety of species (the bubble on the top left). Finally, in the late phase, four new themes emerged that indicated: 1) the appropriate time of implementing pest control strategies (the bubble on the top left); 2) the appropriate tools for pest control (the bubble on the center left); 3) the use of pest management strategies (the bubble on the top left); 4)

the agreement on balancing pest management and people's needs (the bubble on the bottom right). Taken together, the content variation throughout different phases indicates that as the crowdsourcing unfolded, the crowds' interests developed from calling for eradication efforts to striking a balance between protecting human habitat and animal habitat, and ultimately the crowd reached some agreements in terms of tools and timing of pest management. For example, in the first phase, participants articulated, "we should find a balance between pest control and environmental protection", whereas in the second phase, participants considered the importance of scientific research and mentioned "more research on environmental management is needed", and finally in the third phase, some consensus was reached, such as "I agree that using traps at appropriate times will be good for our environment".

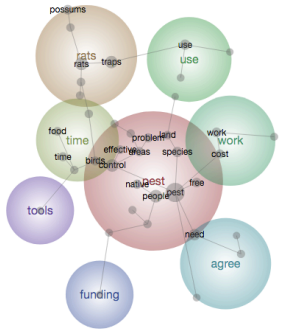

Figure 8. Semantic Network from Environmental Management Crowdsourcing Challenge, Late Phase

5. Discussion

5.1. Findings and Implications

This study mainly demonstrates semantic patterns of crowd members' knowledge collaboration as well as collective reflection when responding to open innovation challenges. A total of 1,116 crowd-sourced online posts generated by organization-sponsored innovation challenges were analyzed, exhibiting the dynamic attributes of crowdsourced knowledge collaboration. First, the emerging themes, concepts, as well as the semantic networks in which those concepts and themes were interconnected jointly indicated that knowledge contributors cognitively reference each other when engaging in collective innovation. Second, a comparison of semantic networks generated across different discussion phases reveals that members of the crowd have collectively undertaken reflective thinking in the course of ongoing discussion.

Several thematic patterns of crowdsourcing have emerged from crowd members' collective knowledge collaboration. First, findings suggested that the frequencies of themes differed across three phases, representing heterogeneous symbolic artifacts that participants had collectively produced. These symbolic artifacts can be viewed as manifestations of the common knowledge shared within the crowd [52]. In the open innovation challenges studied, the crowds collectively generated symbolic artifacts that were commonly used throughout the entire discussion (for example, "banks", "account" in the financial service innovation challenge; "traps", "species", "control" in the environmental management innovation challenge), as well as several symbolic artifacts uniquely used in each phase (for example, "game", "generation", "college" in the financial service innovation challenge; "fences", "research", "land" in the environmental management innovation challenge). Meanwhile, as shown in semantic co-occurrence networks, all the symbolic artifacts were embedded in a socio-semantic network in which all posts were connected to one another, and the network structures were constantly evolving as new topics emerged [52, 62]. Such a topical interlinking tendency [62, 63] facilitates sustaining online crowdsourcing discussion for generating solutions to innovation challenges.

Furthermore, the evolving semantic networks across three crowdsourcing phases revealed that crowd members constantly engage in collective reflection. In alignment with the literature on reflection and collective reflexivity [28,29,64], this study revealed that reflection takes place when diverse members of the crowd engage in a collaboration driven by a common goal, as well as when the actions occur in a friendly and encouraging environment that makes collaborators feel safe to share unique opinions. The innovation challenges studied in the present research both demonstrated that crowd members refer to each other's comments and take each other's perspectives to develop their own thinking, and thus collectively advance their ongoing discussion. In accordance with the theories [37-39,47-50], collective reflection facilitates setting a shared vision amongst crowdsourcing participants. For example, in the innovation challenge where the theme was to seek solutions on financial service business models, a comparison of three co-occurrence networks demonstrated that, as the discussion unfolded, crowd members constantly adjusted their common visions and switched their focus from financial and banking apps to different categories of customers, and finally to the broader banking product markets. Likewise, in the challenge where the theme was to harvest innovative solutions in terms of environmental management, a

comparison of the co-occurrence networks across three phases revealed that the crowd's common vision changed along the progress of collective discussion. In the first phase, the common vision was about eradication; in the second phase, the common vision became the issue of funding; and finally, in the third phase, the common vision was about emerging tools that could be used for pest control and environmental management.

