2008

Wikipedia Usage Patterns: The Dynamics of Growth

Amitava Dutta
George Mason University, adutta@gmu.edu

Rahul Roy
Indian Institute of Management, Calcutta, rahul@iimcal.ac.in

Priya Seetharaman
Indian Institute of Management, Calcutta, priyas@iimcal.ac.in

WIKIPEDIA USAGE PATTERNS: THE DYNAMICS OF GROWTH

Les usages types de wikipédia : la dynamique de croissance

Amitava Dutta
School of Management
George Mason University
Fairfax, VA 22030-4444, U.S.A
adutta@gmu.edu

Rahul Roy
MIS Group
Indian Institute of Management Calcutta
Calcutta 700104, India
rahul@iimcal.ac.in

Priya Seetharaman
MIS Group
Indian Institute of Management Calcutta
Calcutta 700104, India
priyas@iimcal.ac.in

Abstract

Wikis have attracted attention as a powerful technological platform on which to harness the potential benefits of collective knowledge. Current literature identifies different behavioral factors that modulate the interaction between contributors and wikis. Some inhibit growth while others enhance it. However, while these individual factors have been identified in the literature, their collective effects have not yet been identified. In this paper, we use the system dynamics methodology, and a survey of Wikipedia users, to propose a holistic model of the interaction among different factors and their collective impact on Wikipedia growth. The model is simulated to examine its ability to replicate observed growth patterns of Wikipedia metrics. Results indicate that the model is a reasonable starting point for understanding observed Wiki growth patterns. To the best of our knowledge, this is the first attempt in the literature to synthesize a holistic model of the forces underlying Wiki growth.

Keywords: Wikipedia, behavioral factors, system dynamics, simulation, survey data

Résumé

Le Wiki est une puissante plateforme technologique qui a pour but de mettre a profit la connaissance collective. Les écrits actuels identifient certains facteurs comportementaux individuels modulant l'interaction entre les contributeurs et le wiki mais leur effet collectif reste à identifier. Nous utilisons a la fois la méthodologie dynamique du système ainsi qu'un sondage des utilisateurs de Wikipédia pour proposer un modèle holistique de l'interaction entre les différents facteurs et leur impact collectif sur l'expansion de Wikipédia.

Introduction

Wikipedia is an encyclopedic source used by many people of all ages, backgrounds, and nationalities (Voss 2005). Although Wikipedia may be the most recognizable wiki, wiki technology has become extremely popular and has been used by numerous organizations for a variety of purposes (Hof 2004, Morse 2008). Many wiki communities are private, particularly those within firms, where they are often used for internal documentation and knowledge management (McKelvie et al 2007). Regardless of whether a wiki is public or private, a defining characteristic of this technology is that anyone in that wiki community can add to and edit content on the Web site with relative ease.

This openness has generated conflicting views on the value of wikis and the reasons why some are successful while others are not. Anecdotal evidence can be found to support both sides. One view holds that 'wisdom of the crowds' automatically results in careful scrutiny of content as well as a more comprehensive representation of knowledge about the issue being covered in the wiki. After all, if so many readers are examining the content, errors will be caught quickly by someone or the other. Similarly, with so many contributors, one is likely to have more complete information resulting from the fusion of so many diverse sources. Contributions stand on their own merit and if they stand up to inspection, they stay – otherwise they go. In short, this view essentially believes that continuous tinkering only makes things better. The alternate view cast doubt on this outcome. It notes that being accessible to everyone, wikis are susceptible to both malicious actions and incompetence. For controversial topics, or for topics that are not of widespread interest, content quality can be compromised. These opposing views have conflicting implications for the attractiveness of wikis to seekers of information. The first view suggests that a wiki will become more attractive to information seekers over time, while the second view suggests that a wiki may not sustain interest in information seekers and may eventually atrophy.

Contributors who add and update content on a wiki, also face opposing forces. There usually is no monetary reward for contributing, so the incentive for contributing comes largely from intrinsic motivation, defined as interest in or enjoyment of an activity for its own sake (Zhang and Zhu 2006). Research has shown that one of the factors that contributes to intrinsic motivation is a sense of meaningfulness – a sense that what one is doing has purpose or serves greater good. Wikipedia, for example, aims to give everyone free access to the sum total of human knowledge. This lofty goal clearly appeals to many. This sense of purpose is reinforced by seeing their edited articles being read by a large community of information seekers. Moreover, when a contributor is a registered user, he/she can gain recognition from peers for content they have generated provided it stands the test of time. However, if contributions are frequently deleted by others in the community, or if the content does not attract sufficient interest from information seekers, the motivation of contributors may begin to wane.

In summary, a variety of forces act on both seekers and contributors of wiki content (Sanger 2005). However, these forces do not act independently of one another. They interact among themselves and generate feedback effects. For instance, the motivation that a contributor feels from recognition by peer contributors may be diminished if there is no significant readership of the content. Similarly, increased readership of content can motivate more contributors, increasing content quality, which in turn brings in more readers. It is this interaction among counteracting forces and their collective effect that determines whether or not a wiki effort takes root and grows or atrophies and becomes 'extinct'. To the best of our knowledge, a holistic model that integrates these counteracting forces to help us better understand the mechanics of growth or demise of a wiki, has not appeared in the literature. In this paper, we develop such a holistic model using the system dynamics methodology and a small exploratory survey. The model contributes to the literature by revealing how known factors interact and generate causal mechanisms leading to a wikis growth or demise.

The remainder of the paper is structured as follows. In the next section, we survey the literature and document different factors that have been found to influence the behavior of contributors to a wiki. This is followed by findings from a small exploratory survey we conducted for this study, in which the majority of respondents were Wikipedia information seekers. The holistic model that integrates these different factors is then developed. The model is simulated to examine its ability to replicate observed patterns of growth for Wikipedia and demonstrate its potential use for descriptive purposes. The model is then simulated under a hypothetical scenario to demonstrate its potential use for prescriptive purposes. The value of a system dynamics approach is revisited at this point. Limitations of and extensions to the current work are discussed in conclusion.

Content Contributing and Seeking Behavior – A Review

In many public domain wikis, such as Wikipedia, the vast majority of users are seekers of information. A much smaller proportion of users are contributors and editors of content. In private wikis, particularly in corporate settings, the proportion of contributors may be much higher. This partly stems from the kinds of applications for which corporations have used Wikis. They include product development, research and development, and diffusion of best practices. Since the Wiki paradigm of content development, whether in the public or private domain, is radically different from that of the more traditional proprietary and expert based approach to content development (Encyclopedia Britannica, for example), there have been studies reported in the literature that examine the factors the drive individuals to contribute content.

