Journal of Management Information Systems

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/mmis20

Stigmergy in Open Collaboration: An Empirical Investigation Based on Wikipedia

Lei (Nico) Zheng, Feng Mai, Bei Yan & Jeffrey V. Nickerson

To cite this article: Lei (Nico) Zheng, Feng Mai, Bei Yan & Jeffrey V. Nickerson (2023) Stigmergy in Open Collaboration: An Empirical Investigation Based on Wikipedia, Journal of Management Information Systems, 40:3, 983-1008, DOI: 10.1080/07421222.2023.2229119

To link to this article: https://doi.org/10.1080/07421222.2023.2229119

JOURNAL OF MANAGEMENT INFORMATION SYSTEMS
2023, VOL. 40, NO. 3, 983–1008
https://doi.org/10.1080/07421222.2023.2229119

Stigmergy in Open Collaboration: An Empirical Investigation Based on Wikipedia

Lei (Nico) Zheng [a], Feng Mai [b], Bei Yan [b], and Jeffrey V. Nickerson [b]

[a]College of Management, University of Massachusetts Boston, Boston, MA, USA; [b]School of Business, Stevens Institute of Technology, Hoboken, NJ, USA

ABSTRACT

Participants in open collaboration communities coproduce knowledge despite minimal explicit communication to coordinate the efforts. Studying how participants coordinate around the knowledge artifact and its impacts are critical for understanding the open knowledge production model. This study builds on the theory of *stigmergy*, wherein actions performed by a participant leave traces on a knowledge artifact and stimulate succeeding actions. We find that stigmergy involves two intertwined processes: *collective modification* and *collective excitation*. We propose a new measure of stigmergy based on the spatial and temporal clustering of contributions. By analyzing thousands of Wikipedia articles, we find that the degree of stigmergy is positively associated with community members' participation and the quality of the knowledge produced. This study contributes to the understanding of open collaboration by characterizing the spatial-temporal clustering of contributions and providing new insights into the relationship between stigmergy and knowledge production outcomes. These findings can help practitioners increase user engagement in knowledge production processes in order to create more sustainable open collaboration communities.

KEYWORDS

Open collaboration; collaboration communities; stigmergy; work coordination; Wikipedia; community participation; information quality; Moran's I; knowledge production

Introduction

Open collaboration communities such as Wikipedia and open-source software projects are new kinds of knowledge production enabled by information technology. This knowledge production model has attracted considerable attention because both scholars and practitioners want to understand how collective knowledge production can occur on digital platforms without formal organizational structures or coordination mechanisms [41]. Compared to conventional organizations, these online communities are characterized by fluid membership, high turnover rates, dependence on voluntary contributions, and a lack of formal structure and face-to-face interactions [29]. Because of this, coordination methods that are commonly used to organize tasks in conventional organizations are often

CONTACT Lei (Nico) Zheng lzheng9@stevens.edu College of Management, University of Massachusetts Boston, 100 Morrissey Blvd., Boston, MA 02125, USA.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/07421222.2023.2229119

inapplicable, or even in conflict with contributors' needs for openness, autonomy, and egalitarianism [45, 50].

There is growing recognition that the key to successful open collaboration is coordination via shared knowledge artifacts. Through co-created knowledge artifacts, large numbers of geographically dispersed participants can build knowledge on top of each other's work without explicit communication [40] and can self-organize based on the state of the artifacts [4, 5, 44, 50]. Such artifact-centric coordination is helpful because it allows effective coordination in a variety of situations where traditional command and control structures are not tenable [21]. For example, members of a community may be separated from each other because of time zone differences or work patterns. With thousands of people involved in a project, universal point-to-point communication is not feasible [26]. However, despite the increasing recognition and adoption of open collaboration knowledge production [4, 5, 40, 44], how participants self-organize via knowledge artifacts to produce coherent, high-quality knowledge remains unclear.

To shed light on this issue, IS scholars have proposed the concept of *stigmergy* to explain the role of shared knowledge artifacts in open collaboration [4, 9, 21]. Originally a term used in biology [37], stigmergy is a mechanism of coordination in which an action performed by an agent leaves traces on the environment that stimulate succeeding actions by the same or another agent. For example, when termites build nests, the pheromone left on the existing soil pellets stimulates other termite builders to drop their pellets in the same spot, constructing skyscraper-like pillars [25]. This mechanism can also be used to explain the coordination that occurs in self-organized human collectives where participants are physically dispersed and cannot communicate synchronously. For instance, in Wikipedia, editors build knowledge on top of each other's work to develop articles [40, 59].

Although the concept of stigmergy has the potential to advance our knowledge of self-organized coordination in contexts where direct communication is limited, the investigation is still in an early stage. Existing theorization [9, 16, 21] and operationalization [20, 59] efforts have proposed stigmergy as a form of self-organization enabled by knowledge artifacts. Yet the consequences of such coordination remain vague. This paper extends prior research by further clarifying its processes and conducting an empirical study to answer the following research question:

> Research Question: *How is stigmergy related to participation and quality in online knowledge production?*

Our inquiry into this question is motivated by both theoretical and practical considerations. With respect to theory, current research contends that edits in knowledge artifacts without explicit communication constitute stigmergic coordination [20, 59]. However, an absence of explicit coordination is not evidence that such absence is always beneficial. The platform features that afford stigmergy and thereby enable any community member to change content and render the changes universally discoverable can also provoke clashes of opinions, edit wars, or deliberate vandalism [13]. Current stigmergy studies in information systems do not clearly demonstrate that the modification process is coordinated. Nor do the studies weigh in the extent to which stigmergy is effective in affecting desirable outcomes. There is also not currently a way to distinguish between activities with high versus low degrees of stigmergy. For the theory to be improved, we need to dive deeper into the constituent processes of stigmergy in the context of open collaboration.

With respect to practice, answers to the question may help address two major challenges facing open collaboration communities: improving information quality and encouraging participation. On Wikipedia, only 0.55 percent of the articles are considered by human raters to be good [68]. Because articles produced through open collaborations have become a leading source of information both for the general public and for artificial intelligence (AI) systems, increases in the quality of open collaboration can have positive societal consequences [39, 67, 69]. Moreover, open collaboration communities face challenges related to participation. For example, more volunteers are leaving than joining Wikipedia since 2008 [3], putting increased pressure on platforms to keep participants engaged. For platforms that plan to use stigmergy to enhance participation and quality, it is vital to better understand its constituent processes. It also will be helpful to have ways of measuring stigmergy in order to empirically examine how it impacts knowledge production performance in open collaboration platforms.

To address the research question, we propose that underlying stigmergy are two related processes. The first is a *collective modification* process, which allows an artifact to be changed by members of a community [9, 40, 44]. The second is a *collective excitation* process, in which work done by one participant stimulates further contributions from other participants [25, 38]. The collective excitation process highlights the spatial-temporal clustering of contributions inherent in stigmergy. When stigmergy occurs, participants coordinate by interpreting tasks and task dependencies that emerged from changes to a shared artifact [9]. Therefore, participant contributions will be concentrated in adjacent sections of the knowledge artifact and follow each other in time. Stigmergy thus can be observed and measured via the spatial-temporal clustering of participant activities with respect to the shared knowledge artifacts.

We use a qualitative example to illustrate this new theoretical insight and develop a new measure of stigmergy by adapting Moran's I to capture stigmergy's spatial and temporal characteristics [55]. We then analyze a large dataset with over 1 million revisions of 2,275 Wikipedia articles within the Apple Inc. WikiProject between 2001 and 2017. We find that the degree of stigmergy is positively associated with both community members' participation and the information quality of the knowledge produced. Our findings are robust to different model specifications and alternative samples.

This study makes several contributions to the literature on open collaboration. First, previous literature suggests that editors coordinate around the knowledge artifact [4, 5, 40, 44, 50] and stigmergy may play an important role in coordinating the collective effort [9, 19]. However, there is a lack of clarity on the processes at work and their consequences. We extend this line of research [9, 19] by explaining stigmergy among editors as a pair of intertwined processes. There is a process of *collective modification*, which allows artifacts to be constructed through the joint efforts of community members. There is also a process of *collective excitation*, through which members become aware of each other's edits, edit in reaction to those edits, and excite more edits as a result.

Second, we develop a new measure of stigmergy that reflects the spatial and temporal clustering that happens when the modification process changes an artifact, which leads to more edits as a result of the excitation process. Understanding how stigmergy affects online knowledge production is limited due to the lack of reliable empirical measures [9]. With the new measure, our empirical analyses show that stigmergy is strongly and positively associated with both user participation and production quality. The positive associations of stigmergy and knowledge production outcomes show that stigmergy serves not only as a coordination mechanism for modifying an artifact but may also stimulate continuous participation and quality improvement in open collaborations. Lastly, through a large-sample study, we identify new avenues for improving open collaboration and retaining contributors on digital platforms. This study offers prescriptive insights to platform managers and project owners of collaboration platforms who want to increase participation in knowledge production processes.

Background

Coordination in Open Collaboration Communities

Coordination among a group of people is the orchestration of interdependent actions and fundamentally concerns the management of dependencies [30, 52, 53, 61]. Coordination is essential in open collaboration because of task dependency—numerous participants are working on the same artifact together without face-to-face interaction or role/task assignment, but the contributions they make need to fit with one another to produce a coherent final product. Although the participants may never meet or know one another in person, they nonetheless need to coordinate their contributions with others to ensure their work complements rather than duplicates or contradicts others' work to produce a complete final product. Consider the co-editing process of a Wikipedia article as an example: someone adds a piece of information to the article; others will soon review the information, and some of them will modify it by adding external references, neutralizing the tone, and integrating it with the rest of the article—by doing these, they coordinate with each other and manage task dependencies.