This study makes several contributions. Theoretically, it adopts socio-semantic network perspective in examining the dynamic patterns of collective reflection in crowdsourcing for innovation. Viewing collective knowledge collaboration as a socio-semantic network, this study unpacks the dynamics of how knowledge artifacts are networked and evolving along with contributors' collective reflection. It highlights the importance of incorporating a semantic dimension into the research on crowd-level collective reflexivity emerging from open innovation challenges. From the information systems perspective, this study highlights that the crowd's information processing can be evolving and reflective and therefore calls for future research on the dynamic nature of online crowds' collective information processing in open innovation challenges. Practically, findings of this research indicate that crowdsourcing practitioners should be attentive to the semantic connections occurring in knowledge contributors' back-forth online conversations, as the ever-evolving semantic networks manifest the reflection that crowd members collectively undertake. Crowdsourcing helps organizations adjust their business goals and open strategies; for example, organizations in this study can improve the quality of financial services provided to different generations or develop pet-friendly environmental management plans based on crowdsourcing findings. Designers of information systems can benefit from the semantic characteristics of crowd-based knowledge collaboration to implement the infrastructure that better identifies new ideas and facilitates its emergence. Understanding how the crowd's opinion evolves will help business practitioners to effectively integrate the wisdom of crowds into future managerial actions.

5.2. Future Research

This study is limited in several ways and thus calls for further research. To extend the current findings, beyond using a limited number of crowdsourcing innovation challenges held within a specific time frame, semantic network analysis should be applied to a broader context in order to demonstrate the crowd's collective knowledge collaboration in natural settings. Furthermore, based on the findings from this research, future studies should experimentally test the causes and effects of collective reflection in the context of crowdsourcing knowledge collaboration, as well as the role played by various semantic components in leading up to collective innovation on web-based crowdsourcing platforms.

6. References

[1] Afuah, A. and Tucci, C. L. (2012), "Crowdsourcing as A Solution to Distant Search", Academy of Management Review, 37(3), pp. 355-375.

[2] Brabham, D. C. (2013). Crowdsourcing. MIT Press.

[3] boyd, D. M. and Ellison, N. B. (2007), "Social network sites: Definition, history, and scholarship", Journal of Computer-Mediated Communication, 13(1), pp. 210-230.

[4] Faraj, S., Jarvenpaa, S. L., and Majchrzak, A. (2011), "Knowledge collaboration in online communities", Organization Science, 22(5), pp. 1224-1239.

[5] Collins, A. M. and Quillian, M. R. (1969), "Retrieval time from semantic memory", Journal of Verbal Learning and Verbal Behavior, 8(2), pp. 240-247.

[6] Barnett, G. A. and Woelfel, J. (1988). Reading in the Galileo System: Theory. Methods and Applications. Kendall/Hunt, Dubuque, IA.

[7] Chang, T. M. (1986), "Semantic memory: Facts and models", Psychological Bulletin, 99(2), 199-220.

[8] Doerfel, M. L. (1998), "What constitutes semantic network analysis? A comparison of research and methodologies", Connections, 21(2), pp. 16-26.

[9] Doerfel, M. L. and Barnett, G. A. (1999), "A semantic network analysis of the International Communication Association", Human Communication Research, 25(4), pp. 589-603.

[10] Rice, R. E. and Danowski, J. A. (1993), "Is it really just like a fancy answering machine? Comparing semantic networks of different types of voice mail users", The Journal of Business Communication, 30(4), pp. 369-397.

[11] Corman, S. R., Kuhn, T., McPhee, R. D., and Dooley, K. J. (2002), "Studying complex discursive systems. Centering resonance analysis of communication", Human Communication Research, 28(2), pp. 157-206.

