Association for Information Systems
AIS Electronic Library (AISeL)

AMCIS 2007 Proceedings

Americas Conference on Information Systems (AMCIS)

December 2007

Building Successful Knowledge Repository Systems: Knowledge Refinement, Quality and Usage

T. Rachel Chung
University of Pittsburgh

BUILDING SUCCESSFUL KNOWLEDGE REPOSITORY SYSTEMS: KNOWLEDGE REFINEMENT, QUALITY AND USAGE

T. Rachel Chung
Katz Graduate School of Business
University of Pittsburgh
249 Mervis Hall
tchung@katz.pitt.edu

Abstract

Existing research on knowledge repositories has largely focused on supply-side driven research questions, such as employee motivation to contribute knowledge. This research turns attention to the demand-side issue of knowledge usage. Using the Information Systems (IS) success framework (DeLone and McLean 2003), this research theorizes knowledge quality as an important antecedent to knowledge usage. A conceptual framework delineates dimensions of knowledge quality from two perspectives. Moreover, an Input-Process-Output model of knowledge refinement process is proposed, and a causal model is offered to explain determinants of knowledge refinement effectiveness. Qualitative interviews and a survey study have been designed and are soon to be conducted with authors, validators and users of Eureka, a successful global knowledge repository system of the Xerox company. Implications for knowledge management research and managerial practices are discussed.

Keywords: Knowledge refinement, knowledge repository systems, quality control, knowledge management, IS Success

Introduction

Knowledge repositories are "databases allowing the storage and retrieval of explicit research and technical and management knowledge in text format" (King, Marks and McCoy 2002, p. 93). Existing research on repository systems has largely focused on supply-side driven research questions, analyzing issues such as employee motivation to contribute knowledge (Garud and Kumaraswamy 2005; Kankanhalli, Tan and Wei 2005). However, empirical evidence suggests that valuable content of repository systems is not always used as frequently as expected (Goodman and Darr 1998; Gray and Durcikova 2005). This research focuses on the demand-side issue of knowledge usage, as repository systems are successful only to the extent that their content is actively utilized by organizational members, either for replication (Dixon 2000; Markus 2001) or for innovation (Majchrzak, Cooper and Neece 2004). KM efforts to motivate knowledge sharing and contributions to repositories may be effective, but the KM process for which they support will not be successful if organization members are reluctant to access and apply the contributed knowledge to their own work.

Success of knowledge repository systems

We apply Delone and McLean's IS success framework (2003) to understand factors that make a repository system successful. A successful repository system should provide benefits to the individual users and enhance the organization's performance. Usage of the system and user satisfaction with the system positively predicts the benefits user will receive from the system. Three important determinants of system usage and user satisfaction are: information quality, system quality, and service quality (DeLone et al. 2003). Our approach focuses on the role of knowledge quality in driving usage behavior, after system quality and service quality are accounted for. This focus on knowledge quality is driven by the nature of repository systems. The content of a repository system usually consists of explicit knowledge, such as ideas, beliefs, descriptions, procedures, rationale, understanding, or insights that are formally articulated and codified in formats for digital preservation. As such, knowledge, as opposed to information, more appropriately characterizes the content of a knowledge repository system.

Knowledge quality and knowledge usage

Knowledge quality is defined here as the extent to which a knowledge object successfully serves the purposes of users (Kahn, Strong and Wang 2002). This definition is adapted from Kahn et al.'s (2002) definition of information quality (see also Wang and Strong 1996). We believe that a direct application of the information quality construct in the KM domain may not be conceptually appropriate, because knowledge is widely recognized as distinct from data and information. Whereas information derives from data embedded in context, knowledge captures meaningful interpretation of information that usually incorporates personal beliefs and values. This somewhat subjective nature of knowledge makes it difficult to establish objective methods of evaluating the quality of contributed knowledge. Content that is clear and useful for one knowledge worker may be seen as confusing and poorly written by another.