Based on a sample of 168 respondents from users of corporate wikis, Majchrzak et. al. (2006), found several patterns of behavior among the contributors. First, there were different types of contribution ranging from simple spell checking and editing, all the way to rewriting whole sections or rolling back content provided by others. The factors influencing individuals to contribute include earning the respect of others, improving professional status, and making it easier to do their work. The authors concluded that unlike open source software development contexts, corporate wiki contributors saw the participation benefits to be mainly work and organization related. Contributors also saw reputation benefits. This study found corporate wikis to be sustainable, as measured by longevity, number of participants and frequency of accesses.

One of the case studies in Wagner and Majchrzak (2007), however, reports a very different experience. This was more of a public domain application in which a newspaper publisher wanted to involve its readership to co-create an editorial using a wiki platform. The chosen top was Iraq. The registration mechanism was very simple, making it easy for anyone to contribute. For the first twenty four hours, the contributions and contributor interactions were very constructive. However, about thirty hours after the editorial wiki was initiated, it became difficult to keep opposing arguments within the same document, which was then split into two. Shortly thereafter, as the popularity of the site grew, it got the attention of malicious users and there were several incidents of destruction of wiki content. Coupled with technical shortcomings of the software platform that was used (not locking a page during edits), the site was overwhelmed by the contribution process and the experiment was shut down three days after it started. So in this experiment, the wiki became extinct and the benefits of recognition or contributing to a greater good were not enough to sustain the contribution process.

Nov (2007) examined various motivations for individuals to actively contribute on Wikipedia. Among the motivating factors, 'fun', 'ideology', 'enhancement' and 'protective' ranked high though the author did not find correlation between ideology and contribution rates. Intrinsic motivation to contribute in the context of Wikis has been operationalized as free contributor-to-contributor assistance, perceived competence and sense of relatedness (Zhang and Zhu, 2006). Wagner and Prasarnphanich (2007) examined the role of altruism in collaboratively creating content on a wiki platform. In summary, while the literature has identified a variety of factors that influence the contribution process for wikis, the factors that influence seekers of information in the wiki environment also need further investigation. While contributors, by definition, must also seek content before they can edit, individuals who are primarily seekers, are not necessarily influenced by the same factors. Two areas of inquiry in the literature can inform us here. One is technology acceptance models which examine factors influencing user acceptance of IT applications. The second is diffusion of innovations which examines the dynamics of permeation of a new technology within a user community.

Technology Acceptance Models

The original Technology Acceptance Model (TAM) (Davis 1986) captured users' perceptions regarding usefulness and ease of use of the technology and their influence on use through attitude and behavioral intention to use IT. There have been numerous extensions to this original model (e.g. Hu et al 1999; Venkatesh and Davis, 2000). Integrating TAM and other associated theories based on attitudes and beliefs, Venkatesh et al (2003) presented a unified view of user acceptance of IT. In doing so, they developed a theory using performance expectancy, effort expectancy, social influence and facilitating conditions, as predictors of behavioral intention and use behavior. Brown et al (2002), argued that extrinsic motivation (being told to use) plays as important a role as intrinsic motivation (inclination to use) in augmenting the impact of perceived ease of use on behavioral intention to use IT. The TAM literature reveals a collection of variables critical to understanding seeker behavior in wiki environments, including ease of use, ease of access and perceived usefulness. A complementary view of technology adoption, one

that focuses on the pattern of permeation of the technology over time, can be drawn from the diffusion of innovations literature.

Diffusion of Innovations

Lucas et al (2007) pointed out that the emergence of community technology platforms is indeed a new IT innovation and implementation challenge. Wikis are an epitome of such a community technology platform. In the classic diffusion of innovations (Rogers, 2003) model, an application or infrastructure is treated as an innovation and its relative advantage, compatibility, complexity, triability and observability aspects are examined to understand the process of use and diffusion. Some authors (Lou and Scamell, 1996) conclude that users' rejection of a system or lack of use may be due to absence of tasks or other reasons to use the system, inconvenience in accessing the system, or lack of understanding of system capabilities and/or awareness of its operations, while others (Kang and Kim, 1996) have used the twin concepts of internal personal belief and external social belief to understand technological innovation adoption. Diffusion of innovations literature thus emphasizes personal belief, relative advantage, perceived ease of use, triability and trust as factors aiding technology diffusion process while inaccessibility, lack of understanding and complexity act as inhibitors to the diffusion process. Kittur et al (2007) who studied the 'rise of the bourgeoisie' in Wikis, argued that adoption of Wikipedia can be compared to diffusion of an innovation as the constant change in Wikipedia makes it a dynamic social system.

From the foregoing review, it can be seen that the literature on the growth and adoption of wikis has identified a variety of factors that influence the contribution process for wikis, but is relatively thin on the content seeking process and its influencing factors. The literature on TAM and diffusion of innovations may prove useful in addressing the latter. This is the motivation for the exploratory survey of Wikipedia information seekers that follows. However, even after identifying influencing factors for content seekers, what one would have is still an unstructured collection of individual factors without a more complete picture of the architecture of their interaction that results in wiki growth. This is what motivates us to follow up the survey with development of a holistic model of wiki growth using the system dynamics methodology.

Exploratory Survey of Wiki Information Seekers

As noted in the previous section, available findings on factors which influence information seeking behavior on wiki platforms are quite meager. To address this need, we conducted an exploratory online survey using both closed and open ended questions based on findings in the TAM and diffusion of innovations literature summarized in the previous section. The questionnaire was targeted to current graduate students and alumni of the institution of one of the authors, and survey questions referred to a specific public domain wiki, namely Wikipedia. Our reasoning was that Wikipedia, being a public domain wiki, was accessible to all the potential respondents providing a common context for the responses. The questionnaire was divided into two parts, the first intended for information seekers, the second for contributors. About 25% of the sample were in the age group of 22 to 26 years and nearly 63% were in the age group of 27 to 40. Nearly 16% of the sample held a bachelor's degree while 82% held a master's degree or above. Of the usable responses, 119 classified themselves as primarily information seekers on Wikipedia. Only five respondents considered themselves to be primarily contributors to Wikipedia. The low proportion of contributors was not surprising since this wiki application, to create encyclopedic content, is intended mostly for readers, unlike the collaborative applications discussed previously for corporate wikis. We thus report findings from only these 119 seekers to complement contributor behavior findings summarized earlier. The constructs covered in the survey and their basis in the literature reviewed earlier are shown in Table 1. Each construct was operationalized for this particular wiki context.

The questionnaire was pretested with nine faculty members and doctoral candidates who conducted an initial review of these measures to establish face validity. Each pretester was asked to identify the construct corresponding to each of the scale instruments. Concurrence across pretesters ranged between .60 to .90 for the main constructs including ease of use, perceived usefulness, motivation to use and inhibitors to contribution. Cognitive pre-testing helped refine the language of some questions. Due to space constraints, the complete questionnaire is not shown, but sample questions for selected constructs appear in Table 2.