[12] Baumeister, R. F. and Leary, M. R. (1995). "The need to belong: desire for interpersonal attachments as a fundamental human motivation", Psychological Bulletin, 117(3), 497-529.

[13] Nelson, K. M. and Cooprider, J. G. (1996), "The contribution of shared knowledge to IS group performance", MIS Quarterly, 20(4), pp. 409-432.

[14] 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.

[15] Kane, G. C., Johnson, J. and Majchrzak, A. (2014), "Emergent life cycle: The tension between knowledge change and knowledge retention in open online coproduction communities", Management Science, 60(12), pp. 3026-3048.

[16] Wagner, C. and Majchrzak, A. (2006), "Enabling customer-centricity using wikis and the wiki way", Journal of Management Information Systems, 23(3), pp. 17-43.
[17] Chester, A. and Bretherton, D. (2007), "Impression management and identity online", The Oxford Handbook of Internet Psychology, pp. 223-236.
[18] Wallace, P. (1999/2015). The psychology of the Internet. Cambridge University Press.
[19] Chiu, C. M., Hsu, M. H. and Wang, E. T. (2006), "Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories", Decision Support Systems, 42(3), pp. 1872-1888.
[20] Bandura, A. (1989), "Human agency in social cognitive theory", American Psychologist, 44(9), pp. 1175-1184.
[21] Wellman, B. and Wortley, S. (1990), "Different strokes from different folks: Community ties and social support", American Journal of Sociology, 96(3), pp. 558-588.
[22] Bock, G. W., Zmud, R. W., Kim, Y. G. and Lee, J. N. (2005), "Behavioral intention formation in knowledge sharing: Examining the roles of extrinsic motivators, social-psychological factors, and organizational climate", MIS Quarterly, 29(1), pp. 87-111.
[23] Kankanhalli, A., Tan, B. C. and Wei, K. K. (2005), "Contributing knowledge to electronic knowledge repositories: An empirical investigation", MIS Quarterly, 29(1), pp.113-143.
[24] Dholakia, U. M., Bagozzi, R. P. and Pearo, L. K. (2004), "A social influence model of consumer participation in network-and small-group-based virtual communities", International Journal of Research in Marketing, 21(3), pp. 241-263.
[25] Hars, A. and Ou, H. (2002), "Working for free motivations of participating in open source software projects", Proceedings of the 34th Hawaii International Conference on System Sciences, pp. 25-31.
[26] Yoo, W. S., Suh, K. S. and Lee, M. B. (2002), "Exploring the factors enhancing member participation in virtual communities", Journal of Global Information Management, 10(3), pp. 55-71.
[27] Grant, A. M., Franklin, J. and Langford, P. (2002), "The self-reflection and insight scale: A new measure of private self-consciousness", Social Behavior and Personality: An International Journal, 30(8), pp. 821-835.
[28] West, M.A. (1996). "Reflexivity and work group effectiveness: A conceptual integration", The handbook of work group psychology, John Wiley & Sons, Ltd. pp. 555-579.
[29] West, M.A. (2000), "Reflexivity, revolution and innovation in work teams", In M.M. Beyerlein, D.A. Johnson and S.T. Beyerlein (Eds.), Product Development Teams, 5, 1–29. Stamford, CT: JAI Press.
[30] Murray, L., Hourigan, T. and Jeanneau, C. (2007), "Blog writing integration for academic language learning purposes: Towards an assessment framework", Ibérica, Revista de la Asociación Europea de Lenguas para Fines Específicos,14, pp.9-32.
