All That Is Gold Does Not Glitter: A Framework for Content Quality in an Online Collaborative Content-Orientated Platform

Tomislav Viljevac, Max Rohde, and David Sundaram

Abstract—**The ubiquitous nature of the Internet has resulted in an ever increasing use of and reliance upon, online content created through various modes and media. Many contemporary online platforms such as Wikipedia enmesh rich collaboration with extensive content-repositories. However, such online collaborative content-orientated platforms (OCCP) have been criticized for the reliability and quality of the content created and hosted. To this end, research has provided significant insights into the enabling factors for social transparency and trust within these platforms. We expand on these insights through an extensive survey of mechanisms employed in successful and popular OCCPs. We reflect upon the findings of this survey using the lenses of activity theory and legitimate peripheral participation. A framework is synthesized from this discussion, which stresses the importance of the interaction between quality assurance and the attraction of new content; two fundamental, yet potentially conflicting, processes that drive the success of OCCPs.**

Index Terms—**online collaborative content-orientated platform, knowledge management, wiki, framework, content quality.**

I. INTRODUCTION

The ubiquity of the Internet and the resulting ease of access to extensive repositories of information have led to an increased use and reliance on the content provided by this medium. Content-oriented platforms are typically websites allowing individuals singularly or collaboratively, to create and edit pages presenting various forms of content [1]-[3]. Online collaborative content-orientated platforms (OCCP) encompass any web content host page that allows a group to collaboratively create content for specific interests[1]. While the online encyclopedia, Wikipedia, is the most prevalent example of a platform comprised of collaboratively edited content, there are myriads of smaller platforms in the form of discussion forums, video tubes and message boards that serve large and small specific interests groups. Although differences exist in some socio-technical approaches to assessing quality and attracting content to a platform, such as rolling back in wikis to a previous state or granting permission to edit already submitted content, wikis and other platforms can be deemed OCCP's by that in practice they

focus on content creation by anybody using simple means [1]-[3]. The importance and popularity of these platforms has spurred a large number of scholarly investigations concerned with how the quality of contributions on such platforms can be measured and which mechanisms can be employed to uphold and advance the quality of contributions [4]-[15]. Other streams of literature investigate why people are motivated to contribute to online platforms or reference their content [8], [9], [12], [16]-[18].

Nevertheless, many platforms based on collaborative content are deemed unsuccessful. In particular, organizational discussion forums, wikis and other forms of knowledge repositories often contain little and out-dated information [18, 19]. These platforms fail in two ways: they fail to attract a sufficient amount of content from users, and they fail to provide content of sufficient quality. While valuable insights are presented in the literature to address each of these challenges, it is seldom considered how these two dimensions interact. We therefore endeavor to untangle the interactions between these two dimensions of participation and assuring content quality in the context of information handling processes (e.g. create, read, update, delete) involved in collaborative content creation [15]. Literature on communities of practice [8], [12], [17], [18] and participation in wikis [4], [5], [8], [16] looks at motivating participation and highlights the potential value to content quality if platforms can cater to the needs of various user groups, such as zealots, professionals and novices [5]. It has further been suggested to explore these issues in the context of social theories such as Legitimate Peripheral Participation (LPP) and Activity Theory (AT). These frameworks are used to explain socio-technical interaction and how casual users become on-going contributors and members of a community of practice [8]. Varying degrees of motivation dictate the extent of one's interaction with a platform, and vice versa. The platform must have mechanisms to address this occurrence. Meeting users motivational factors for participation can help attract content and provide value by inspiring zealots that are proven to provide higher quality content [5].

The notion of creating trust through providing transparency of information is an important aspect to consider in an OCCP. If trust can be created through various mechanisms on the platform, a direct effect on quality can be observed. Increasing trust enables positive judgment of quality [7], [11], [13], [17], [20]. An example of information transparency is providing information about the origin of content. Effective communication on collaborative content-orientated platforms stems from making socially significant information visible, by accounting for the who, what, where, when and how of information [13], [14], [20]. Vast literature and existent platforms provide many examples of methods that help to assure quality and provide visibility of social data for the success of the online collaborative content-orientated

Tomislav Viljevac, Max Rohde and David Sundaram are with Department of Information Systems and Operations Management University of Auckland Auckland, New Zealand (e-mail: tvil007@aucklanduni.ac.nz, m.rohde@auckland.ac.nz, d.sundaram@auckland.ac.nz).

platform. These methods will be discussed in their respective sections.