Table 1: Questionnaire Constructs and Conceptual Basis	
CONSTRUCT	**LITERATURE REFERENCE**
Ease of Use: (Ease of Navigation, operation, accessibility, availability)	Venkatesh, Morris, Davis and Davis (2003); Taylor and Todd (1995)
Perceived Usefulness: (Accuracy, Preciseness, Completeness and Usefulness of Information)	Baroudi and Orlikowski (1988)
Motivation to Use: (Peer Influences, Work Pressure, Perceived Efficiency, Easy Accessibility)	Taylor and Todd (1995)
Attitude towards the System	Venkatesh, Morris, Davis and Davis (2003)

Table 2: Selected Constructs and Sample Questions	
CONSTRUCT MEASURED	**SAMPLE**
Ease of Use – Seeker	o Wikipedia is easy to navigate. I am able to move from one page to another easily. o I can read and understand the standard structure (such as links, reference, classification of pages, etc.) of Wikipedia pages easily.
Perceived Usefulness	o I perceive information from Wikipedia to be accurate o I perceive information from Wikipedia to be in-depth
Motivation to Use	o I am motivated to use Wikipedia as it can be accessed electronically from my computer itself o Using Wikipedia significantly enhances my work quality o I use Wikipedia because it is free
Inhibition to Contribute	o I do not feel inclined to 'give back' to Wikipedia even though I seek information from it. o Seeing my name in the contribution page in Wikipedia is not enough incentive for me to contribute to Wikipedia. o I feel I don't have the expertise to contribute knowledge.

About 70% of respondents said that they used Wikipedia "out of informational curiosity", around 68% used it to find information to do their professional work and 63% to improve their general knowledge. When asked about their most common area of search, nearly 30% said they search for information related to history. Other common search topics included finance, geography, music and science and technology. Several respondents also stated that Wikipedia helps them look for "any information under the sun" and "tit-bits about anything that I hear". This reflects the intrinsic encyclopedic nature of Wikipedia.

Factors Influencing Use of Wikipedia

Descriptive statistics collected from the survey produced the following findings regarding factors influencing seeker behavior on Wikipedia.

Ease of Use: 94% of the respondents felt that Wikipedia is easy to navigate and a similar number felt that it is easy to locate desired information on Wikipedia while 86% felt Wikipedia's structure is easy to comprehend.

Perceived Usefulness: 82% of the respondents said that Wikipedia was accurate while only 51% felt it was complete. Thus, Wikipedia is not perceived as being an in-depth guide on any topic, but rather as an initial source of content for further investigation through other sources.

Motivation to Use: Compared to a hard copy encyclopedia, Wikipedia is both free and easy to access. Over 90% of the respondents indicated that these two were motivating factors for them to use Wikipedia. About (78%) indicated that the fact that many people have looked at the content and corrected mistakes, was another motivating factor, and

smaller proportion (51%) felt motivated by Wikipedia having survived since 2001. However, Wikipedia's being a non-profit organization was not an important motivating factor.

Inhibitors to Contribution: The survey asked respondents who perceived themselves to be seekers rather than contributors, if their inhibitors to contribution included inability to spare the time (61% agreed), perception of lack of knowledge (53% agreed) and lack of professional compulsion (48% agreed). The lack of personal benefit was not seen as an inhibitor to contribution. In fact, many respondents expressed that even though they feel like giving back, they hesitate because they feel they lack the required in-depth knowledge or they do not have the time and the disciplined commitment to participate.

These survey findings, though preliminary in nature, complement those summarized earlier about contributors and the factors that motive them. Taken together, we now have a more complete list of influencing forces that affect the two communities associated with a wiki, namely seekers and contributors, and are in a position to synthesize a holistic structure of interaction among these different factors, which ultimately drive the growth of a wiki.

Mechanics of Wiki Growth

We use the system dynamics methodology (Coyle 1998; Forrester 1961; Richardson 1996) to synthesize a holistic model of the forces identified in the previous sections. In doing so, we will capture the interaction among different factors that affect seekers and contributors, as well as the interaction between these two communities. System Dynamics (SD) is a mathematical language to represent the causal structure of a system and has been used extensively in numerous application domains (Coyle et al 1999; Woodside 2006). In our context, the 'system' consists of the Wiki technology platform, seekers and contributors of content. The distinctiveness of SD is that it connects causal structure to system behavior. In our context, system behavior would include such items as the patterns of page seeks, number of articles contributed, and number of edits to existing articles over time. If these patterns grow (or remain steady as appropriate) over time, the Wiki would be considered to be self sustaining. If the patterns decay, then of course the system behavior is one of atrophy. Figure 1 shows our SD model of Wiki growth. The basic constructs of the SD methodology are mentioned now. Additional details may be found in the references given at the start of this section. An SD model like Figure 1 is called a causal loop diagram (CLD). Links show cause-effect relationships and their polarities show the direction of effect. A positive (negative) polarity means that cause and effect change in the same (opposite) direction. A closed sequence of links forms a feedback loop. By following the polarities of links around a loop, loop polarity can be determined. An odd (even) number of negative links results in a negative (positive) feedback loop. A positive loop generates reinforcing forces leading to exponential behavior. A negative loop generates balancing forces leading to self-stabilizing behavior. So a CLD can be used to qualitatively reason about the behavior of systems by inspecting the interaction among its feedback loops. This holistic causal structure also reveals the macro-mechanism by which individual factors interact to result in different system behaviors.

Some of the variables that appear in Figure 1 will be seen to be derived from the survey findings summarized in the previous section, albeit at a coarser level of aggregation. Others, particularly those related to contributors, have been adapted from the related areas of literature reviewed earlier (Bryant et. al. 2005, Viegas et. al. 2004). Hence the model is one which synthesizes the individual factors identified previously, into an overall causal structure. We can begin the synthesis by first focusing on the creation and modification of articles themselves. The variable *Edits* in Figure 1 is intended to cover both the introduction of new articles and modifications made to existing articles. As noted earlier, the literature (Majchrzak et al 2006) identifies several different kinds of edits. However, to keep the model tractable, we aggregate these into one variable as a first approximation. More *Edits* generate more *Articles*, represented by the positive polarity link from the former to the latter. More *Articles* in turn generates more *Edits*, since each article is forwarded to peers who engage in verification, correction and extension. This cycle is a basic activity characterizing Wikipedia contributors. *Articles* generate *Wikipedia Page Views*. Some of these page views generate *Edits* for creation of new articles.