Although coordination is essential in open collaboration or any form of collaborative work, it may take different forms compared to traditional organizational contexts where people interact directly face-to-face or using virtual communication tools. Explicit coordination methods—schedules, plans, hierarchies, and feedback through communication developed for conventional organizations [27, 28]—are less applicable to open collaboration communities for several reasons. First, contributors to open collaboration communities are volunteers who work across temporal and geographical distances. Even though they may use communication media such as mailing lists and notice boards to coordinate [7, 11, 32], they often cannot avail themselves of the standard working hours and regular meetings that are common practices in conventional organizations. Second, collaboration in online communities lacks a formal organizational structure, monetary incentives, and leaders who control compensation and assign tasks [32], making coordination through hierarchies less feasible [50]. Third, collaboration in online communities usually involves a large number of individuals with diverse knowledge backgrounds [2, 11]. However, with the ongoing turnover of participants [29, 58] and low levels of task familiarity [35], it can be challenging for participants to use explicit coordination methods [65] or develop a shared cognition (e.g., transactive memory system, situational awareness) [17, 23, 28] through interpersonal interactions.

Stigmergy in Open Collaboration Communities

With limited explicit coordination, the literature posits that the shared knowledge artifacts co-created by the participants play a central role in coordinating open collaboration [4, 5, 40, 44]. Specifically, participants coordinate and self-organize by interacting with the knowledge artifact such as a Wikipedia article [40, 41]. These interactions leave digital traces that others can see, enabling participants to learn about the task, task dependencies, and information about others in order to coordinate [9, 19]. For example, in Wikipedia, editors' contribution behavior depends on the life cycle of an article [42], and more developed content reduces editors' tendency to contribute [44]. Roles and routines are frequently redefined based on changes to the knowledge artifact [4, 50]. Despite the increased recognition of the importance of knowledge artifacts in coordination, the mechanism that enables artifact-centric coordination is not obvious.

In an attempt to explain artifact-centric coordination, scholars have theorized the mechanism as stigmergy [9, 19, 21], analogous to the biological mechanism that coordinates behaviors of social insects such as ants and bees. Stigmergy in biology is defined as the "stimulation of workers by the performance they have achieved" [37, p.43, 25, p.852]. For example, in ant foraging, the pheromone trail laid by a returning forager stimulates other outgoing foragers to follow the path to the most promising food source [36, 62]. When adopted by the IS field, stigmergy refers to a coordination mechanism in which "actors are leaving traces of their actions in the code (artifact) and they are reading and reflecting on the code written by others (changes made to the artifact) in order to take coordinated action" [9, p. 20]. Stigmergy has been observed in traditional groups such as construction workers [14–16], as well as in open-source software projects [9], data science teams [21], and Wikipedia [26, 59]. For example, when a new version of the iPhone came out, someone added a sentence in the iPhone article in Wikipedia that said, "*In early September, Apple released iPhone 7*." After reviewing the sentence and its context, other editors soon adopted various ways to improve it. They extended this sentence by adding more descriptions, referring to external sources, and revising the writing style, which eventually expanded the initial seven-word sentence into a comprehensive introduction to the new iPhone. As a result, without any planning or direct communication, editors were able to coordinate with others and manage task dependencies (e.g., add descriptions, add citations, etc.) by following and observing the emerging changes in the knowledge artifact [9].

Stigmergy offers a compelling explanation of how artifact-centric coordination unfolds dynamically. Most existing studies focus on theoretical development illustrated by case studies [9, 19]. This is because the process of stigmergy is difficult to observe and measure as participants dynamically adjust their behaviors based on their observations and interpretations of others' behavior [9]. In the next section, we further develop the concept of stigmergy by differentiating its two underlying processes. We then propose hypotheses regarding the impact of stigmergy on online knowledge production outcomes.

Theory and Hypothesis Development

Processes Underlying Stigmergy

Existing information systems (IS) literature theorizes stigmergy as coordination and self-organizing via the shared knowledge artifact. This is what we call the *collective* *modification* process because this process invites everyone to make changes and observe the changes. This process appears well-understood in the IS literature [4, 5, 40, 44]. In the biological literature, stigmergy also involves a *collective excitation* process that regulates behaviors. As Heylighen writes, "A process is stigmergic if the work ('ergon' in Greek) done by one agent provides a stimulus ('stigma') that entices other agents to continue the job" [38, p. 7]. In ant colonies, such excitation is achieved by the pheromones diffusing chemical substances in the environment to stimulate others to follow the action [37]. For example, termite builders use soil pellets soaked with pheromones to construct nests. The nest-building takes two phases [25]. Initially, individual builders randomly drop pellets on the ground. The dropped pellets keep emitting pheromones that attract other builders to leave their pellets in the same spot. When one of the piles reaches a critical size, a coordination phase takes place. In the coordination phase, every termite drops mud around the pile, and pillars emerge.

This example illustrates that successful collective action through stigmergy is demonstrated by the clustering of activities along two dimensions: space and time. Spatially, the existence of an initial pile stimulates builders to accumulate more pellets in the same spot through a positive feedback loop as the accumulation of pheromones reinforces the attractiveness of the pile. The accumulation of pellets also clusters in time because the pheromone scent is strong and exciting at first but decays over time. Empirical studies identify strong spatial-temporal clustering behavior in ants' stigmergy [25, 64].

We propose that the same collective excitation process also manifests in the stigmergy in open collaboration. Contributions can excite more contributions by enacting a positive feedback loop between contribution and attention. Knowledge contributors to open collaboration communities are often consumers of that knowledge; they are attracted to the community by high-quality knowledge created by others [44]. Cumulative contributions and content growth thus not only improve the quality of the knowledge artifact, but also excite contributions from members of the communities [1, 74]. At the same time, the contributions reveal more information about the task and task dependencies, which, by themselves, are stimulating [14, 40]. For example, when an editor adds a sentence such as "*Apple designs iPhone*" to the company's Wikipedia page, the edit itself excites more edits that modify and extend the statement.

The exciting effect of activities may occur when an action is performed in the shared knowledge artifact, regardless of whether participants intentionally leave traces for coordination [1]. Participants may be attracted because somebody left a message saying, "more revision is needed here," or simply because a recent change was made by someone who made no attempt to exhort others. Moreover, the collective excitation process may be amplified by modern information system features such as notifications, awareness displays, easy version comparisons, and hyperlinked networks [22, 56, 57, 74]. For example, changes made to a Wikipedia article can attract the attention of other editors who receive notifications about those changes or come across the article following article hyperlinks. Once editors are attracted to the page, they are further influenced by others' behavior by observing each other's edits [22, 56]. Therefore, with the excitation process at work, participants pay attention to each other's activity, and the rate of activity functions as a signal that coordinates community members' actions.

Stigmergy, User Participation, and Information Quality

We argue that the effects of stigmergy on knowledge production outcomes can be best understood through the examination of its relationship with two important knowledge production outcomes, namely user participation [1, 44, 71, 74] and information quality of the coproduced knowledge [5, 6, 43, 46, 66]. The processes of collective modification and collective excitation serve as the underlying processes driving this relationship. Because of these two processes, more participants are attracted to the knowledge artifact by responding to modification activities clustered in time and space. In this way, participants coordinate by being aware of what is happening (the collective modifications) and when and where it is happening (the collective excitations).

Since user activity serves to excite other users' actions in stigmergy, the collective excitation process can generate a positive influence on user participation by encouraging repeated interactions and reinforcing members' tendency to contribute. In the context of Wikipedia, such interactions are enabled through the articles that editors contribute to. Editors follow digital traces made by other contributors and allocate their effort to extend or modify one another's work. Observing someone editing the same article at about the same time can entice the editors to further contribute, because it indicates their work has been noticed and provides an immediate reinforcement for their contributions [8, 10, 63]. Newcomers are also more willing to contribute when receiving feedback from others, regardless of the content and sentiment of the message [73]. Thus, frequent interactions create a positive reinforcement loop that motivates editors to spend greater effort and contribute more [56]. Therefore, our first hypothesis is:

Hypothesis 1 (H1): Stigmergy is positively associated with an increase in user participation.

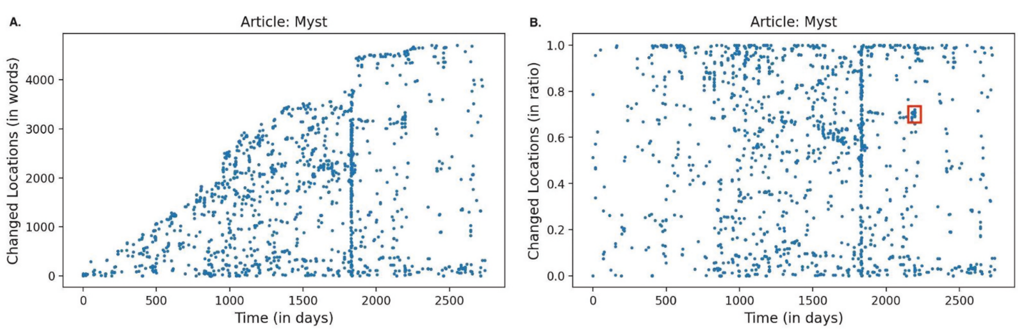
The collective excitation process may also improve the quality of knowledge production by facilitating collaborative learning among participants. Collaborative learning occurs when a group of people gain new knowledge in a common task, capitalizing on each other's resources and skills [24]. In the context of Wikipedia, editors learn and build shared knowledge by interacting with each other via co-editing the knowledge artifact [5, 40, 44]. It is through such interactions that editors exercise, verify, solidify, and improve their knowledge. On the one hand, previous edits carry task-relevant information and provide good examples of how to engage with the artifact [40]. To illustrate, an editor may learn the proper syntax to add a citation to a Wikipedia article by observing previous edits that add citations to the page. On the other hand, later edits facilitate the learning of previous contributors by providing general feedback about the quality of their edits [54]. A later edit rejecting or correcting their edits reveals signs of low-quality contributions and ways to improve them; an edit accepting or extending their edits shows signs of high-quality contributions and encourages further high-quality edits. This collective learning process is further accelerated by the excitation process that encourages more frequent and rapid feedback. For example, software developers may find it frustratingly slow to learn if they can only receive feedback from others days after deploying their code [48]. By contrast, in the presence of stigmergy, edits spur more edits in a short time span and later edits provide immediate feedback to previous edits, thus facilitating fast learning. As a result, editors can exchange information more frequently and learn more efficiently by constantly observing and interpreting changes made by others [21, 22]. This increased collective learning is likely to increase the quality of the artifact. As a result, we hypothesize:

Hypothesis 2 (H2): Stigmergy is positively associated with an improvement in the quality of a peer-produced artifact.