We first elicit mechanisms that would help attract content (CAM) and assure content quality (QAM). Informed by the identification of these mechanisms, a framework of OCCP success is produced. This framework considers the perspectives of attracting and supporting content contribution by users and assuring its quality. Section II describes the research approach undertaken. The identified CAMs and QAMs will then be presented and discussed in Sections III and IV. Section V will introduce and discuss the framework and how the CAMs and QAMs combine together. The final section VI will summarize the paper.

II. ELICITATION OF CONTENT ATTRACTION AND QUALITY ASSURANCE MECHANISIMS

A. Platform Examination

In order to elicit mechanisms for content attraction and quality assurance, we have undertaken four intertwined activities: (1) the examination of OCCPs, (2) a survey of relevant literature, (3) the initial identification of control functions, and (4) the synthesis of findings across examined platforms and streams of literature (Fig. 1).

The OCCPs examined are Wikipedia, StackOverflow, Amazon.com, HotCopper and Youtube. These platforms were chosen as they are largely centered on creating content through collaboration. We wanted to identify common approaches in signaling quality to users and attracting content. Amazon may be an online retail store but collaboration on creating content is an integral part of the purpose, function and ultimately success of the platform. Similarly, Youtube was regarded as a collaborative content orientated platform, exhibiting aspects of creating and collaborating on content similar to other platforms. From the mechanisms of quality attraction and assurance observed, it is evident that the Youtube community endeavors to provide social transparency and promote entrusting of the platform by both readers and authors.

HotCopper.com.au is a leading forum for investors in the Australian Stock Exchange. By purpose and function the forum is a good example of a collaborative content-orientated platform having content and knowledge creation and collaboration as a quintessential aspect. Overall these platforms resemble mainly aspects of wikis by acting as a medium of discussion, a repository of information and as tools for collaborating [9].
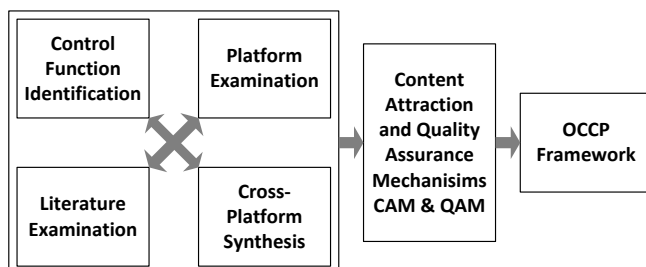

Fig. 1. Research approach outline

B. Control Function Identfication and Literature Examination

The world's oldest wiki, started in 1995 by Ward Cunningham provided a base list of control functions as indicators for the kinds of controls and mechanisms for content creation and quality assurance that can be expectant of an OCCP today [9]. It was identified as being the first OCCP, and we evaluated platforms in light of these wiki controls introduced by the Wikiwiki community. Examples of controls include usernames, recent changes display and edit text icons and boxes.

Literature was reviewed to ascertain additional socio-technical mechanisms proposed and employed by platforms. Almost all literature proposed methods that increase transparency of information. Furthermore, mechanisms were employed to support social interactions such as the governance structure of article deletion in Wikipedia, and the procedures around assigning reputation metrics to users.

C. Cross Platform Synthesis

The platforms were compared amongst one another to see which mechanisms were common across the sample, and to what extent content attraction and content quality assurance was supported by each platform. Comparison between platforms helped to create the framework by making it possible to see trends in the application of mechanisms; proving existence of a mechanism across all platforms signals the importance of the mechanism to further validate it as being a CAM or QAM. On the other hand, a mechanism featuring only in one platform does not necessarily signal that this mechanism is not of high applicability to being a CAM or QAM. It may be that it is specific to the purpose of a platform, and thus is still significant in providing overall attraction and quality assurance. Naturally, across different platforms and mostly evident with Wikipedia, different functions and purposes of the platform will call for unique platform specific mechanisms. The same mechanism may contribute differently in other platforms, as quality is community-of-interest specific and socially constructed [12]. Moreover, platforms that featured more mechanisms than others, could signal a higher success at attracting and assuring quality of content, and thus be a platform of higher quality content. Cross-platform comparison enabled easier judgments on overall quality of a platform due to relative comparison

D. Toward a CAM, QAM and OCCP Success Framework

Lastly, grouping the mechanisms as attracting content, assuring quality of content, or both, is the foundation of the OCCP success framework. Based on the four mechanism derivation techniques (Fig. 1), the mechanisms were deemed to be significant to either attracting content, such as directly allowing users to submit to a web page or making the process of contribution easier, or assuring quality, such as providing hyperlinks to content citations. Some mechanisms were deemed to be both. The mechanisms highlighted were hoped to support the underlying context of LPP and AT.