Our exploratory survey had indicated that information seekers were motivated by the perceived accuracy and completeness of Wikipedia content. We can think of a composite variable, which we name *Perceived Usefulness*, that is a conceptual composite of physical article properties such as size, hyperlinks and content references, and more intangible properties such as perceived accuracy and completeness. It is known that not all edits are beneficial (Wagner and Majcharzak 2007) and there will be instances of defacing and destructive edits. However, considering Wikipedia as a whole, a large proportion of edits are indeed beneficial. This justifies the positive link from *Edits* to

Perceived Usefulness. In the reverse direction, as *Perceived Usefulness* increases, the need for *Edits* would decrease since articles in the Wikipedia will be seen to be more complete and/or accurate. This explains the negative link from *Perceived Usefulness* to *Edits*. Although the currency of each article in Wikipedia will vary, it is safe to assume that taken as an aggregate collection, the content as a whole must be subject to obsolescence. We represent this through *PU Atrophy*, the deterioration of *Perceived Usefulness*. The negative loop between these two variables represents the simple process of deterioration.

Surrounding the mechanics of content editing just discussed is the mechanics of interaction between the Wikipedia content and content seekers/contributors (contributors are often referred to as Wikipedians). Factors moderating this interaction had been summarized in the previous section. Building on the causal structure identified above, the magnitude of *Edits* also depends on the number of *Wikipedians*, and their *Inclination to Contribute*. In this model we define *Wikipedians* to be the number of registered users. The pool of Wikipedia contributors actually includes a sizeable number of anonymous users, but that data is not readily available. In any case, for modeling purposes, this can be compensated somewhat by scaling up the number of *Wikipedians* suitably to get an estimate of the total pool. In the previous section, we noted how the literature found peer recognition to be a major motivator for Wiki contributors. This effect is represented by a small substructure starting with the positive link from *Wikipedia Page Views* to *Wikipedians Joining* followed by a positive loop between *Wikipedians Joining* and *Wikipedians*. The greater the number of new contributors joining, the greater the number of total contributors, which further increases the potential for peer recognition and attracts even more new contributors. Of course, as new contributors join, this reduces the number of remaining potential contributors and this effect is captured by the negative loop between *Wikipedians Joining* and *Potential Wikipedians*. While contributors are motivated by peer recognition, it is also known that if edit frequency is too high, motivation to contribute drops (Zhang and Zhu, 2006). The reason for this behavior is that excessive level of edits is often seen as an indicator that one's contribution is much more likely to be modified/deleted by others, hence reducing the attraction of contributing in the first place. In Figure 1, we show this effect by two alternate causal influence routes between *Edits per Article* and *Inclination to Contribute*. The preceding narrative presented the mechanics of contributor activity in the Wiki context, and was based on behavioral factors identified in the literature that were summarized in the previous section.

In our causal model, seeker activity manifests itself as Wikipedia Page Views. Based on findings from our exploratory survey reported earlier, this depends on the number of Articles, Trust in Wikipedia, Ease of Use and Ease of Access. As the repository of Articles expands, the site gets more page views. As Trust in Wikipedia goes up repeat page visits occur pushing up Wikipedia Page Views. Provided Perceived Usefulness is satisfactory, increased Wikipedia Page Views help to build more Trust in Wikipedia. Hence the positive polarity links between Trust in Wikipedia and Wikipedia Page Views. The latter also has a link from Perceived Usefulness. For a given level of page views, improved Perceived Usefulness increases trust.

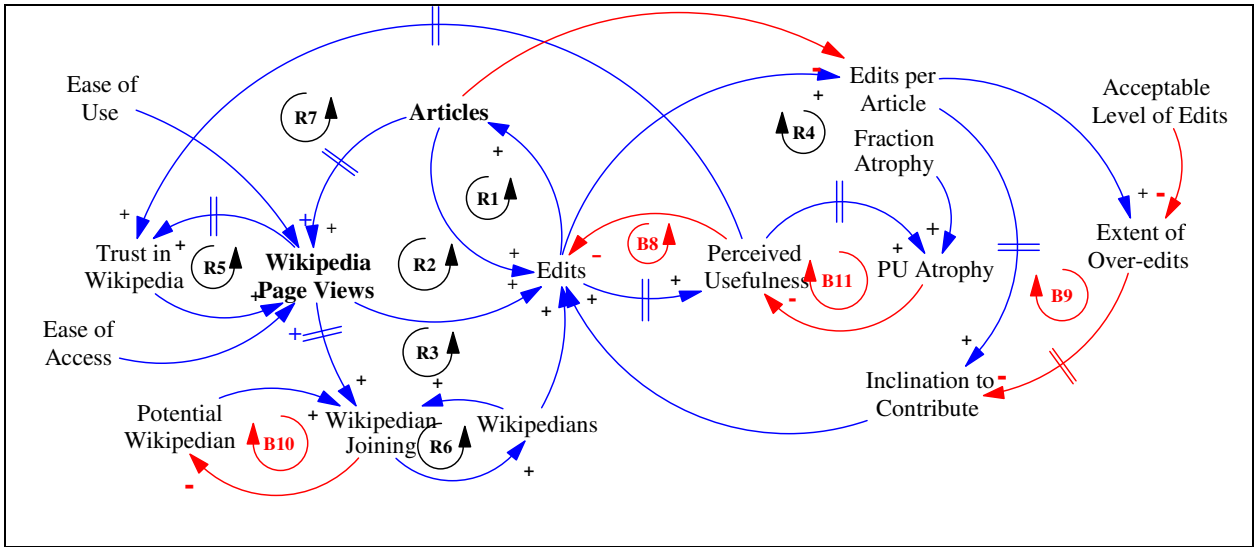

Figure 1: Seeker and Contributor Influences in Wikipedia Growth – Holistic Causal Structure

The preceding narrative describes the causal effects present in the holistic model shown in Figure 1. The structure is based on behavioral factors for contributors identified previously in the literature and behavioral factors influencing seekers identified in an exploratory survey we conducted for this work. The model is clearly coarse but, to the best of our knowledge, it is the first attempt to synthesize individual seeker and contributor factors into a more complete structure of the mechanics underlying the expansion or extinction of a Wiki application. We have identified individual causal links in the preceding narrative, but the macro building blocks in a system dynamics model are the feedback loops. Numerous loops can be identified in Figure 1, and the more significant ones are shown in Table 3. The first column shows the loop number as it appears in Figure 1, the second shows the variables constituting the loop, the third column indicates if the loop is positive or negative in polarity, the fourth column indicates whether or not the loop involves delayed effects, and the last column shows variables outside a loop that impacts its strength because of inbound links ending somewhere on the loop.