Measuring Stigmergy

We adapt Moran' I to gauge the level of spatial-temporal clustering in Wikipedia edits to measure stigmergy. Moran's I is a correlation coefficient that measures how closely points are clustered together in a two-dimensional space [55]. In contrast to previous attempts that measured stigmergy by distinguishing communicative and noncommunicative group activities [20, 59], our approach leverages the insight that the excitation process of stigmergy leads to a high dependence of group activities within the knowledge artifact in space and time. When members coordinate in a stigmergic fashion, they pay attention to each other's activities and are excited by each other's actions. Thus, previous contributions increase the probability of observing a later contribution around the same location in space soon after, leading to strong dependencies of user contributions in the spatial-temporal space or clustered activities [25,56]. Conversely, when there is a low level of stigmergy, contributions do not excite each other, resulting in evenly or randomly distributed activities in both space and time.

To capture patterns indicating stigmergy in a Wikipedia article using Moran's I, we first map a Wikipedia article's editing activities as a point pattern along the spatial and temporal dimensions using edits' timestamps and changed locations. For example, Figure 1A shows the point pattern of the Myst article, in which each dot represents one edit. The x-axis, or the temporal axis, represents when each edit is made in the number of days since the page's creation. The y-axis, or the spatial axis, records where the edit is made in the article, determined by comparing the text of the current revision with its previous revision. For example, if a revision states, "*Apple is an IT company*," and its succeeding revision states "*Apple is an American IT company*," the changed location for the later revision will be 4 as it inserts a new word in position 4. Since an article's length increases over time as the article



Figures 1. A and B. Spatial-temporal point patterns of the Myst Article.

develops, we remove the influence of article length by converting the raw changed locations into a ratio of the raw location divided by the total article length as shown in Figure 1B. When stigmergy is present, we would expect to see edits cluster together temporally and spatially in the 2D spatial-temporal point pattern, as illustrated by the red box in Figure 1B.

To calculate Moran's I based on the spatial-temporal point pattern of editing activities, we first separate the points into smaller spatial-temporal units by taking bins on the temporal axis (x-axis) and spatial axis (y-axis). We set the number of bins on the temporal axis equal to the article age in days and the number of bins on the spatial axis to 10. This way, each spatial-temporal unit i represents one day temporally and 10 percent of the total article length spatially. We use x_i to denote the number of edits occurring in the unit. Consistent with other uses of Moran's I [31], we define a weight matrix W_{ij} in which the weight w_{ij} equals 1 if two units x_i and x_j are neighbors (adjacent in the spatial-temporal space); otherwise 0. Then, Moran's I is defined as:

$$M = \frac{N}{W}\frac{\sum_i\sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

where N is the number of units indexed by i and j, and W is the sum of all w_{ij}.

Moran's I has a value in the bounded range [-1,1], and its magnitude indicates the point pattern's degree of clustering. An activity pattern with spatial-temporal randomness will yield a neutral $M = 0$, whereas $M = 1$ is the most clustered pattern, and $M = -1$ is an evenly dispersed (rhythmic) pattern. Higher values correspond to higher dependency between edits in both time and space, thus indicating higher stigmergy. The articles in our sample have a Moran's I ranging from -0.49 to 0.736 with a mean value of 0.107, which is significantly greater than zero.[1] The majority of the articles (98.42 percent) exhibit a Moran's I value greater than 0, indicating the prevalence of spatial-temporal clustering patterns on Wikipedia.

Figure 2 compares the editing timestamps of two pairs of Wikipedia articles with a similar number of edits but different M, where each dot represents one edit in the article history. Figures 2A and 2C show two articles with higher M. The dark clustered points are the results of intensive edits in approximate positions in a short time span. In Figures 2B and 2D, the articles have lower M, which have lower levels of clustering in time and space, that is, editing patterns are closer to random.

We take a two-pronged approach to show that M can capture the clustering of editing activities on Wikipedia. First, we analyze the edits in the Myst article using a qualitative approach and illustrate that stigmergy is reflected by clustered activity in space and time. Then, in Online Supplemental Appendix B, we perform a simulation analysis which shows that in the absence of stigmergic activities along either or both the space and time dimensions, M would have been lower. These analyses suggest the validity of the measure to reflect the level of stigmergy through editing patterns at the article level. However, it should be noted that the measure does not directly gauge the stigmergic process. A more direct measure would require somehow instrumenting the editors to know exactly what they viewed before editing. This is difficult because, for privacy reasons, Wikimedia makes publicly visible who has edited content but not who has viewed content.

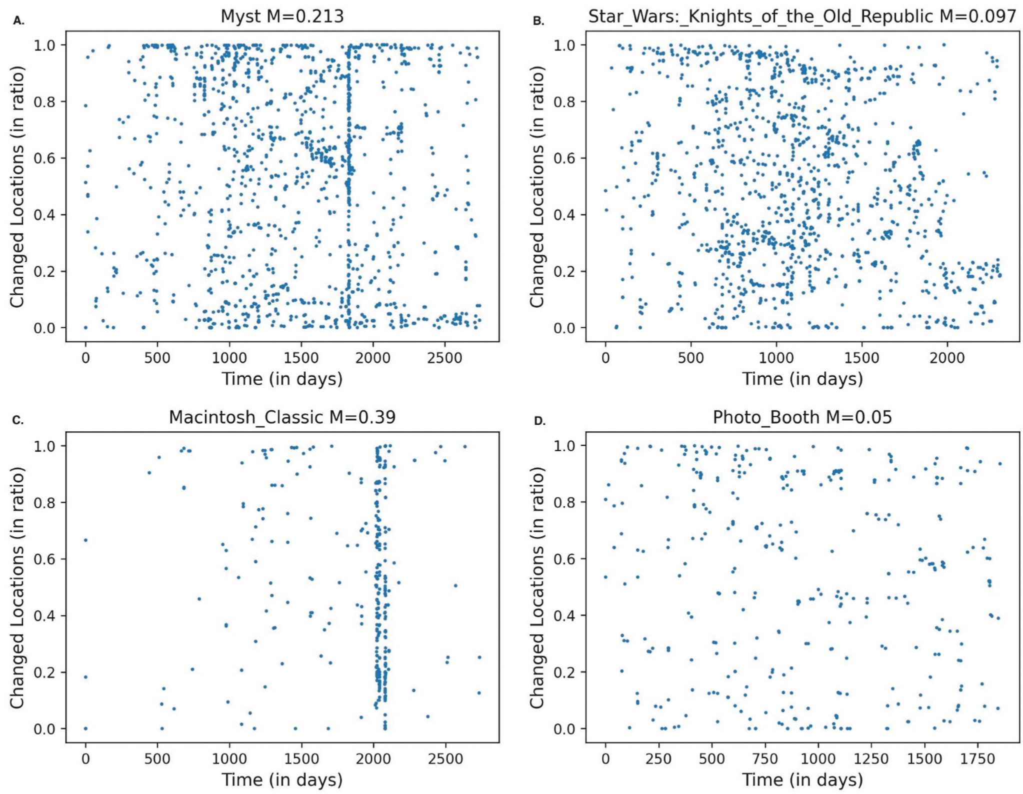


Figure 2. A, B, C, D. Editing point patterns and Moran's I of Wikipedia articles. *Notes*: * Each dot represents one edit in the article history. A. Myst page with a high Moran's I M = 0.213, Total Edits = 1496; B. Star Wars: Knights of the Old Republic page with a moderate Moran's I M = 0.097, Total Edits = 1438; C. Macintosh Classic page with a high Moran's I M = 0.39, Total Edits = 376; D. Photo Booth page with a low Moran's I M = 0.048, Total Edits = 377.