[31] Liu, C. and Tsai, C. (2005), "Peer assessment through web-based knowledge acquisition: Tools to support conceptual awareness", Innovations in Education and Teaching International, 42(1), pp. 43–59.
[32] Wiecha, J. M., Gramling, R., Joachim, P. and Vanderschmidt, H. (2003), "Collaborative e-learning using streaming video and asynchronous discussion boards to teach the cognitive foundation of medical interviewing: a case study", Journal of Medical Internet Research, 5(2), e13.
[33] Weingart, L. R. (1992). Impact of group goals, task component complexity, effort, and planning on group performance. Journal of Applied Psychology, 77(5), pp. 682-693.
[34] Hacker, W. (2003), "Action regulation theory: A practical tool for the design of modern work processes?", European Journal of Work and Organizational Psychology, 12(2), pp. 105-130.
[35] Schippers, M. C., Den Hartog, D. N., Koopman, P. L. and Wienk, J. A. (2003), "Diversity and team outcomes: The moderating effects of outcome interdependence and group longevity and the mediating effect of reflexivity", Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior, 24(6), pp. 779-802.
[36] Dyer, W. G. (1995). Team building: Current issues and new alternatives (3rd ed.), New York, NY: Addison-Wesley Publishing Company.
[37] De Dreu, C. K. (2002), "Team innovation and team effectiveness: The importance of minority dissent and reflexivity", European Journal of Work and Organizational Psychology, 11(3), pp. 285-298.
[38] Gibson, C. and Vermeulen, F. (2003), "A healthy divide: Subgroups as a stimulus for team learning behavior", Administrative Science Quarterly, 48(2), pp. 202-239.
[39] Tjosvold, D., Tang, M. M. and West, M. (2004), "Reflexivity for team innovation in China: The contribution of goal interdependence", Group & Organization Management, 29(5), pp. 540-559.
[40] González-Romá, V. and Hernández, A. (2014), "Climate uniformity: Its influence on team communication quality, task conflict, and team performance", Journal of Applied Psychology, 99(6), pp. 1042-1058.
[41] Williams, K. Y. and O'Reilly, C. (1998), "Forty years of diversity research: A review", Research in Organizational Behavior, 20, pp. 77-140.
[42] Agrell, A. and Gustafson, R. (1994), "The Team Climate Inventory (TCI) and group innovation: A psychometric test on a Swedish sample of work groups", Journal of Occupational and Organizational Psychology, 67(2), pp. 143-151.
[43] Senge, P. M. (1997), "The fifth discipline", Measuring Business Excellence, 1(3), pp. 46-51.
[44] Hoegl, M. and Parboteeah, K. P. (2006), "Team reflexivity in innovative projects", R&D Management, 36(2), pp. 113-125.
[45] Schippers, M. C., West, M. A. and Dawson, J. F. (2015), "Team reflexivity and innovation: The moderating role of team context", Journal of Management, 41(3), pp. 769-788.
[46] Paulus, P. B. and Yang, H. C. (2000), "Idea generation in groups: A basis for creativity in organizations", Organizational Behavior and Human Decision Processes, 82(1), pp. 76-87.
[47] Carter, S. M. and West, M. A. (1998), "Reflexivity, effectiveness, and mental health in BBC-TV production teams", Small Group Research, 29(5), pp. 583-601.