Table 3: Analysis of Feedback Loops in Holistic Model

Loop Id	**Loop Characteristics**			
	Variables	**Type of Loop**	**Delays**	**NonLoop Variables Contributing to Strength of Loop**
R_1	Articles, Edits	Reinforcing	None	Wikipedia Page Views, Wikipedians, Inclination to Contribute, Perceived Usefulness
R_2	Articles, Wikipedia Page Views, Edits	Reinforcing	One	Articles, Trust in Wikipedia, Wikipedians, Inclination to Contribute, Perceived Usefulness
R_3	Articles, Wikipedia Page Views, Wikipedians Joining, Wikipedians, Edits	Reinforcing	Two	Articles, Trust in Wikipedia, Potential Wikipedians, Inclination to Contribute, Perceived Usefulness
R_4	Edits, Edits per Article, Inclination to Contribute	Reinforcing	One	Articles, Extent of Over-edits, Wikipedians
R_5	Wikipedia Page Views, Trust in Wikipedia	Reinforcing	One	Perceived Usefulness, Articles, Ease of Use, Ease of Access.
R_6	Wikipedians Joining, Wikipedians	Reinforcing	None	Potential Wikipedians, Wikipedia Page Views
R_7	Trust in Wikipedia, Wikipedia Page Views, Edits, Perceived Usefulness	Reinforcing	One	Wikipedians, Inclination to Contribute, PU Atrophy
B_8	Edits, Perceived Usefulness	Balancing	One	Articles, Wikipedia Page Views, Wikipedians, Inclination to Contribute, PU Atrophy
B_9	Edits, Edits per Article, Extent of Over-edits, Inclination to Contribute	Balancing	One	Articles, Wikipedians, Acceptable Level of Edits
B_{10}	Potential Wikipedians, Wikipedians	Balancing	None	Wikipedia Page Views
B_{11}	Perceived Usefulness, PU Atrophy	Balancing	One	Fraction Atrophy

As noted at the start of this section, a reinforcing or positive loop would move the system away from stability while a balancing (negative) loop would pull it back towards stability. Increase/decrease in variables in the last column of Table 3 would accelerate/retard the effects of any loop. For example consider the reinforcing loop R_1. The strength of this reinforcing loop depends on *Wikipedia Page Views*, *Wikipedians, Inclination to Contribute, Perceived Usefulness*. Let's assume a situation where all these variables are held constant. Under that condition the value of *Articles* would grow at a certain constant rate. Now suppose the value of *Wikipedians* is increased. Since the link between *Wikipedian* and *Edits* is positive, the increase would accelerate the growth.

Qualitative Reasoning Using the Causal Structure

Based on the feedback loops present in the model above, it is possible to suggest qualitative explanations for growth of a wiki site. In the period immediately following the site launch, there would just be a few *Articles* and only a small number of promoters would constitute the *Wikipedian* pool. But the *Inclination for Contribution* of this pool would be high. Consequently the number of *Edits* and *Articles* would start growing being driven by R_1. Contemporaneously, with increased *Articles,* the number of *Wikipedia Page Views* would start growing (R_2). Also, loop R_6 would pull a few *Potential Wikipedians* into the pool of *Wikipedians*. Meanwhile with increased *Edits*, *Inclination to Contribute* would increase and that would push the rate of growth of *Articles* even further. Depending on the length of delays associated with some of the loops, other reinforcing loops would become active and contribute to the rapid growth. Collectively then, the feedback loops identified here would cause *Edits* and *Articles* to grow, and at an increasing rate at that.

The growth of *Articles* however would also be retarded by various balancing loops. Of these, the ones without any delay would act first - the only one being B_{10}. Next to follow would be B_8, B_9, and B_{11}. Recent statistics on the Wikipedia English site show a slowing down in growth of both *Wikipedians* and *Articles*. It does seem that loop B_{10} has started dominating the behavior for Wikipedia specifically. Although our model does not explicitly consider *Wikipedians* churn, actual Wikipedia statistics shows a section of the contributors departing. This does not augur well for the site. It needs a steady number of *Edits* to prevent *Perceived Usefulness* from dipping due to *PU Atrophy* (B_{11}). A lower level of *Perceived Usefulness* can erode the *Trust in Wikipedia* and discourage content seekers away from Wikipedia, leading to atrophy of the site (R_2 and R_5).

Erosion of *Edits* can also come about by a decline in *Inclination to Contribute* (B_9). It may be argued that this loop is self correcting and therefore has a stabilizing effect on *Edits*. However, the loop has a delay that can be reasonably assumed to be fairly long since changes in *Inclination to Contribute* take time to develop. Hence the effect of this loop would only be felt after a substantial delay. In the mean time, *Edits* is going to be reduced to a low level because of the other feedback effects discussed above. The preceding narrative shows how it is possible to suggest, simply through a CLD, mechanisms by which the influencing factors interact to cause certain types of growth patterns in a wiki site. However, qualitative reasoning only takes us so far in that we are able to reason about individual feedback loops and their effects. Humans are very poor at deducing the effects of interacting variables, let alone interacting feedback loops (Moxnes 2004). This is where the computational nature of SD models shows its strength.

Reproducing Observed Growth Patterns

In the system dynamics methodology, a Causal Loop Diagram can be converted into an equivalent stock-flow model, in which cause-effect relationships can then be quantified (Richardson 1996). The resulting stock-flow model is then simulated to computationally deduce system behavior resulting from the proposed causal model. Variables that represent accumulations over time are stocks, those that represent rates of change are the flows. We converted the model in Figure 1 to its equivalent stock-flow form and then implemented it using the Vensim® software package. As part of implementing any stock-flow model, the functional form of cause effect relationships must be specified quantitatively. Due to space constraints, it is not possible to show each and every relationship. However, the important ones are summarized below.

The variables *Articles*, *Trust*, *Wikipedia Page Views*, *Wikipedians* that appear in the CLD in Figure 1 now become the stocks in the corresponding stock-flow model. Note that *Wikipedia Page Views* is measured as the fraction of Internet Page views that visit Wikipedia, because that is the metric that is available from the public Internet traffic source www.alexa.com for calibration and validation.