Extended Example: The Myst Article

To further illustrate why temporal and spatial clustering can be used to measure stigmergy, we take a closer look at what happens in a highly stigmergic cluster in the Myst article. Figure 3 summarizes a series of stigmergic edits among five editors in the cluster on the Myst article in the red box in Figure 1B. The editing sequence started when an anonymous editor wrote "*On the 2nd of May, 2009, Myst was released as a 700mb download for the iPhone and iPod Touch via the App Store*." New issues emerged because of the newly added sentence. Editors noted the new sentence overlaps with the existing content in the article. They also noted that the claim lacks an external reference. Also, they found out that the sentence was inaccurate because the size of the software was slightly over 700 MB and the release date of the software was earlier than claimed in the edit. Moreover, the format of the date (*On the 2nd of May*) was not consistent with the rest of the article. In the next 34 hours, four different editors came to resolve these issues. As shown in Figure 3, these editors made their contributions on top of each other in a stigmergic fashion by directly modifying previous edits. Moreover, the edits excited other edits, as the newly added content introduced new issues that needed to be resolved. For example, when another anonymous editor added another date (*April 22*) as the software's release date, the new information

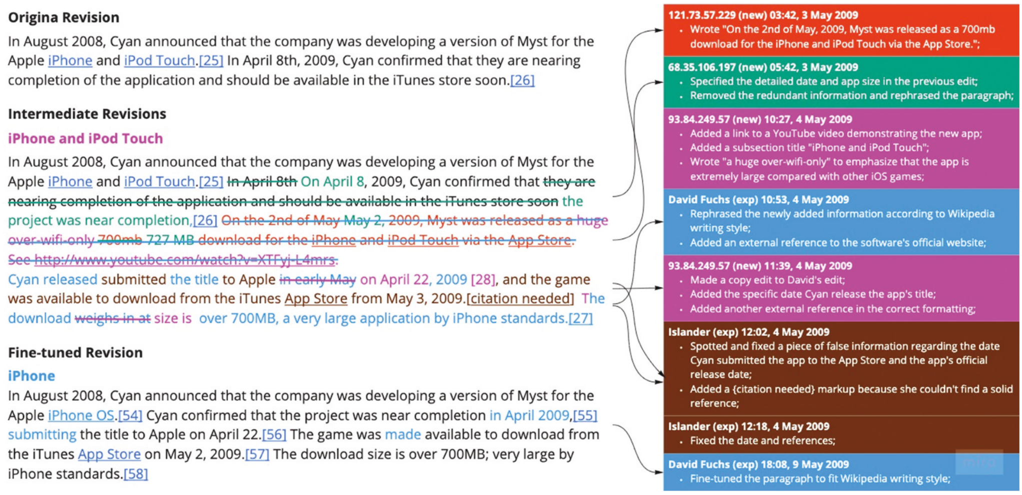


Figure 3. Interpreting the clustering editing activities.

created confusion about the actual release date of the software. This issue was soon addressed by Islander, an editor who found that both dates were valid because one was the time the developer released the title (*April 22*) and the other was the time the game was made available in the iOS store (*May 2*). Islander clarified this information, but her contribution created another issue as she could not find an appropriate reference. As a result, she put a {{citation needed}} tag as a placeholder, and possibly to request assistance from others until the issue was resolved.

The aforementioned example also illustrates the two processes through which stigmergy may impact knowledge production: the collective modification process that allows editors to build on each other's work, and the collective excitation process that motivates editors to remain engaged in consecutive modifications. Both processes can be seen in the interactions of an experienced editor (David Fuchs) and a new anonymous editor (93.84.249.57). When the new editor first came to the page, she found that the newly added information lacked sufficient references. Thus she added a link to a YouTube video in the format of "*See* https://. . . ." The edit created a new issue as both the source and the format of the reference did not comply with the Wikipedia writing style. This issue was soon addressed by an experienced editor, David Fuchs, who was one of the page's leading contributors and likely received notifications about the page's content change. David fixed the problem within 30 minutes by adding a link to the software's official website with the correct formatting, and his copy-edit served as feedback for the new editor and stimulated the editor's learning and contribution. Within 40 minutes, 93.84.249.57 came back and further copy-edited David's revision. She also added another external reference, but this time, in the correct format.

Method

Data

We conduct our empirical analysis using the Apple Inc. WikiProject, an active subcommunity focusing on developing information related to Apple Inc. and its products. A WikiProject relies on the collaborative effort of a group of editors who develop, maintain, and organize articles associated with a specific topic area. At the time of writing, the Apple Inc. WikiProject consisted of 3,260 articles written by 366,294 unique contributors. We gather the full-text history of 1,158,279 revisions of 3,260 articles in the Apple Inc. WikiProject from 2001 until October 2017 using publicly provided APIs. For each revision of an article, we record the editor's identity, the text of the edit, the editor's description of the edit, and the time of the edit. We also gather the daily page view history of all the articles starting from July 2015.

We conduct the following data cleaning procedures for our sample. Bots are widely used in Wikipedia to perform routine tasks such as correcting typos, checking for copyright violations, and countering vandalism [72]. To address the influence of such agents, we first remove all bot edits (approximately 4.42 percent of all edits) by matching contributors' usernames with a list of registered bots in English Wikipedia.[2] Second, extremely short articles have limited content or edits for us to analyze their stigmergic processes. Therefore, we exclude articles that have fewer than 30 edits, leaving 2,275 articles in our sample.

Measures

Dependent Variables

User Participation. Following prior studies, we measure user participation using the average duration of time spent by users on developing the knowledge artifact. Specifically, we segment user participation into edit sessions [5, 33], and compute the total amount of time a user devotes to the writing of an article by summing up the time duration of their individual edit sessions. Each edit session is defined as a sequence of edits made to an article by an editor in which time intervals between every two consecutive edits are less than one hour. We choose one hour as the cutoff time following [33], who demonstrate that editors generally make consecutive edits within one hour by analyzing one million randomly sampled revisions. Then, user participation is calculated as the total participation time divided by the total number of distinct editors on that article. We use average rather than total user participation time to adjust for the fact that total user participation time is likely to be higher when more users participate.

Information Quality. To assess information quality, we took advantage of Wikipedia's article assessment project, which has evaluated over 5,971,036 articles by peer review in a consistent and uniform manner.[3] The peer-rated article quality is widely used in related studies and is a good proxy for information quality [43, 46]. The peer reviewers assessed all articles in Wikipedia on a six-point scale (from lowest to highest quality: Stub, Start, C, B, Good, Featured). To avoid quality changes caused by further editing after the peer reviewers' assessments, we remove all the revisions that occurred after the date the article was evaluated. This leaves a total of 656,738 revisions. We code the peer-rated article quality using six-point scale integers, ranging from 1 (Stub) to 6 (Featured). Table A1 in Online Supplemental Appendix A summarizes the number of articles in each quality level in our sample.

Independent Variables

Stigmergy. We apply the Moran's I of an article to capture the editing pattern indicating stigmergy. As explained previously, we convert edits within an article into a point pattern by mapping edits into a 2D spatial-temporal space and computing the Moran's I for each article.

Control Variables

We control for the following variables: *article age, article popularity, number of distinct contributors, anonymity, revert ratio*, and *article topics*. The longer an article exists and the more popular the article is, the more editors may contribute to it, which in turn leads to higher information quality [43]. To control for the two variables, we first add *article age*, which is computed as the number of months since the article's creation. We also use the article's average daily pageviews as a proxy for an article's *topic popularity*.

Moreover, the number of distinct contributors and their membership status were shown to be associated with information quality [5, 58]. We control for their effects using the *number of distinct contributors* and *article anonymity* (the percentage of unregistered editors). In Wikipedia, people can make edits with or without logging into the Wikipedia system. Registered editors tend to be more familiar with the community norms and rules and are more likely to make edits consistent with the mainstream perspective. Unregistered editors tend to contribute novel ideas that expand article exploration [5].

We further control for the effect of task conflicts by adding the *revert ratio*, the percentage of reverts over all edits. Task conflict is a common phenomenon in open collaboration contexts as people with different backgrounds and viewpoints come together in developing the same knowledge artifact [47]. In Wikipedia, an example of excessive conflict is an edit war, in which editors revert each other's previous edits, leading to meaningless or even harmful edits. Empirical evidence suggests that such excessive task conflicts impede successful collaboration and undermine information quality [6].

Finally, we control different *article topics* using their categories. We first retrieve all categories and their subcategories following the category hierarchy of the Apple Inc. WikiProject. For each article, we define a set of dummy variables that indicate whether the article belongs to a topic category. For instance, if an article belongs to the *Apple Inc. hardware* category or its subcategories, the associated variable will be one, otherwise zero. We control for five categories that are most frequently used within the project (*hardware, software, people, products*, and *platforms*) and aggregate articles that do not belong to any of them into *others*. Table 1 presents the summary statistics and correlations between the key variables.

Data Analysis

To test our hypotheses on the relationship between stigmergy and knowledge production outcomes, we use the following regression specification to analyze the data at the article level. The regression analysis takes into account the nature of the dependent variable. When the dependent variable Y is *user participation*, we estimate a log-level Ordinary Least Squares (OLS) regression because the dependent variable is continuous and highly skewed. When the dependent variable Y is *information quality*, due to its ordered and categorical nature, we estimate an Ordered Logistic regression following previous research [43]. To

TABLE 1. Summary statistics and correlations.

Variables	Mean	STD	1	2	3	4	5	6	7	8	9
1. Quality Rating	2.08	0.83	1.00								
2. User Participation	13.85	11.58	0.13	1.00							
3. Stigmergy	0.11	0.10	0.32	0.28	1.00						
4. Age (in month)	65.20	34.66	0.25	-0.21	-0.23	1.00					
5. Topic Popularity	219.44	820.59	0.28	0.04	0.26	0.07	1.00				
6. Distinct Contributors	151.70	314.50	0.40	-0.04	0.19	0.26	0.61	1.00			
7. Anonymity (percent)	0.32	0.14	0.06	-0.27	-0.02	0.10	0.07	0.27	1.00		
8. Revert Ratio (percent)	0.02	0.03	0.14	0.01	0.07	0.07	0.07	0.16	0.06	1.00	
9. Average Stigmergy	0.10	0.03	0.16	0.06	0.18	-0.05	0.13	0.11	0.08	0.03	1.00
Observations	2,275										

ensure the variable coefficients are comparable, we standardize all independent variables. We enter the variables in a stepwise fashion, starting with the control variables, followed by the main effect of *stigmergy*.

$$Y_i = \beta_0 + \beta_1 Stigmergy_i + \beta_2 Age_i + \beta_3 TopicPopularity_i + \beta_4 DistinctContributors_i + \beta_5 Anonymity_i + \beta_6 RevertRatio_i + \sum_{j=1}^{5} \gamma_j ArticleTopics_j + \varepsilon_i$$

We find no significant concern of multicollinearity in our model by checking the variance inflation factors (VIFs). We find that the VIFs associated with variables are all less than 3.35, which is well below a critical value of 10 [12].