[48] De Dreu, C. K., Nijstad, B. A. and van Knippenberg, D. (2008), "Motivated information processing in group judgment and decision making", Personality and Social Psychology Review, 12(1), pp. 22-49.
[49] Hülsheger, U. R., Anderson, N. and Salgado, J. F. (2009), "Team-level predictors of innovation at work: a comprehensive meta-analysis spanning three decades of research", Journal of Applied Psychology, 94(5), pp. 1128-1145.
[50] Wong, A., Tjosvold, D. and Su, F. (2007), "Social face for innovation in strategic alliances in China: The mediating roles of resource exchange and reflexivity", Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior, 28(8), pp. 961-978.
[51] Hookway, N. (2008), "Entering the blogosphere': some strategies for using blogs in social research", Qualitative Research, 8(1), pp. 91-113.
[52] Roth, C. and Cointet, J. P. (2010), "Social and semantic coevolution in knowledge networks", Social Networks, 32(1), pp.16-29.
[53] Kumar, R., Raghavan, P., Rajagopalan, S. and Tomkins, A. (1999), "Trawling the Web for emerging cyber-communities", Computer Networks, 31(11-16), pp.1481-1493.
[54] Glance, N., Hurst, M., & Tomokiyo, T. (2004). Blogpulse: Automated trend discovery for weblogs. WWW 2004 workshop on the weblogging ecosystem: Aggregation, Analysis and Dynamics.
[55] Mishne, G. and De Rijke, M. (2006), "A study of blog search", European Conference on Information Retrieval, Springer, Berlin, Heidelberg, pp. 289-301.
[56] Balog, K., Mishne, G. and De Rijke, M. (2006), "Why are they excited? identifying and explaining spikes in blog mood levels", Demonstrations, pp.207-210.
[57] Bansal, N. and Koudas, N. (2007), "Blogscope: a system for online analysis of high volume text streams", Proceedings of the 33rd international conference on Very large data bases VLDB Endowment, pp. 1410-1413.
[58] Gruhl, D., Guha, R., Liben-Nowell, D. and Tomkins, A. (2004), "Information diffusion through blogspace", Proceedings of the 13th international conference on World Wide Web ACM, pp. 491-501.
[59] Schoop, M., Moor, A. D. and Dietz, J. L. (2006), "The pragmatic web: a manifesto", Communications of the ACM, 49(5), pp. 75-76.
[60] Cahier, J. P., Zaher, L. H. and Zacklad, M. (2007), "Information seeking in a socio-semantic web application", Proceedings of the 2nd International Conference on Pragmatic Web, ACM, pp. 91-95.
[61] Smith, A. E. and Humphreys, M. S. (2006), "Evaluation of unsupervised semantic mapping of natural language with Leximancer concept mapping", Behavior Research Methods, 38(2), pp. 262-279.
[62] Roth, C. (2013), "Socio-semantic frameworks", Advances in Complex Systems, 16, pp. 1350013.
[63] Dokoohaki, N. and Matskin, M. (2008), "Effective design of trust ontologies for improvement in the structure of socio-semantic trust networks", International Journal on Advances in Intelligent Systems, 1, pp. 23-42.
[64] West, M. A. and Richter, A. W. (2008), "Climates and cultures for innovation and creativity at work", Handbook of Organizational Creativity, pp. 211-236.
[65] Chesbrough, H. (2003). Open innovation: The new imperative for creating and profiting from technology. Boston, MA: Harvard Business School Press.
[66] Parker, G., and Van Alstyne, M. (2018). Innovation, openness, and platform control. Management Science, 64(7), pp.3015-3032.
[67] Dahlander, L. and Gann, D. M. (2010). "How open is innovation?", Research Policy, 39(6), pp.699-709.
[68] Afuah, A., and Tucci, C.L. (2012). Crowdsourcing as a Solution to Distant Search. Academy of Management Review, 37(3), pp. 355–375.
[69] Jeppesen,L.B. and Lakhani,K.R. (2010). Marginality and Problem-Solving Effectiveness in Broadcast Search. Organization Science, 21(5), pp. 1016–1033.
[70] Piezunka, H., and Dahlander, L. (2015). Distant search, narrow attention: How crowding alters organizations' filtering of suggestions in crowdsourcing. Academy of Management Journal, 58(3), pp. 856-880.
[71] Ali, S., Green, P., and Robb, A. (2013). "Measuring Top Management's IT Governance Knowledge Absorptive Capacity," Journal of Information Systems, 27(1), pp. 137-155.
[72] Bassellier, G., and Benbasat, I. (2004). "Business Competence of Information Technology Professionals: Conceptual Development and Influence on It-Business Partnerships," MIS Quarterly, 28(4), pp. 673-694.
[73] Dahlander, L. and Piezunka, H. (2014), "Open to suggestions: How organizations elicit suggestions through proactive and reactive attention", Research Policy, 43(5), pp. 812-827.
[74] Fleming, L. (2001), "Recombinant Uncertainty in Technological Search", Management Science, 47(1), pp. 117-132.
[75] Zarrouk, M., Lafourcade, M., and Joubert, A. (2013). "Inference and reconciliation in a crowdsourced lexical-semantic network", Computing System, 17(2), pp.147-159.
[76] Xu, Z., Liu, Y., Xuan, J., Chen, H., and Mei, L. (2017). "Crowdsourcing based social media data analysis of urban emergency events". Multimedia Tools and Applications, 76(9), pp. 11567-11584.