Following Kahn, Strong and Wang (2002), we consider two views of quality: conforming to specifications, and meeting or exceeding consumer expectations. Producers of knowledge, such as the authors and those that validate authors' contributions, usually take the former view, whereas users that consume the knowledge tend to take the latter. Each view proffers a number of dimensions, or knowledge quality attributes, that together represent the superordinate construct of knowledge quality (Wang et al. 1996):

> *H1: Knowledge quality consists of producer dimensions and user dimensions.*

From the producer's view, a unique aspect of knowledge quality is argument quality, the degree to which arguments presented in the codified knowledge are persuasive (Sussman and Siegal 2003). An argument is of high quality when the information used for the argument is complete, consistent, and accurate (Sussman et al. 2003). Another important aspect is causal ambiguity, the extent to which factors that cause the success or failure of an idea or practice can be determined with precision (Powell, Lovallo and Caringal 2006; Szulanski 1996; Szulanski, Cappetta and Hensen 2004). When the success of an idea is causally ambiguous, or when clear proof of the usefulness of the practice is difficult to obtain, knowledge is less accessible to potential users (Szulanski 1996). Other dimensions that characterize high knowledge quality include: free-of-error, defined as the extent to which knowledge is correct and reliable, concise representation, defined as the extent to which knowledge is compactly represented, and completeness, defined as the extent to which knowledge is not missing and is of sufficient breadth and depth for the task at hand (Kahn et al. 2002). Thus:

> *H1a: Producer dimensions of knowledge quality consist of argument quality, causal ambiguity, free-of-error, concise representation, and completeness.*

From the user's view, knowledge is of high quality when it is useful and usable (Kahn et al. 2002). The usefulness aspect involves appropriate amount of information, defined as the extent to which the volume of knowledge is appropriate for the task at hand, relevancy, defined as the extent to which knowledge is applicable and helpful for the task at hand, and understandability, defined as the extent to which knowledge is easily comprehended. The usability aspect involves believability, defined as the extent to which knowledge is regarded as true and credible, ease of manipulation, defined as the extent to which knowledge is easy to manipulate and apply to different tasks, reputation, defined as the extent to which knowledge is highly regarded in terms of its source or content, and value-added, defined as the extent to which knowledge is beneficial and provides advantages from its use (Kahn et al. 2002). Thus:

> *H1b: User dimensions of knowledge quality consist of appropriate amount of information, relevancy, understandability, believability, ease of manipulation, and value-added.*

Taken together, knowledge quality is conceptualized as a reflective aggregate construct with a constellation of dimensions from both the knowledge producer and the knowledge user's views. This conceptual framework is summarized in Figure 1.

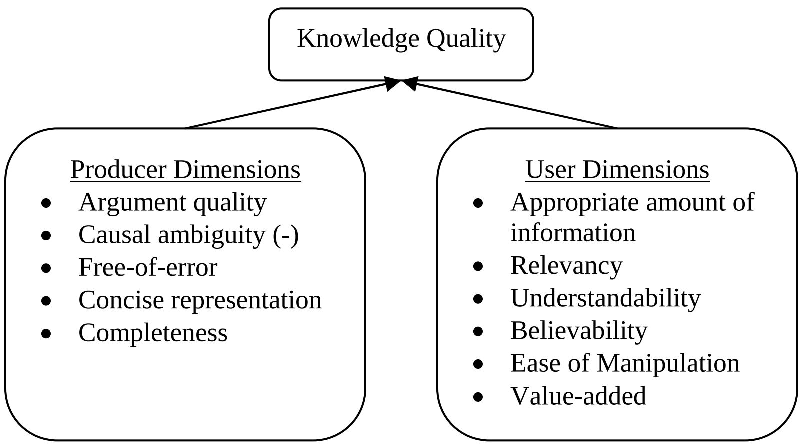

Figure 1. A conceptual framework of knowledge quality

The quality of a knowledge object should positively influence the extent to which the knowledge object gets used for several reasons. First, knowledge quality, particularly in terms of argument quality and argument usefulness, is an important predictor of the likelihood that a piece of advice gets accepted and eventually adopted by the advice recipient (Sussman et al. 2003). Moreover, when knowledge quality is high, recipients are more likely to modify their existing work routine and incorporate the new knowledge into their work practices (Kane, Argote and Levine 2005). On the other hand, low knowledge quality in terms of causal ambiguity increases the probability that the user's efforts to adopt the knowledge for a new task may fail, which in turn decreases the willingness of the potential user to adopt and apply the knowledge (Szulanski 1996). Based on this line of reasoning we hypothesize the following:

> *H2: Knowledge quality is positively associated with knowledge usage.*

A Refinement-Based Model of Knowledge Quality

Given that knowledge quality is critical for the success of repository systems, a conceptual model is much needed for understanding factors that contribute to knowledge quality. We develop such a model by focusing on knowledge refinement processes. Knowledge refinement refers to the process of evaluating, analyzing and optimizing the knowledge object to be stored in a repository (Alavi 2000; Cho, Chung, King and Schunn in press; Markus, Majchrzak and Gasser 2002; Qian and Bock 2005; Zack 1999)..