Results

Table 2 presents the OLS and Ordered Logistic Regression results. The two left columns illustrate the results of *user participation*. Model 1 only contains the control variables. In Model 2, we add the variable of interest, *stigmergy*, into the regression and find that the R^2 nearly doubled, showing the predictive power of *stigmergy* on user participation time. Supporting H1, the effect of *stigmergy* is positive and significant. Specifically, one standard deviation increase in *stigmergy* (0.1) is associated with a 16.1 percent increase in *average user participation time* (2.23 minutes per editor). Furthermore, since all independent variables are standardized, we could directly compare the effect size of *stigmergy* and other predictors of user participation. We find that stigmergy has a larger effect size than all the other controls, suggesting it is the greatest contributor to sustained user participation.

The two right columns summarize the results of *information quality*. The values provided are odds ratios. Model 3 contains only the control variables. Supporting H2, Model 4 shows that *stigmergy* has a significant positive effect ($oddsratio = 1.679, p < 0.01$) on *information quality*. One standard deviation increase in *stigmergy* (0.1) is associated with a 1.679 increase in the odds of the article being at a higher quality level. Furthermore, the effect size of stigmergy exceeds several control variables including *article age, article anonymity*, and *revert ratio*.

TABLE 2 Regression of user participation and information quality on stigmergy.

	Participation		Quality	
Variables	(1)	(2)	(3)	(4)
Stigmergy		0.161***		1.679***
		(0.011)		(0.099)
Age (in months)	-0.077***	-0.020***	1.174***	1.435***
	(0.007)	(0.008)	(0.062)	(0.084)
Topic Popularity	0.058***	0.031**	1.828***	1.702***
	(0.013)	(0.013)	(0.128)	(0.124)
Distinct Contributors	-0.032*	-0.074***	3.180***	2.875***
	(0.017)	(0.015)	(0.277)	(0.251)
Anonymity (percent)	-0.114***	-0.082***	0.543***	0.593***
	(0.012)	(0.010)	(0.034)	(0.035)
Reverts Ratio (percent)	0.021***	0.018***	1.061	1.059
	(0.007)	(0.006)	(0.054)	(0.056)
Constant	2.656***	2.643***		
	(0.021)	(0.019)		
Article Topics	Yes	Yes	Yes	Yes
Observations	2,275	2,275	2,275	2,275
R-squared	0.202	0.390	0.196	0.217

Notes: * The sample includes 2,275 articles from February 2001 to October 2017.
All variables are standardized. ***p < 0.01. **p < 0.05. *p < 0.1.

Robustness Checks

We conduct a number of robustness checks. First, in the main regression of *user participation*, we use a 1-hour cutoff for the user edit session to estimate our main dependent variable. To ensure the robustness of our analysis, we use 30-minute and 3-hour as two alternative cutoffs of user edit sessions and re-estimate the dependent variable. We find that using different cutoffs does not change the direction or significance level of stigmergy.

Second, we address the possibility that the observed high spatial-temporal dependency among edits is *solely* driven by external shocks instead of a natural synergy among editors. By definition, stigmergy triggered by external shocks is still stigmergy as long as the edits in the articles stimulate additional actions from editors. However, we may also observe high levels of spatial and temporal clustering in activities as a result of many editors flooding to edit the article when external events occur [51, 74]. Even though editors who converge on Wikipedia due to external shocks may inevitably need to coordinate with others in a stigmergic fashion, for the sake of generalizability, it is important to ensure that the effect of stigmergy is robust in the absence of external influences. Accordingly, we rerun our analysis using an alternative sample, Ancient Philosophy WikiProject, which is unlikely to be subject to external shocks. The Ancient Philosophy project has 1,325 articles that focus on developing information related to the philosophical schools and philosophers in early human history. It is less influenced by external shocks since few new facts would have emerged in short time windows. We gather the articles' full text history, daily page views, and peer-rated quality. We perform the same data screening procedures as we did for the Apple Inc. project, leaving 671 pages that satisfy the requirement of our analysis.[4] Then we parse all related variables and re-estimate our main regressions.

Table 3 shows the main regression results estimated using the Ancient Philosophy WikiProject. We see that the effects of *stigmergy* on *user participation* and *information quality* are still significant and positive. At the same time, using the alternative sample, the

TABLE 3. Results from alternative sample (ancient philosophy wikiproject).

	Participation		Quality	
Variables	(1)	(2)	(3)	(4)
Stigmergy		0.255***		1.714***
		(0.020)		(0.168)
Age (in month)	-0.042**	-0.032**	1.223**	1.260**
	(0.018)	(0.015)	(0.114)	(0.119)
Topic Popularity	0.214***	0.130***	1.543***	1.307*
	(0.043)	(0.035)	(0.248)	(0.207)
Distinct Contributors	-0.120**	-0.140***	2.717***	2.716***
	(0.060)	(0.050)	(0.525)	(0.514)
Anonymity (percent)	-0.311***	-0.180***	0.375***	0.476***
	(0.027)	(0.021)	(0.045)	(0.060)
Reverts Ratio (percent)	0.038*	0.017	1.321**	1.272**
	(0.022)	(0.017)	(0.144)	(0.140)
Constant	2.747***	2.748***		
	(0.016)	(0.013)		
Observations	671	671	671	671
R-squared	0.316	0.550	0.144	0.167

* The sample includes 671 articles in the Ancient Philosophy Project from February 2001 to March 2022. All variables are standardized. ***p < 0.01. **p < 0.05. *p < 0.1.

effect of stigmergy on user participation is still robust to different cutoffs (30 mins and 3 hours). Notably, we see that the effect sizes of stigmergy on both user participation and information quality increase when using the alternative sample. Specifically, the effect of stigmergy on user participation increases by 58 percent (from 0.161 to 0.255) and the effect on quality increases by 2 percent (from 1.679 to 1.714). This suggests that there is a stronger association between stigmergy and collaboration outcomes for editors working on articles less influenced by external shocks and surges of public attention.

Second, we use a fixed effect panel regression to further alleviate endogeneity concerns. Instead of assigning a single measure of stigmergy to an entire article, we measure stigmergy for each article-year combination. With such a panel dataset, we can examine whether article quality improves as its stigmergy changes over the entire life cycle of the article. In addition, by introducing article fixed effects into the model, we can account for article-specific unobserved confounders such as different initial article developmental stages.

Specifically, we construct a 16-year panel of yearly article-level data. Similar to the cleaning procedures executed at the article level, we remove all article-year observations whose total edits are less than 30. After this procedure, our panel contains 3,592 article-year observations of 844 articles. An article-year observation includes the average user participation time in that year, the article quality change (ΔQ), the degree of stigmergy in that year, and all other controls. As most articles have only been evaluated once or twice by Wikipedia peers, we use the Objective Revision Evaluation Service (ORES)[5] predicted article quality to assess the change in quality rating for an article in an observation year. The ORES service is an automated tool developed by the Wikimedia Foundation to estimate article quality. It uses machine learning models trained on peer-rated articles to assign quality ratings to article revisions. The ORES predicted quality has a rating scale that resembles the peer-rated quality scale (from lowest to highest quality: Stub, Start, C, B, Good, Featured) and has been used in previous research to assess article quality [5]. Thus, the quality change (ΔQ) is measured as the difference between the average quality of the first 5 revisions (Q_{before}) and the last 5 revisions (Q_{after}) within the article-year observation ($\Delta Q = Q_{after} - Q_{before}$). Then,

TABLE 4. Panel regression of user participation and information quality on stigmergy.

	Participation		Quality	
Variables	(1)	(2)	(3)	(4)
Stigmergy		0.167***		0.243***
		(0.009)		(0.027)
Age (in month)	-0.296*	-0.170	1.525***	1.707***
	(0.160)	(0.150)	(0.581)	(0.549)
Topic Popularity	0.202	0.150	-0.037	-0.112
	(0.191)	(0.172)	(0.489)	(0.431)
Distinct Contributors	0.049***	-0.029**	0.167***	0.055
	(0.012)	(0.011)	(0.046)	(0.046)
Anonymity (percent)	-0.151***	-0.115***	-0.156***	-0.103***
	(0.010)	(0.008)	(0.028)	(0.027)
Revert Ratio (percent)	-0.022***	-0.016**	-0.104***	-0.095***
	(0.008)	(0.006)	(0.024)	(0.023)
Constant	2.564***	2.580***	0.243***	0.266***
	(0.032)	(0.029)	(0.085)	(0.076)
Article Fixed Effect	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes
Observations	3,592	3,592	3,592	3,592
R-squared	0.598	0.710	0.286	0.309

* The sample includes 3,592 article-year observations in Apple Inc. Project from February 2001 to October 2017.
All variables are standardized. ***p < 0.01. **p < 0.05. *p < 0.1.

we run a two-way fixed effect model specified as follows (Y is either *UserParticipation* or ΔQ):

$$Y_{i,t} = \beta_0 + \beta_1 Stigmergy_{i,t} + \beta_2 Age_{i,t} + \beta_3 TopicPopularity_{i,t} + \beta_4 DistinctContributors_{i,t} + \beta_5 Anonymity_{i,t} + \beta_6 RevertRatio_{i,t} + ArticleFE_i + YearFE_t + \varepsilon_{i,t}$$

Table 4 shows the estimation results of the panel regressions. We find that stigmergy has remained positively and significantly associated with both user participation and information quality. The observed relationships are thus robust to unobservable article-specific characteristics. The results also demonstrate that the relationship manifests over time, rather than just cross-sectionally.

Two other robustness checks are presented in the Online Supplemental Appendices C and D. In the Online Supplemental Appendix C, we conduct a comparative analysis by constructing alternative Moran's I measures that consider clustered activities only in the temporal but not the spatial dimension, and only in the spatial but not temporal dimension. The analysis shows that the association with article quality is weaker when clustering in either dimension is low. In the Online Supplemental Appendix D, we present an instrumental variable analysis to further alleviate endogeneity concerns.