The rate of change of *Articles* is proportional to *Edits*. Thus $\frac{d(Articles)}{dt} = \alpha(t) * Edits$, where $\alpha(t) = a + (1 - a) * \exp(-t/b))$. Note the time variant parameter $\alpha(t)$ declines with time to capture the fact that initially, the majority of *Edits* are for creating new articles. But as time progresses, more effort is spent in improving the quality of content.

$$Edits = \beta * Net\ Internet\ Page\ Visits * Wikipedia\ Page\ Views +$$

$$\text{Min}\ (Articles * \frac{Gap\ in\ Links\ per\ Article}{Links\ Added\ perEdit}, Wikipedian * Inclination\ to\ Contribute)$$

The first term in the expression represents Edits undertaken by Anonymous contributors. The number of such contributors is proportional to the section of the Page Visits ending up in Wikipedia. The second term in *Edits* represents contributions by *Wikipedians* to adjust the gap in *Perceived Usefulness*. Since the total number of such edits can not exceed the product of *Wikipedian* and *Inclination to Contribute* of an average Wikipedian, we take the minimum of these two. We define *Inclination to Contribute* as

*Inclination to Contribute = Average of Edits per Wikipedian * Extent of Over-edits*

As explained earlier, the term *Extent of Over-edits* captures the displeasure induced by over edits. We have captured this phenomenon with the following function:

$$Extent\ of\ Over\text{-}Edits = 1 - \frac{1}{(1 + \exp(-(Edits\ per Wikipedian / Average\ Edits\ per\ Wikipedian - 2)))}$$

The gradient of the *Extent of Over-Edits* implies a slow decline of the variable with small increment in the ratio of *Edits per Article* to *Average Edits per Article*, but much faster decline at higher value of the ratio. At even higher values of the ratio, the variable asymptotically approaches zero eliminating the *Inclination to Contribute*.

$$Trust\ in\ Wikipedia = P(\text{Repeat Visit by an Average Internet Surfer})$$

$$= P(\text{Satisfaction with content} \mid \text{Visit to a Wikipedia Site}) * P(\text{Visiting Wikipedia})$$

$$= Perceived\ Usefulness * Wikipedia\ Page\ Views$$

Wikipedia Page Views is an important state variable in the model and we model it as

$$\frac{d(Wikipedia\ \ Page\ \ Views)}{dt} = \gamma * \frac{1}{Trust}\frac{d(Trust)}{dt} + \rho * \frac{1}{Articles}\frac{d(Articles)}{dt}$$

In other words we assume that the rate of change of *Wikipedia Page Views* is a linear combination of the trends in *Trust* and *Articles*. A linear combination captures the mutual complementarity of *Trust* and *Articles,* both of which help to ensure viewership for a reference site.

Changes in *Wikipedian* happens through contagion based adoption that is given by:

$$\frac{d(Wikipedian)}{dt} = \mu * Wikipedia\ \ Page\ Views * Potential\ Wikipedian + \sigma * Potential\ Wikipedian * \frac{Wikipedian}{TotalPopulation}$$

Here μ represents the fraction of *Potential Wikipedians* becoming *Wikipedians* based on occasional page views. σ represents the strength of word-of-mouth. The model parameters were estimated using publicly available Wikipedia statistics and page view data about Wikipedia available from http://www.alexa.com .

The estimated values of the different parameters noted above are shown in Table 4.

Table 4. Stock-Flow Model Parameter Estimation

Parameter	a	b	β	γ	ρ	μ	σ	c
Estimate	0.19	18	0.70	0.15	0.002	0.005	0.14	3

The estimated model was simulated and a comparison of simulated and actual patterns for some key variables is shown in Figures 2-5. The actual patterns were again derived from publicly available data on www.alexa.com

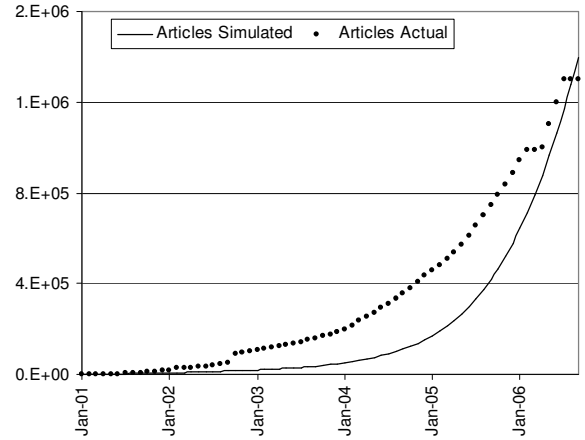

Figure 2. Number of Articles

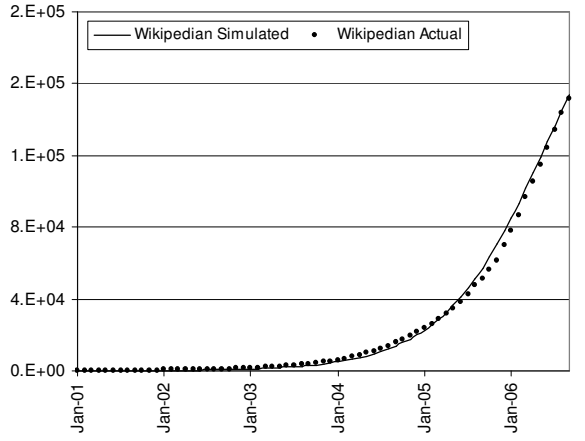

Figure 3. Number of Contributors

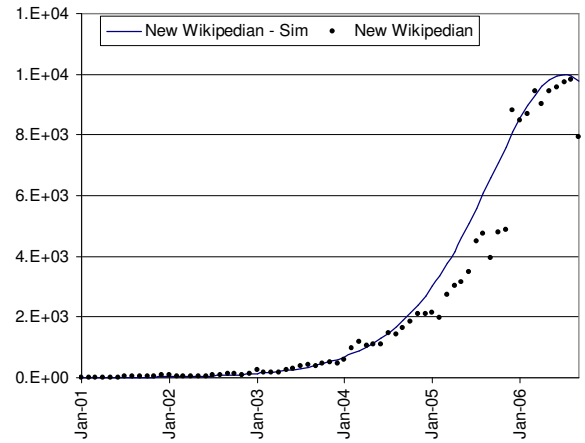

Figure 4. Number of New Contributors

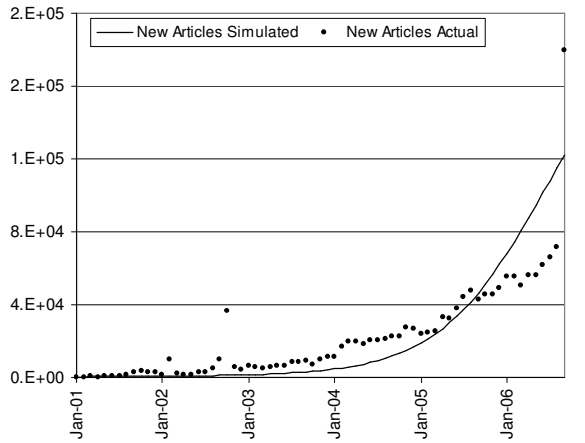

Figure 5. Number of New Articles

The first observation from Figures 2-5 is that the holistic model developed here is able to reproduce the functional shape of all four variables, these being *Articles, Contributors, New Contributors* and *New Articles*. The fact that the functional shapes of these key variables could be reproduced simultaneously is significant because it clearly suggests that the feedback structures constituting the holistic model are consistent with real causal mechanisms underlying this phenomenon. This is an important first step in developing a realistic causal model of Wikipedia growth. Of course, as the Figures also show, further refinement of model parameters is needed in order to improve the quantitative fit between observed and simulated patterns.