III. ATTRACTING CONTENT CONTRIBUTION

Factors that influence contribution and patronage to an online collaborative content-orientated platform are numerous and have been grouped in this study to be intrinsic and extrinsic factors, of which some are the willingness to learn, altruism and the rewards of reputation [8], [16]-[19]. Mechanisms need to address the physical and social aspects of interacting on a platform in a community as per AT and LPP, and the CAMs are often direct motive to contribute and provide the means for submission of content.

TABLE I: SUMMARY COMPARISON OF CROSS-PLATFORM CAMS AND QAMS

	Wikipedia	StackOverflow	Amazon	YouTube	Hot Copper	Literature
CAM	7	10	8	7	9	2
QAM	28	30	27	26	28	22
Both	5	8	6	3	5	5
Total	40	48	41	36	42	29

Out of a total of 71 mechanisms identified, 10 were deemed to be as content attracting mechanisms, yet another further 10 were considered as acting as both a mechanism of attraction and quality assurance (Table I). The types of mechanisms identified as strictly attracting content are those that enable direct submission of content to the platform, such as controls for editing text or the ability of updating a user profile. Even simply listing out a need for contributors and collaboration on a specific topic in Wikipedia may prompt the reader to sign up and share knowledge, just as supplying FAQs on how to edit articles can be a control to not deter contribution. However difficult and subjective it may be to determine whether a mechanism be determined as strictly an attraction mechanism, a CAM is deemed to be so, based on the extent of the mechanism initially assuring the quality of the content. The attraction mechanisms simply act as a segue of content contribution. Mechanisms that have been highlighted as being both CAM and QAM tend to work backwards in that quality assurance happens first, and has influenced a judgment of the platform and it's quality by the user and thus has influenced their motive to contribute.

Key mechanisms that are deemed both a CAM and a QAM involve aspects of reputation such as visibility of ratings by other users [4], [15]. Discussion pages and chats, that offer alternative simpler, less structured manner of content contribution are also deemed both CAM and QAM. Discussions and chats can signal quality as activity and applying extra rigor to the collaboration process deems content current and alive. Moreover, the ability to format and tailor content helps users who want to be more involved and comprehensive. They go beyond the basic functionality and can address AT and LPP in OCCP interaction [8]. This motivates the individual's intentions for contribution, but also ensures some quality control is applied at the input process. Similarly specifying the mission statement of the community is a further way to align users and inspire contribution.

IV. QUALITY ASSURANCE

Literature concerning content quality focuses generally on reputation [4], [5], [12], [15], [20] as well as factors in combination of the author, the number of authors, edit history, edit frequencies [20] and the length of the edited content [7]. The underlying notion in these studies is that rigor and frequency of authoring instances plays an integral part in reputation. Indeed, the majority of the 51 various quality assurance mechanisms elicited from platforms (Table. 1) involve capturing and presenting metrics on edit histories, authors and their reputation. Blumenstock [7] discusses word length as an indicator of content quality

where the length of an article would signal higher quality content. Similarly, Druck et al. [21] discuss how a simple quality metric of measuring the frequency of a submission being reverted can influence judgment of quality. A proposed quality assurance approach was to attribute authors a score if their content endured existence, and was not subject to edits by other authors [4], [15]. Another key mechanism for quality assurance is history flow, that encompasses the amount of content edited, the time, and also the author so as to create a dashboard visually displaying how content has collaboratively grown [22].

Platform analysis has yielded examples that consider the activity of participants. Wikipedia visitors can compare revision histories of particular pages of content. All platforms in fact featured reputation metrics for users, such as simply registering their submission being helpful and on topic. StackOverflow provided a net score on users contributing content depending if their answers were voted to be of use to other users. Similarly the employment of badges and bounties to achieve in reward for quality submissions is clearly on display. It is worth noting that propositions for quality assurance identified in literature considered synthesizing multiple mechanisms and metrics however, this was not entirely evident in current platforms observed. All these examples exemplify how transpiring the source and frequency of information has an impact on the users judgment of content quality. Quality assurance mechanisms in combination with content attraction mechanisms therefore instill trust as they enable judgment of quality amongst users and affect their overall motivation to be part of a collaborative content community.