Understanding the Processes

Given the positive associations between stigmergy and knowledge production outcomes, what are some ways to amplify it? To test the impact of changes to the collective modification process we first explore the use of *inline cleanup tags*.[6] These tags are left in articles to indicate if it needs further work, such as improving clarity (e.g., {{expand acronym}}) or verifiability (e.g., {{citation needed}}). We note that the use of such tags is on the border between explicit coordination and stigmergy—a task is suggested, as a trace for others to discover and act on, whereas the specific person in charge is not specified as it would in explicit coordination. We first collect all task-relevant templates that are used in English Wikipedia and compute their usage frequency in article revisions using regular expressions. Surprisingly, the majority of articles (82.3 percent) in the sample did not use any of these templates in all their revisions. We then re-estimate the effect of stigmergy in articles that do not contain any task-relevant templates for coordination versus those that do. The results can be seen in Table E1 in the Online Supplemental Appendix E. We find that the stigmergy's effects on user participation and information quality are still positive and significant without the templates (Online Supplemental Table E1, columns 1 and 3). In fact, the effect size is slightly larger than the articles that used the templates (Online Supplemental Table E1, columns 2 and 4). While this result should not be interpreted as indicating these cues are redundant, it at least suggests that the effect of stigmergy is equally important for ordinary edits without explicit tasks suggested.

We turn to the second approach of highlighting changes to amplify the collective excitation processes. The excitation process relies on users' timely awareness and can be amplified by modern information system features such as notifications and change alerts/comparisons. We test the effectiveness of a change to the excitation process using a pre-post study based on the adoption of email notifications. In April 2012, the English Wikipedia introduced its email notification feature which enables editors to receive email notifications whenever changes were made to articles on their personal watchlist, which contains pages that editors would like to "keep track of, and react to, what's happening to pages they have created or are otherwise interested in."[7] This feature changed the pattern of excitation because editors would be notified and attracted to recent edits in an article even without logging in to Wikipedia.

We use a binary variable *Notification* to indicate such feature changes: this variable equals 1 when the observational period is at or later than 2012 and is 0 for any period before 2012. We then test the interaction effect of *Notification* on *Stigmergy* using panel regression. We find that the interaction terms were positive and significant, indicating that the effects of stigmergy on both user participation and information quality are stronger after the deployment of the email notification system. In other words, the collective excitation process can be amplified by fostering a collective spatial-temporal awareness; one way to do this is through notifications that specify events and their locations. Online Supplemental Appendix E Table E2 presents the regression results.

Discussion

Open collaboration communities such as Wikipedia have fundamentally changed how knowledge is created, distributed, and consumed. Compared with conventional organizations, these communities produce knowledge differently because the majority of contributors don't work simultaneously and rarely communicate with each other. With limited explicit coordination [28, 50], a persistent research question is how members in these communities coordinate to produce coherent, high-quality knowledge. Although prior work has acknowledged the role of the shared knowledge artifact in open collaboration [4, 5, 40, 44], the processes and consequences of stigmergy are not fully understood. In this work, we extend the existing literature by elaborating the concept of stigmergy from biology; we articulate two intertwined processes involved in stigmergy and investigate its consequences on open collaboration outcomes.

We argue that stigmergy in open collaboration is not only a collective modification process that updates artifacts, but also a collective excitation process that stimulates contributions from the community. As such an excitation process manifests in the spatial-temporal clustering of group activities, we develop a measure to gauge stigmergy and investigate its association with knowledge production outcomes. Our empirical analysis shows that successful online coordination exhibits stigmergy, and that concentrated edits in a short time span can drive greater user participation and higher information quality.

Theoretical Contributions

This study contributes to our knowledge of open collaboration by clarifying the conceptualization of stigmergy. Despite the increasing recognition of artifact-centric coordination [4, 5, 40, 50] and calls for further investigations [42, 44], stigmergy has been proposed relatively recently as an explanation of how such coordination happens in open collaboration [9, 19]. The relationship between stigmergy and open collaboration outcomes is also not well-understood due to the early stage of theories and measurement methods.

Our work helps explain artifact-centric coordination by distinguishing two processes that constitute stigmergy: collective modification and collective excitation. Stigmergy is not just about collective modification via a shared knowledge artifact as proposed in previous literature [9, 16, 19], but also involves a collective excitation process in which recent changes in a particular location motivate community members to make more changes. Our analysis to probe the process suggests that stigmergy is likely to occur regardless of the contributors' intentions: although in many situations participants may leave explicit traces, such as using task-relevant templates in Wikipedia to exhort others' contributions, changes without direct cues can also trigger collective excitation. Moreover, system notifications alerting interested parties of recent changes may further promote awareness and attract contributions.

We develop a new measure of stigmergy by analyzing the spatial and temporal characteristics of collaborative editing patterns. The collective excitation process in stigmergy is manifested through temporal and spatial clustering of community member activities. Crowston and Rezgui [20, 59] have proposed a different measure of stigmergy. The measure is designed to distinguish between coordination that uses explicit communication and coordination that doesn't. That is, it focuses on the collective modification process. By contrast, our proposed measure integrates the collective modification and excitation processes. The measure is relatively simple to apply and captures simultaneously time-based and space-based clustering of human activity. It provides ways of predicting quality, which can be useful for designing system features that increase the quality of online knowledge production. It can also be applied by scholars to advance research on the antecedents and consequences of stigmergy or artifact-centric coordination in general on digital platforms.

Our study found a positive association between stigmergy and knowledge production outcomes. Our empirical findings show the potential benefits of stigmergy, in which members coordinate by changing artifacts and reacting to the changes made by others. One related concept in the management literature is the feedback loop between knowledge consumption and knowledge production. Improvement in content quality caused by incremental knowledge production can attract more consumers of that knowledge and thus more potential contributors [1, 44, 74]. By providing quantitative evidence that stigmergy is associated with higher user participation, this study shows that when people coordinate in a stigmergic fashion, the new activity itself may attract more attention and stimulate more edits. Thus, stigmergy can not only serve to coordinate the modification of artifacts but also can excite participation. In this way, stigmergy can help open collaboration communities overcome problems related to fluid membership, geographical and temporal dispersion, and limited explicit communication to achieve better outcomes.

For the field of information systems, there are many ways this work might be used to develop a better understanding of coordination in general, and coordination as specifically applied to online communities. Although the current study investigates stigmergy in the open collaboration context, stigmergy can be a general way for people to coordinate when explicit coordination methods are limited. At the very basic level, many information systems can be seen through a stigmergic lens. Examples may include software change management systems, enterprise architecture document repositories, version control platforms and code repository systems like GitHub, collaborative editing tools such as Google Docs, open innovation systems like Thingiverse, generative AI remix systems like Midjourney, and technical blogs, among many others. For instance, when timelines are used in social media, there is a signal being sent not just by the content published on the timeline but by the frequency of updates. When members of an online community focus on particular locations in artifacts, there is a signal that may attract others to work at those specific locations.

Practical Implications

This work has implications for practitioners, specifically the designers of open collaboration platforms. Interfaces might be constructed to enhance the collective excitation process. For example, users might be nudged toward particular artifacts, and locations in artifacts, at approximately the same time so that they might be more likely to excite each other to participate longer and improve the quality of the artifact. We noted that the introduction of email notifications to editors had a positive effect on the Wikipedia community: many more features along such lines might be developed. It is also possible that the collective modification process could be engineered to promote stigmergy. For example, an article with a problematic section might temporarily block all other sections from being edited to channel attention toward that section: diffused attention would become focused attention. Once the quality of the section improved, the block on other sections might be lifted.

Some tasks in open collaboration communities may sit incomplete waiting for the person with the right expertise to contribute, making progress frustratingly slow compared with progress in conventional organizations [40]. To remedy this, algorithms can be designed that dynamically compute the coordination requirements and suggest potential participants. In Wikipedia, algorithms based on static information are currently able to provide general recommendations about article-related tasks, such as adding more citations or improving article organization [18]. Our findings suggest that these algorithms can also leverage activity bursts [34, 60] and the emergence of local clusters [70] to strengthen stigmergy among editors. For instance, burst detection algorithms [49] and local Moran's I [70] might be used to bring articles to the attention of community leaders so they can either encourage attention or discourage conflicts.

Limitations

Our study has two limitations. First, we only establish an associational relationship between stigmergy and outcomes of open collaboration. Despite efforts to address endogeneity concerns using instrumental variables and panel regression, our study relies on observational data; the absence of a clear natural experiment limits the ability to draw causal conclusions. For example, our measure does not distinguish stigmergy arising organically from interactions among participants from stigmergy triggered by external shocks. Open environments such as Wikipedia can be affected by external shocks that alter the behaviors of contributors in different ways and potentially confound the observed effects. Future studies may use field or lab experiments to manipulate the degree of stigmergy in order to determine its causal effect and isolate factors contributing to the observed patterns of stigmergy.

Second, measuring stigmergy based on Moran's I only gauges a pattern consistent with stigmergy but does not directly capture the processes that led to its emergence. A more direct measure would be based on the degree to which editors pay attention to recent changes and act upon them. Additionally, although we find that the positive associations of stigmergy and knowledge production outcomes hold regardless of whether people use explicit traces signaling coordination needs, the proposed stigmergy measure does not account for these explicit traces. The role of such traces could be further explored in future studies, given AI tools have been developed to create such traces to facilitate online knowledge production [18]. Lastly, our measure uses the relative location of content changes in the knowledge artifact. This method allows for comparison across articles of different lengths, but may introduce bias if articles grow suddenly. For example, there will be a jump in the measure if a large amount of text is added all at once to the end of the article. Future studies may improve the measurement of stigmergy using website heatmap tools that track user attention (e.g., Hotjar, CrazyEgg) or by analyzing granular clickstream data. This would allow for a more direct assessment of stigmergy and increase its robustness to large edits.