Predicting Patterns of Evolution

The preceding simulations represent a descriptive use of the holistic model in that we were trying to get the model to mimic observed behavior. Now that is has undergone an initial calibration, the model was subjected to an artificial 'shock' input to see what it predicts in terms of dynamic behavior for some of these key variables. The objective of this exercise is to demonstrate the model's potential use for predictive purposes and what-if scenario analysis. The scenario generated here corresponds to a situation in which many articles of poor quality are suddenly added to the Wikipedia. In the experiment, this sudden influx is timed to occur in September 2006. Figures 6-8 show the dynamic behavior of selected variables predicted by the model in response to this shock input. For comparison, the

corresponding base runs have also been shown. Base runs show the predicted dynamic behavior of these variables for the calibrated model in the absence of this shock. Notice that the simulation has been extended out to year 2009 for purposes of prediction. As would be expected, the number of *Articles* ramps up sharply in late 2006 since that is the shock that has been introduced. *Perceived Usefulness* takes a nosedive in late 2006 but then the drop steadies out after a while. *Wikipedia Page Views* also starts to decline although not immediately after the shock. The Wikipedia paradigm holds that at their inception, articles could in fact be poorly written or have mistakes in them. The process of collective editing is supposed to remedy that problem over time as Wikipedians make their individual contributions. What this hypothetical scenario shows is that it is actually possible for this public editorial process to get overloaded, resulting in a Wiki's atrophying away and losing readership. The value of having a causal model is that it is possible to reveal the mechanism by which this behavior is occurring. In the current scenario, loops R_7, B_8, B_9, B_{11} can be seen to act together to cause this atrophy. The large influx of poor quality articles causes a rapid drop in *Perceived Usefulness* which, after some delay, reduces readership. This has a snowball effect in that it reduces *Inclination to Contribute* as there are fewer people to recognize these efforts. Hence the improvement in article quality slows down causing further erosion in readership and so on. Once the underlying mechanism for the atrophy is identified it is possible to make policy interventions to minimize unwanted behavior or enhance desirable ones. For example, the rapid atrophy witnessed in this hypothetical scenario occurred essentially because of a mismatch between the speeds at which article quality improves due to *Edits,* and the speed at which *Perceived Usefulness* change in the minds of readers. A possible policy intervention to prevent such atrophy would be to limit the number of new articles that can be posted in any time period, analogous to the Securities and Exchange Commission shutting down trading in a stock if certain threshold of activity is crossed.

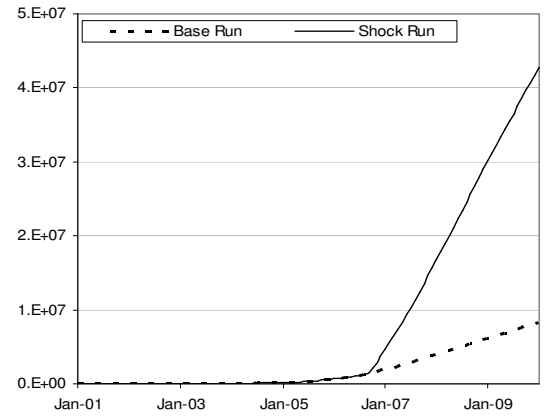

Figure 6. Predicted Number of Articles

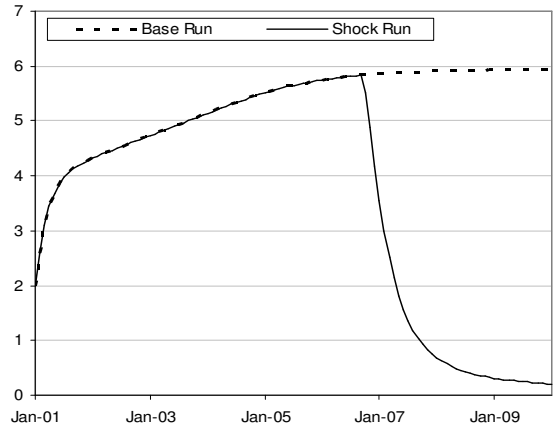

Figure 7. Predicted Perceived Usefulness

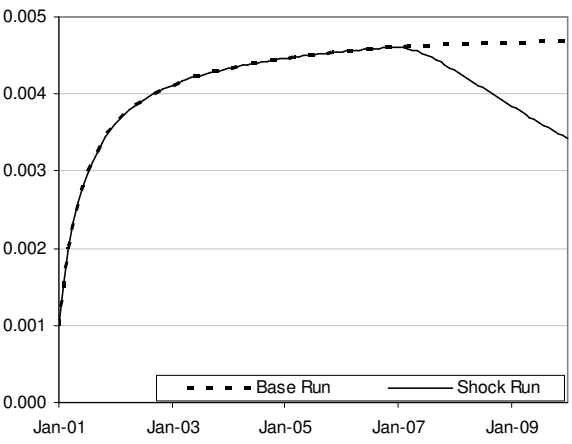

Figure 8. Predicted Page View

Of course, many other scenarios can be played out using the model. For instance, there has been talk of imposing additional controls on the editorial process in Wikipedia type settings, by introducing the concept of 'experts' (Lapp 2007). These experts would vet first time submissions and have special privileges to make changes to existing documents. The proposed change is controversial as it goes against one of the founding principles of Wikipedia, but is thought to be necessary to combat incompetence and/or malicious actions. The effects of such a move could be simulated after making slight modifications to the current model to incorporate the role of experts. In summary, the experimental results shown in this section illustrate how the holistic model of Wiki growth can be used for descriptive as well as prescriptive purposes.

Conclusion

Current literature has identified factors that influence the interaction between contributors of content and Wiki platforms in a variety of corporate applications. In this study, we have first complemented prior findings about content contributors with an exploratory survey of content seekers in the Wikipedia context. We next used the system dynamics methodology to develop a holistic model linking these different factors to patterns of Wikipedia growth. The model was tested by examining its ability to replicate observed growth patterns. The contribution of such a holistic model is that we now know not only factors affecting Wikipedia growth, but also the overall mechanics by which they do so. In other words, the model proposes "how" these factors might be leading to Wiki growth patterns. We have also demonstrated how the model can be used for descriptive as well as prescriptive purposes. There are, of course, several limitations in the current study, and avenues for further refinement.