V. FRAMEWORK FOR SUCCESS THROUGH CONTENT ATTRACTION AND QUALITY ASSURANCE

The final resultant artefact produced is the OCCP Success framework (Fig. 2) integrating the dimensions of content attraction and content quality assurance. The framework conceptualizes the interactions and relationships of the motivation elements that are intrinsic and extrinsic, with the dimensions of content attraction and quality assurance and the dimension of social transparency. It depicts the interaction that quality assurance and content attraction is influenced with initial and subsequent quality assurance and content attraction. Social transparency motivates judgments of quality and resultantly can affect intrinsic and extrinsic facets of motivation to interact with an OCCP. As users contribute, the socio-technical mechanisms CAMs and QAMs, are also enablers of learning by peripheral participation and the means to bend dimensions of community, divisions of labor, tools, and rules that are involved when subjects interact socio-technically with objects in a context [8]; in this instance the particular content in an OCCP. Supporting the physical and social participation and assuring quality of content through these mechanisms, helps toward ensuring longevity and success of the online collaborative content-orientated platform.
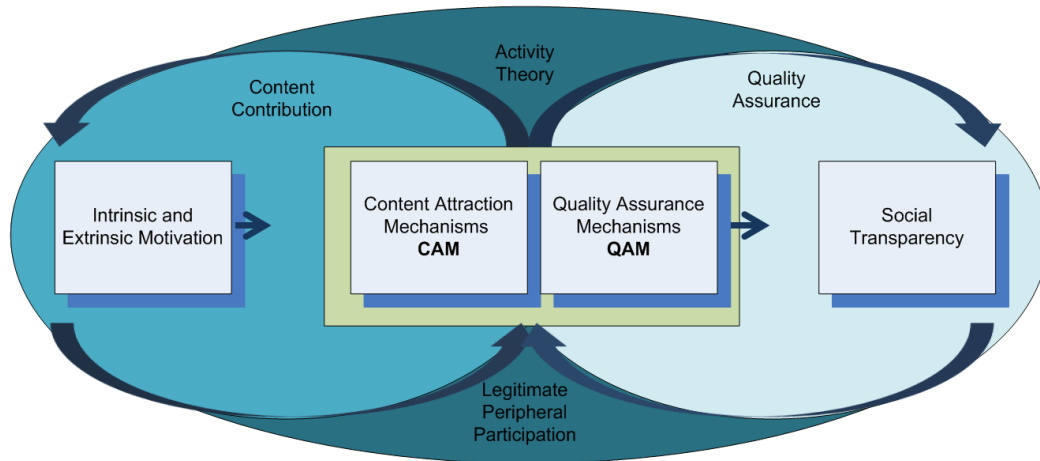
Fig. 2. OCCP success framework

VI. CONCLUSION

Increased use and reliance on online content of collaborative platforms brings about the need to address the quality of content in conjunction with enabling the provision of content. This paper has discussed the recurrent call for transpiring information in online platforms that impact on the trust and quality judgments of content. We have suggested harnessing content attraction mechanisms alongside quality assurance mechanisms that work collectively to provide transparency of information origin. Literature reviewed for the purposes of eliciting mechanisms has provided many propositions that are currently implemented to limited extent with regards to synthesizing mechanisms. Platforms observed have been shown to exemplify aspects of content attraction and quality assurance mechanisms in different forms, yet they all share the common goal of providing content that can be relied upon, irrespective of the extent to which rigor in assuring and attracting quality has been applied. The paper finally proposes a framework that helps work towards successful longevity of platforms, through assuring quality, and resonating aspects of the platform with those of individuals who contribute or simply reference the content.

REFERENCES

[1] K. R. Parker and J. T. Chao, "Wiki as a Teaching Tool," *Interdisciplinary Journal of Knowledge and Learning Objects,* vol. 3, pp. 57-72, 2007.

[2] T. Franklin and M. Van Harmelen," Web 2.0 for Learning and Teaching in Higher Education," Franklin Consulting and Mark van Harmelen, London2007.

[3] C. Wagner and N. Bolloju, "Supporting Knowledge Management in Organizations with Conversational Technologies: Discussion Forums, Weblogs, and Wikis," *Journal of Database Management* vol. 16, 2005.

[4] T. Adler, L. de Alfaro, I. Pye, and V. Raman, "Measuring author contributions to the Wikipedia," *the 2008 International Symposium on Wikis and Open Collaboration (WikiSym'08)*, Porto, Portugal., 2008, September.