Potential Extensions

There are several other ways to extend this study. One aspect of the measure of stigmergy that may be further developed is the analysis of spatial location, which is complex in the context of open collaboration. Such production is not characterized by Euclidean space, but by a semantic space that can be measured in many ways, including through links, word embeddings, and shared membership in a document. On a larger scale, articles themselves can be thought of as locations, and so articles have adjacency in semantic space to each other, which can be measured through link structures, folksonomies, and document embeddings. Many editors work across multiple articles at more or less the same time, and studying such editors may suggest ways of thinking about spatial behavior beyond the confines of individual articles.

Conclusion

Humans have the power of language, and so can explicitly coordinate. But there are many situations in which such coordination is impossible or inefficient. In such cases, coordination can be affected by making changes to shared artifacts that others will notice and act on. This artifact-centric coordination process is stigmergy, a term first coined in biology. Since humans are also biological creatures, it is sensible to ask if, like other biological creatures, humans are attuned to spatial-temporal changes in the environment, especially modifications of shared knowledge artifacts. This study shows they are. This work is a step toward better understanding stigmergy in humans, which in turn may help designers build better ways to allow large numbers of people to coordinate effectively while building collective knowledge.

Notes

1. A one sample t-test produces a t-statistic of 54.19 ($p < 0.01$).
2. https://en.wikipedia.org/wiki/Category:All_Wikipedia_bots
3. Wikipedia article content assessment: https://en.wikipedia.org/wiki/Wikipedia:Content_assessment
4. Many pages are excluded from this analysis due to their quality assessment date being shortly after the page's creation date, leaving the number of total edits less than 30.
5. The ORES tool; see https://www.mediawiki.org/wiki/ORES
6. We thank an anonymous reviewer for this suggestion.
7. See https://en.wikipedia.org/wiki/Help:Email_notification and https://en.wikipedia.org/wiki/Help:Watchlist

Acknowledgment

The authors would like to thank Deborah Gordon for her insights on stigmergy.

Funding

This material is based upon work supported by the National Science Foundation under grants 2128906, 2113906, 1909803, and 1717473.

Notes on contributors

Lei (Nico) Zheng (lzheng9@stevens.edu; corresponding author) is an Assistant Professor at the College of Management, University of Massachusetts Boston. She received her Ph.D. in Data Science from Stevens Institute of Technology. Dr. Zheng's research interests include online knowledge collaboration, the use of bots, human-machine collaboration, and online community governance. Her work has been published in the proceedings of *International Conference on Information Systems, Workshop on Information Technologies and Systems*, and in the *Proceedings of the ACM on Human-Computer Interaction*.

Feng Mai (fmai@stevens.edu) is an Associate Professor in the School of Business at Stevens Institute of Technology. He received his Ph.D. from the University of Cincinnati. His research interests include FinTech, natural language processing, and online communities. Dr. Mai's work has been published in *Journal of Management Information Systems, Review of Financial Studies, Journal of Operations Management, Marketing Science*, and other journals.

Bei Yan (byan7@stevens.edu) is an Assistant Professor in the School of Business at Stevens Institute of Technology. She received her Ph.D. degree from University of Southern California. Dr. Yan's research focuses on technology-supported collaboration and influence processes of groups, including small teams interacting with intelligent personal assistants to online communities applying bots. Her work has been published in the *Journal of Management Information Systems*.

Jeffrey V. Nickerson (jnickers@stevens.edu) is the Steven Shulman ’62 Chair for Business Leadership and Professor of Digital Innovation at Stevens Institute of Technology. His research focuses on how humans and machines work together on creative endeavors. He is currently an investigator in a National Science Foundation-funded project in the Future of Work program. His work has been published in the *Journal of Management Information Systems* and *MIS Quarterly*.

ORCID

Lei (Nico) Zheng http://orcid.org/0000-0002-1147-3773
Feng Mai http://orcid.org/0000-0001-6897-8935
Bei Yan http://orcid.org/0000-0001-9737-5308
Jeffrey V. Nickerson http://orcid.org/0000-0001-8590-3898

References

1. Aaltonen, A.; and Seiler, S. Cumulative growth in user-generated content production: Evidence from Wikipedia. *Management Science*, *62*, 7 (July 2016), 2054–2069.
2. Alnuaimi, O.A.; Robert, L.P.; and Maruping, L.M. Team size, dispersion, and social loafing in technology-supported teams: A perspective on the theory of moral disengagement. *Journal of Management Information Systems*, *27*, 1 (July 2010), 203–230.
3. Angwin, J.; and Fowler, G.A. Volunteers log off as Wikipedia ages. *Wall Street Journal*, *23*, (2009), A1.
4. Arazy, O.; Daxenberger, J.; Lifshitz-Assaf, H.; Nov, O.; and Gurevych, I. Turbulent stability of emergent roles: The dualistic nature of self-organizing knowledge coproduction. *Information Systems Research*, *27*, 4 (December 2016), 792–812.
5. Arazy, O.; Lindberg, A.; Rezaei, M.; and Samorani, M. The evolutionary trajectories of peer-produced artifacts: Group composition, the trajectories’ exploration, and the quality of artifacts. *MIS Quarterly*, *44*, 4 (December 2020), 2013–2053.
6. Arazy, O.; Nov, O.; Patterson, R.; and Yeo, L. Information quality in Wikipedia: The effects of group composition and task conflict. *Journal of Management Information Systems*, *27*, 4 (April 2011), 71–98.
7. Bailey, D.E.; Leonardi, P.M.; and Chong, J. Minding the gaps: Understanding technology interdependence and coordination in knowledge work. *Organization Science*, *21*, 3 (2010), 713–730.
8. Bhattacharyya, S.; Banerjee, S.; Bose, I.; and Kankanhalli, A. Temporal effects of repeated recognition and lack of recognition on online community contributions. *Journal of Management Information Systems*, *37*, 2 (April 2020), 536–562.
9. Bolici, F.; Howison, J.; and Crowston, K. Stigmergic coordination in FLOSS development teams: Integrating explicit and implicit mechanisms. *Cognitive Systems Research*, *38*, (June 2016), 14–22.
10. Burke, M.; Kraut, R.; and Marlow, C. Social capital on Facebook: Differentiating uses and users. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems - CHI ’11*. ACM Press, Vancouver, BC, Canada, 2011, p. 571.
11. Butler, B.S. Membership size, communication activity, and sustainability: a resource-based model of online social structures. *Information Systems Research*, *12*, 4 (December 2001), 346–362.
12. Chatterjee, S.; and Hadi, A.S. *Regression Analysis by Example*. Somerset: Wiley, 2015.
13. Chhabra, A.; Kaur, R.; and Iyengar, S.R.S. Dynamics of edit war sequences in Wikipedia. In *Proceedings of the 16th International Symposium on Open Collaboration*. 2020, pp. 1–10.
14. Christensen, L.R. Practices of stigmergy in architectural work. In *Proceedings of the 2007 International ACM Conference on Supporting Group Work - GROUP ’07*. Sanibel Island, FL: ACM Press, 2007, p. 11.

15. Christensen, L.R. Stigmergy in human practice: Coordination in construction work. *Cognitive Systems Research*, *21*, (March 2013), 40–51.
16. Christensen, L.R. Practices of stigmergy in the building process. *Computer Supported Cooperative Work (CSCW)*, *23*, 1 (February 2014), 1–19.
17. Cooke, N.J.; Salas, E.; Cannon-Bowers, J.A.; and Stout, R.J. Measuring team knowledge. *Human Factors*, *42*, 1 (2000), 151–173.
18. Cosley, D.; Frankowski, D.; Terveen, L.; and Riedl, J. SuggestBot: Using intelligent task routing to help people find work in Wikipedia. In *Proceedings of the 12th International Conference on Intelligent User Interfaces - IUI '07*. Honolulu, HI: ACM Press, 2007, p. 32.
19. Crowston, K.; Osterlund, C.S.; Howison, J.; and Bolici, F. Work features to support stigmergic coordination in distributed teams. *Academy of Management Proceedings, 2017*, 1 (August 2017), 14409.
20. Crowston, K.; and Rezgui, A. Effects of Stigmergic and Explicit Coordination on Wikipedia Article Quality. In *IEEE Computer Society*, 2020, pp. 2346–2355.
21. Crowston, K.; Saltz, J.S.; Rezgui, A.; Hegde, Y.; and You, S. Socio-technical affordances for stigmergic coordination implemented in MIDST, a tool for data-science teams. *Proceedings of the ACM on Human-Computer Interaction*, *3*, CSCW (November 2019), pp. 1–25.
22. Dabbish, L.; and Kraut, R. Research note—Awareness displays and social motivation for coordinating communication. *Information Systems Research*, *19*, 2 (June 2008), 221–238.
23. DeChurch, L.A.; and Mesmer-Magnus, J.R. The cognitive underpinnings of effective teamwork: a meta-analysis. *Journal of Applied Psychology*, *95*, 1 (2010), 32.
24. Dillenbourg, P., ed. *Collaborative Learning: Cognitive and Computational Approaches*. Amsterdam; New York: Pergamon, 1999.
25. Dorigo, M.; Bonabeau, E.; and Theraulaz, G. Ant algorithms and stigmergy. *Future Generation Computer Systems*, *16*, 8 (June 2000), 851–871.
26. Elliott, M. Stigmergic Collaboration: The Evolution of Group Work: Introduction. *M/C Journal*, *9*, 2 (May 2006). https://doi.org/10.5204/mcj.2599
27. Espinosa, J.A.; Lerch, F.J.; and Kraut, R.E. Explicit versus implicit coordination mechanisms and task dependencies: One size does not fit all. In E. Salas and S.M. Fiore (eds.), *Team Cognition: Understanding the Factors That Drive Process and Performance*. Washington, D.C: American Psychological Association, 2004, pp. 107–129.
28. Espinosa, J.A.; Slaughter, S.A.; Kraut, R.E.; and Herbsleb, J.D. Team knowledge and coordination in geographically distributed software development. *Journal of Management Information Systems*, *24*, 1 (July 2007), 135–169.
29. Faraj, S.; Jarvenpaa, S.L.; and Majchrzak, A. Knowledge collaboration in online communities. *Organization Science*, *22*, 5 (October 2011), 1224–1239.
30. van Fenema, P.C.; Pentland, B.; and Kumar, K. Paradigm shifts in coordination theory. In *Academy of Management Annual Meeting, New Orleans*. Citeseer, 2004.
31. Fischer, M.M., and Getis, A. *Handbook of Applied Spatial Analysis: Software Tools, Methods, and Applications*. Berlin: Springer, 2010.
32. Forte, A.; Larco, V.; and Bruckman, A. Decentralization in Wikipedia governance. *Journal of Management Information Systems*, *26*, 1 (July 2009), 49–72.
33. Geiger, R.S.; and Halfaker, A. Using edit sessions to measure participation in Wikipedia. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*. 2013, pp. 861–870.
34. Goh, K.-I.; and Barabási, A.-L. Burstiness and memory in complex systems. *EPL (Europhysics Letters)*, *81*, 4 (February 2008), 48002.
35. Goodman, P.S.; and Leyden, D.P. Familiarity and group productivity. *Journal of Applied Psychology*, *76*, 4 (1991), 578.
36. Gordon, D.M. Ant encounters. In *Ant Encounters*. Princeton, NJ: Princeton University Press, 2010. 39. https://www.google.com/books/edition/AntEncounters/MabwdXLZ9YMC?hl=en&gbpv=0
37. Grassé, P.-P. The nest reconstruction and inter-individual coordination in Bellicositermes natalensis and Cubitermes sp. The theory of stigmergy: An attempt to interpret the behavior of constructing termites. *Insectes sociaux*, *6*, 1 (1959), 41–80.