Many repository systems implement clear refinement processes. A common approach is to commission a centralized review committee of domain experts to select, refine and approve knowledge that enters repository systems (Goodman et al. 1998; Markus 2001; Zack 1999). At the other extreme is a decentralized system that assigns randomly chosen novice peers for refinement (Cho et al. in press) or an open-access system allowing anyone interested to participate in the refinement effort on a voluntary basis, as seen in the Wikipedia project[1] (Voss 2005). Somewhere in the middle is Xerox's refinement mechanism for their Eureka knowledge repository system: More than 1,100 validators, or refiners, dispersed globally are responsible for refining knowledge submissions from more than 23,000 users worldwide (Bobrow and Whalen 2002; Boucher 2006).

Universal to these different refinement approaches is joint participation of the refiner and the author in the knowledge creation process. The author externalizes tacit ideas into an explicit format, while the refiner helps the author optimize the quality of the explicated knowledge. Such a refinement process often involves the collaboration of the refiner and the author with varying degrees of interaction between the refiner and the authors.

Conceptualizing the refinement process using the input-process-output (I-P-O) framework of organizational teams and groups (Hackman 1987; McGrath 1984; Steiner 1972), we view the author's contribution as the input, the collaboration between the refiner and the author as the process, and the refined knowledge object as the output of the collaborative process. This view suggests that the goal of effective refinement is to produce knowledge objects of optimal quality (See Figure 2). Knowledge refinement effectiveness, based on this view, is defined as the degree to which the refinement process produces knowledge of desirable quality.

Figure 2. An Input-Process-Output Model of Knowledge Refinement Process

Modeling knowledge refinement as a collaborative process, we seek theoretical explanations of knowledge refinement effectiveness in the literature on collaborative projects. Here we focus on two particularly relevant constructs: shared understanding and arduous relationship. Shared understanding, defined as the extent to which the refiner and the author dyad share common work values, norms, and problem-solving approaches (Gerwin and Moffat 1997; Ko, Kirsch and King 2005; Nelson and Cooprider 1996) captures the cognitive harmony between the two collaborative partners. Arduous relationship, defined as an emotionally laborious and distant relationship (Ko et al. 2005; Szulanski 1996), captures the social harmony between the two.

There is compelling evidence that shared understanding improves knowledge refinement effectiveness. In a laboratory experiment, Cho et al. (in press) demonstrate that when the refiner and the author are more similar in expertise, the quality of the refined knowledge is rated much higher than when the refiner has significantly much more expertise than the author does. When the refiner and the author are similar in expertise, they are much more likely to share a common language when discussing and exchanging ideas. They are also more likely to understand each other's perspectives. In contrast, when the refiner is much more expert than the author, they can lack common vocabulary which is crucial for effective communication. They are also more likely to misunderstand each other which leads to frustration and hurts the team's performance (Nelson et al. 1996). Based on these ideas we propose the following hypothesis:

H3: Shared understanding is positively associated with knowledge refinement effectiveness.

On the other hand, arduous relationship may hurt the effectiveness of knowledge refinement. An arduous relationship exists when communication between partners is demanding, and the collaboration between the two presents a multitude of challenges (Ko et al. 2005; Szulanski 1996). The refinement process often involves frequent and multiple interactions between the refiner and the author (e.g., Cho et al. in press). When the organization's climate is collaborative and encourages members to actively facilitate colleagues' work, the repository system tends to demonstrate better knowledge quality (Qian et al. 2005). Conversely, the presence of an arduous relationship makes it difficult for the refiner and the author to collaborate effectively, because relationship conflicts have consistently demonstrated a strong and negative impact on team performance (De Dreu 2003). Since knowledge refinement effectiveness is essentially an index of the performance of the refiner-author team, we propose the following hypothesis:

H4: Arduous relationship is negatively associated with knowledge refinement effectiveness.