First, the model has been deliberately kept at a coarse level of aggregation in order to have a parsimonious structure during this initial investigation. Having seen its ability to replicate observed dynamics of important variables in the Wikipedia context, the model can now be refined by using more disaggregated findings from the survey as well as from the literature. For instance, even in Wikipedia, articles are not homogeneous in terms of potential for controversy, volatility of content, interest to the readership or the degree of specialization. One cannot expect the dynamics of all these different types of content to be the same. There may be cultural differences among different Wikipedia communities that can affect the strength of specific causal relationships if not their functional form. The purpose for which a Wiki is established may also affect the strength of specific causal relationships in the model. Corporate Wikis established for a very narrow specific purpose, say the drafting of a new product release report, would probably have different dynamics than Wikipedia. Even in the Wikipedia environment, all contributors are not the same. There are registered users, but then there also are administrators and one has to be admitted to the latter group through a voting process. In other words, even the community of contributors is not homogeneous. So there are several avenues for refining the current model to make it more tailored to specific wiki contexts. In summary, the model developed here appears to be a reasonable starting point from which to better understand not just factors driving wiki growth, but the mechanics by which that happens.

References

Baroudi, J.J., Orlikowski, W.J., "A short form measure of user information satisfaction: A psychometric evaluation and notes on use", *Journal of Management Information Systems*, 4(4), 1988, 44-59.

Brown, S. A., Massey, A. P., Montoya-Weiss, M. M. and Burkman, J. R. (2002)., "Do I really have to? User Acceptance of Mandated Technology," *European Journal of Information Systems*, 11(4), 283-295.

Bryant, S., A. Forte & A. Bruckman. 2005. Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia, in *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*, Florida, USA.

Coyle R. "The practice of systems dynamics: milestones, lessons and ideas from 30 years of experience," *System Dynamics Review*, 14(4), (1998), 343-365.
Coyle J.M., Exelby D. and Holt J,.. "System dynamics in defence analysis: Some case studies," *The Journal of the Operational Research Society*. Oxford: Apr 1999.Vol.50, Iss. 4; pg. 372
Davis, F. D. (1986). A Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.
Davis, F. D., Bagozzi, R. P., Warshaw, P. R. "User Acceptance of Computer Technology: A Comparison of Two Theoretical Models" *Management Science*, 35(8), August 1989, 982-1003.
Du, H.S. and Wagner, C. Weblog Success: Exploring the Role of Technology, *International Journal of Human Computer Studies*, 64(9), 789-798, 2006.
Forrester J.W., *Industrial Dynamics*, Cambridge, MA:MIT Press (1961); Pegasus Communications, Waltham, MA.
Hof, R.D. Something Wiki This Way Comes: They're Web Sites Anyone Can Edit – and They Could Transform Corporate America, *Business Week*, June 7, 2004, p. 128.
Hu, P. J., Chau, P. Y. K., Sheng, O. R.L. and Tam, K.Y. (1999). Examining the Technology Acceptance Model Using Physician Acceptance of Telemedicine Technology. *Journal of MIS*, 16(2), 91-112.
Kang B. and Kim H. (1996), A demand–based model for forecasting innovation diffusion, *Computers and Industrial Engineering*, 1996, 30(3): 487–499.
Kittur, A., E. Chi, B. A. Pendleton, B. Suh, and T. Mytkowicz. Power of the few vs. wisdom of the crowd: Wikipedia and the rise of the bourgeoisie. *Annual ACM Conference on Human Factors in Computing Systems* (CHI 2007), April-May 2007.
Lapp A. Wikipedia's Opponent; Wikipedia vet Larry Sanger launches Citizendium, an online encyclopedia with tougher oversight, *PC Magazine*, 2007, 26,(11); p. 1
McKelvie G., Dotsika F., Patrick K., "Interactive business development, capturing business knowledge and practice: A case study," *The Learning Organization*, 14(5), 2007, 407-422.
Majchrzak A., Wagner C., Yates D., Corporate wiki users: results of a survey, *Proc. of the 2006 international symposium on Wikis*, 2006
Morse G., "Conversation: Jimmy Wales", *Harvard Business Review*, 86(4), Apr 2008, p. 26
Moxnes E. Misperceptions of basic dynamics: the case of renewable resource management, *System Dynamics Review*. Chichester: Summer 2004. 20(2); p. 139
Nov, O. (2007). What Motivates Wikipedians. *Communications of the ACM* 50 (11) 60-64.
Raman M. and Ryan T., Designing Online Discussion Support Systems for Academic Setting- "The Wiki Way", In *Proceedings of the Tenth America Conference On Information Systems*, New York, Aug. 4-8, 2004.
Richardson G. Modeling for Management. Aldershot UK: Dartmouth Publishing Co., 1996.
Rogers, E. M. (2003). Diffusion of Innovations (5th Ed.). New York: The Free Press.
Sanger L., (2005), The Early History of Nupedia and Wikipedia, in *Open Sources 2.0: The Continuing Evolution*, (DiBona et al eds.) O'Reilly.
Socialtext (2006), Dresdner Kleinwort Wasserstein, Case study available for download at www.socialtext.com/files/DrKW%20Case%20Study.pdf, last accessed: April 29th, 2008.
Taylor, S., Todd, P. A. "Understanding Information Technology Usage: A Test of Competing Models" *Information Systems Research*, 6(2), 1995,144-176.
Venkatesh, V. and Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, 46(2), 186-204.
Venkatesh, V., Morris, M. G., Davis, G. B., Davis, F. D. "User Acceptance Of Information Technology: Toward A Unified View", *MIS Quarterly*, 27(3), 2003, 425-478.
Viegas, F., M. Wattenberg & K. Dave. (2004), Studying cooperation and conflict between authors with history flow visualizations, *CHI 2004*.
Voss, J. (2005), Measuring Wikipedia. *Proceedings of 10th International Conference of the International Society for Scientometrics and Informetrics*, (Stockholm, Sweden), 2005
Wagner, Christian, and Prasarnphanich, P., Innovating Collaborative Content Creation: The Role of Altruism and Wiki Technology. *40th Annual Hawaii International Conference on System Sciences* (HICSS'07) . 2007. 18-27.
Wagner C. and Majchrzak A., Enabling Customer-Centricity Using Wikis and The Wiki Way, *Journal of Management Information Systems*, 2006-7, 23(3), pp. 17–43.
Woodside A.G., Advancing systems thinking and building microworlds in business and industrial marketing, *The Journal of Business & Industrial Marketing*. Santa Barbara: 2006. Vol. 21, Iss. 1; p. 24 (6 pages)
Zhang X. and Zhu F., Intrinsic Motivation of Open Content Contributors: The Case of Wikipedia, WISE-2006, http://digital.mit.edu/wise2006/papers/3A-1_wise2006.pdf.