38. Heylighen, F. 2007. Why is open access development so successful? Stigmergic organization and the economics of information. In B. Lutterbeck, M. Bärwolff, & R. A., Gehring. (Eds.), *Open Source Jahrbuch*. (pp. 165–180). Lehrmanns Media. https://researchportal.vub.be/en/publications/why-is-open-access-development-so-successful-stigmergicorganizat
39. Hinnosaar, M.; Hinnosaar, T.; Kummer, M., and Slivko, O. Wikipedia matters. *Journal of Economics & Management Strategy* 2021. https://onlinelibrary.wiley.com/action/showCitFormats?doi=10.1111%2Fjems.12421
40. Howison, J.; and Crowston, K. Collaboration through open superposition: A theory of the open-source way. *MIS Quarterly*, *38*, 1 (January 2014), 29–50.
41. Kane, G.C.; Alavi, M.; Labianca, G. (Joe); and Borgatti, S.P. What's different about social media networks? A framework and research agenda. *MIS Quarterly*, *38*, 1 (January 2014), 274–304.
42. Kane, G.C.; Johnson, J.; and Majchrzak, A. Emergent life cycle: The tension between knowledge change and knowledge retention in open online coproduction communities. *Management Science*, *60*, 12 (December 2014), 3026–3048.
43. Kane, G.C.; and Ransbotham, S. Research note—Content and collaboration: An affiliation network approach to information quality in online peer production communities. *Information Systems Research*, *27*, 2 (June 2016), 424–439.
44. Kane, G.C.; and Ransbotham, S. Content as community regulator: The recursive relationship between consumption and contribution in open collaboration communities. *Organization Science*, *27*, 5 (October 2016), 1258–1274.
45. Ke, W.; and Zhang, P. The effects of extrinsic motivations and satisfaction in open source software development. *Journal of the Association for Information Systems*, *11*, 12 (December 2010), 784–808.
46. Kittur, A.; and Kraut, R.E. Harnessing the wisdom of crowds in Wikipedia: Quality through coordination. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, 2008, pp. 37–46.
47. Kittur, A.; Suh, B.; Pendleton, B.A.; and Chi, E.H. He says, she says: Conflict and coordination in Wikipedia. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '07*. San Jose, CA: ACM Press, 2007, pp. 453–462.
48. Krancher, O.; Luther, P.; and Jost, M. Key affordances of platform-as-a-service: Self-organization and continuous feedback. *Journal of Management Information Systems*, *35*, 3 (July 2018), 776–812.
49. Lappas, T.; Arai, B.; Platakis, M.; Kotsakos, D.; and Gunopulos, D. On burstiness-aware search for document sequences. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '09*. Paris: ACM Press, 2009, p. 477.
50. Lindberg, A.; Berente, N.; Gaskin, J.; and Lyytinen, K. Coordinating interdependencies in online communities: A study of an open source software project. *Information Systems Research*, *27*, 4 (December 2016), 751–772.
51. Maldeniya, D.; Budak, C.; Robert Jr., L.P.; and Romero, D.M. Herding a deluge of good Samaritans: How GitHub projects respond to increased attention. In *Proceedings of The Web Conference 2020*. Taipei: ACM, 2020, pp. 2055–2065.
52. Malone, T.W.; and Crowston, K. The interdisciplinary study of coordination. *ACM Computing Surveys*, *26*, 1 (March 1994), 87–119.
53. Marks, M.A.; Mathieu, J.E.; and Zaccaro, S.J. A Temporally based framework and taxonomy of team processes. *Academy of Management Review*, *26*, 3 (July 2001), 356–376.
54. Moon, J.Y.; and Sproull, L.S. The role of feedback in managing the internet-based volunteer work force. *Information Systems Research*, *19*, 4 (December 2008), 494–515.
55. Moran, P.A. Notes on continuous stochastic phenomena. *Biometrika, 37*, 1/2 (1950), 17–23.
56. Olson, J.F.; Howison, J.; and Carley, K.M. Paying attention to each other in visible work communities: Modeling Bursty Systems of multiple activity streams. In *2010 IEEE Second International Conference on Social Computing*. IEEE, Minneapolis, MN, USA, 2010, pp. 276–281.

57. O'Riordan, S.; Emerson, B.; Feller, J.; and Kiely, G. The road to open news: A theory of social signaling in an open news production community. *Journal of Management Information Systems*, *40*, 1 (January 2023), 130–162.
58. Ransbotham, S.; and Kane, G.C. Membership turnover and collaboration success in online communities: Explaining rises and falls from grace in Wikipedia. *MIS Quarterly*, *35*, 3 (2011), 613.
59. Rezgui, A.; and Crowston, K. Stigmergic coordination in Wikipedia. In *Proceedings of the 14th International Symposium on Open Collaboration*. Paris: ACM, 2018, pp. 1–12.
60. Riedl, C.; and Woolley, A.W. Teams vs. crowds: A field test of the relative contribution of incentives, member ability, and emergent collaboration to crowd-based problem-solving performance. *Academy of Management Discoveries*, *3*, 4 (December 2017), 382–403.
61. Rosen, M.A.; Fiore, S.M.; Salas, E.; Letsky, M.; and Warner, N. *Tightly Coupling Cognition: Understanding How Communication and Awareness Drive Coordination in Teams*. University of Central Florida Orlando, 2008.
62. Sumpter, D.J.; and Beekman, M. From nonlinearity to optimality: Pheromone trail foraging by ants. *Animal behaviour*, *66*, 2 (2003), 273–280.
63. Tan, X. (Jane); Jin, F.; and Dennis, A.R. How appreciation and attention affect contributions to electronic networks of practice. *Journal of Management Information Systems*, *39*, 4 (October 2022), 1037–1063.
64. Theraulaz, G.; and Bonabeau, E. A Brief history of stigmergy. *Artificial Life*, *5*, 2 (April 1999), 97–116.
65. Van de Ven, A.H.; Delbecq, A.L.; and Koenig Jr, R. Determinants of coordination modes within organizations. *American Sociological Review*, (1976), 322–338.
66. Velichety, S.; Ram, S.; and Bockstedt, J. Quality assessment of peer-produced content in knowledge repositories using development and coordination activities. *Journal of Management Information Systems*, *36*, 2 (April 2019), 478–512.
67. Wang, K.; and Nickerson, J.V. A Wikipedia-based method to support creative idea generation: The role of stimulus relatedness. *Journal of Management Information Systems*, *36*, 4 (2019), 1284–1312.
68. Wikipedia Contributors. Wikipedia:Good articles. *Wikipedia*, 2022. https://en.wikipedia.org/w/index.php?title=Wikipedia:Good_articles&oldid=956946482.
69. Young, A.G.; Wigdor, A.D.; and Kane, G.C. The gender bias tug-of-war in a co-creation community: Core-periphery tension on Wikipedia. *Journal of Management Information Systems*, *37*, 4 (October 2020), 1047–1072.
70. Zhang, C.; Luo, L.; Xu, W.; and Ledwith, V. Use of local Moran's I and GIS to identify pollution hotspots of Pb in urban soils of Galway, Ireland. *Science of the Total Environment*, *398*,1–3 (2008), 212–221.
71. Zhang, X. (Michael); and Zhu, F. Group size and incentives to contribute: A natural experiment at Chinese Wikipedia. *American Economic Review*, *101*, 4 (June 2011), 1601–1615.
72. Zheng, L. (Nico); Albano, C.M.; Vora, N.M.; Mai, F.; and Nickerson, J.V. The roles bots play in Wikipedia. *Proceedings of the ACM on Human-Computer Interaction*, *3*, CSCW, November 2019, pp. 215:1-215:20.
73. Zhu, H.; Zhang, A.; He, J.; Kraut, R.E.; and Kittur, A. Effects of peer feedback on contribution: a field experiment in Wikipedia. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 2253–2262.
74. Zhu, K.; Walker, D.; and Muchnik, L. Content growth and attention contagion in information networks: Addressing information poverty on Wikipedia. *Information Systems Research*, *31*, 2 (June 2020), 491–509.