[5] D. Anthony, S. Smith, and T. Williamson, "Explaining Quality in Internet Collective Goods: Zealots and Good Samaritans in the Case of Wikipedia," *Fall 2005 Innovation & Enterpreneurship Seminar at MIT,* 2005.

[6] O. Arazy, *et al.*, "Recognizing contributions in wikis: Authorship categories, algorithms, and visualizations," *J. Am. Soc. Inf. Sci. Technol.,* vol. 61, pp. 1166-1179, 2010.

[7] J. Blumenstock, "Size matters: word count as a measure of quality on wikipedia," in *WWW '08: Proceeding of the 17th international conference on World Wide Web*, *Beijing, China*, pp. 1095-1096, 2008.

[8] S. L. Bryant, A. Forte, and A. Bruckman, "Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia," *the Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*, Sanibel Island, Florida, USA, 2005.

[9] W. Cunningam. (7th December, 1995). [Online]. Available: http://c2.com/doc/pages.html

[10] B. Luyt, T. C. H. Aaron, L. H. Thian, and C. K. Hong, "Improving Wikipedia's accuracy: Is edit age a solution?," *Journal of the American Society for Information Science & Technology,* vol. 59, pp. 318-330, 2008.

[11] B. Suh, E. Chi, A. Kittur, and B. Pendelton, "Lifting the Veil: Improving Accountability and Social Transparency in Wikipedia with WikiDashboard," *the CHI 2008: Proceedings of the ACM Conference on Humanfactors in Computing Systems, New York*, 2008.

[12] B. Stvilia, M. Twidale, L. C. Smith, and L. Gasser, "Information quality work organization in wikipedia," *Journal of the American Society for Information Science and Technology,* vol. 59, pp. 983-1001, 2008.

[13] T. Erickson, C. Halverson, W. A. Kellogg, M. Laff, and T. Wolf, "Social translucence: designing social infrastructures that make collective activity visible," *Commun. ACM,* vol. 45, pp. 40-44, 2002.

[14] B. Shneiderman, "Designing trust into online experiences," *Comm. ACM,* vol. 43, pp. 57-59, 2000.

[15] T. Adler and L. de Alfaro, "A content-driven reputation system for the wikipedia," in *Proc. WWW '07: Proceedings of the 16th international conference on World Wide Web*, Banff, Alberta, Canada, pp. 261-270, 2007.

[16] H. Baytiyeh and J. Pfaffman, "Why be a Wikipedian," *the Proceedings of the 9th international conference on Computer supported collaborative learning*, vol. 1, 2009.

[17] A. Ardichvili, V. Page, and T. Wentling, "Motivation and barriers to participation in virtual knowledge-sharing communities of practice," *Journal of Knowledge Management,* vol. 7, pp. 64-77, 2003.

[18] M. E. Rohde and D. Sundaram, "Challenges in Knowledge Management," *AMCIS 2010 Proceedings*, 2010.

[19] M. W. Molly and S. Faraj, "Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice," *MIS Quarterly,* vol. 29, 2005.

[20] A. Lih, "Wikipedia as Participatory Journalism: Reliable Sources? Metrics for evaluating collaborative media as a news resource," in *Proc. 5th International Symposium on Online Journalism*, 2004.

[21] G. Druck, G. Miklau, and A. McCallum, "Learning to Predict the Quality of Contributions toWikipedia," *Association for the Advancement of Artificial Intelligence,* 2008.

[22] F. Vigas, M. Watteberg, and K. Dave, "Studying cooperation and conflict between authors with history flow visualizations," in *Proc. CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*, Vienna, Austria, pp. 575-582, 2004.

T. Viljevac is a current Master's student with the University of Auckland Business School, working on collaborative platform success from the perspective of content quality driven user participation. His other interests in the information systems domain include adaptive business information systems, semantic web technologies and application of collaborative platform and web 3.0 to digital marketing.

M. Rohde is a PhD student in Information Systems at the University of Auckland, New Zealand. His interests include the design and implementation of information-intensive systems for knowledge-driven organizations, in particular the effective deployment of modern database systems and cloud technologies.

D. Sundaram is an engineer by background, a teacher, researcher, and consultant by profession, and a lifelong student. He is passionate about the modeling, design, and implementation of flexible and evolvable information, visualisation, decision, knowledge, social and ubiquitous systems.