Theoretically, knowledge refinement effectiveness is conceptualized in terms of output knowledge quality, and so by definition knowledge refinement effectiveness should be positively associated with knowledge quality. Empirically there is

[1] According to Wikipedia's protection policy published at http://en.wikipedia.org/wiki/Wikipedia:Protection_policy, there are exceptions to this decentralized refinement mechanism. These are pages protected from editing by Wikipedia administrators, including those suffering from heavy and continuous vandalism, such as the article on George W. Bush.

evidence that, when the knowledge refinement process is effective, the repository system's knowledge quality is enhanced (Qian et al. 2005). Furthermore, when Cho et al. (in press) compare a peer-based refinement mechanism to an expert-driven one, they found the peer-based design to be more effective, producing knowledge objects that are of better quality. These empirical observations lead us to proffer the following hypothesis:

H5: Knowledge refinement effectiveness is positively associated with knowledge quality.

Figure 4 presents the integrated research model that incorporates all the hypotheses derived from theory development.

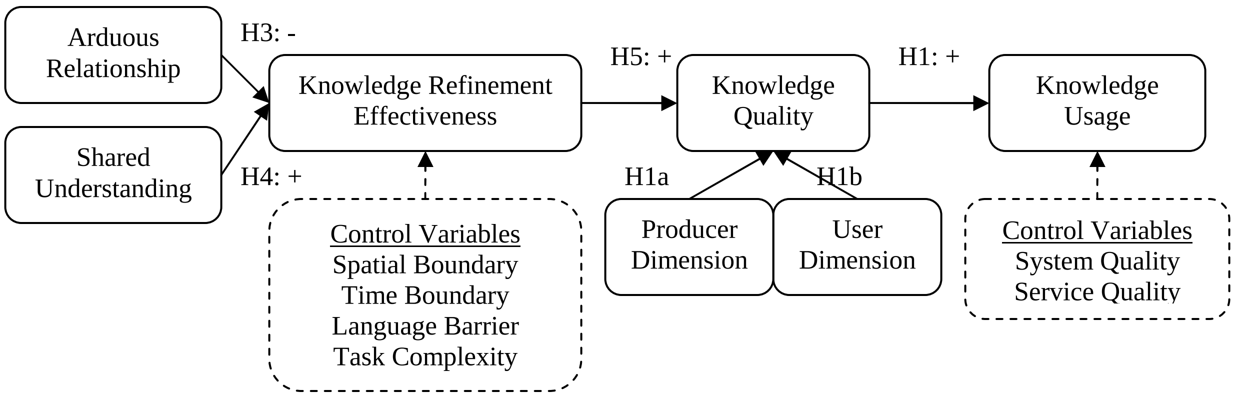

Figure 4. Integrated Research Model and Hypotheses

Method

The two studies described below is designed to evaluate this research model empirically using data to be collected from users of Eureka.

The Xerox Eureka system. Eureka is a knowledge repository system that supports knowledge sharing among Xerox service technicians (Bobrow et al. 2002; Boucher 2006; Hickins 1999) that enables the "globalization of local knowledge" (Von Krogh, Ichijo and Nonaka 2000). When a tip is submitted, the Eureka system coordinator assigns a validator to work with the author in developing and refining the tip. The validator may decide to reject a tip submission on a number of grounds: the tip is already included in Eureka, the tip is part of another tip already present in Eureka, or the tip is invalid. If the tip is indeed worthy of inclusion, the validator goes on to work with the author on refining the tip. The refinement process could include one or more revisions of the submission. The refinement process could last as short as a day and as long as 687 days (mean = 60.8 days; median = 21 days) for the 3485 tips refined and validated in year 2006.

Study 1 – Qualitative interviews. Since very little empirical data are available in the literature that sheds light on the refinement process, the first objective of this study is to conduct site visits and interviews with Eureka authors and validators about the refinement process.

Study 2 – Questionnaire survey. Validated measurement items will be adapted from existing literature for the present survey. The survey will be administered online and will include two versions. The producer version includes measurement items for knowledge quality – producer dimensions, shared understanding, arduous relationship, and knowledge refinement effectiveness. This version will also include measures for a set of control variables. Spatial boundary, defined as geographical differences between the refiner and the author (e.g., those in different cities), and temporal boundary, defined as the workday differences between the refiner and the author (e.g., those in different time zones), both significantly affect team member performance (Espinosa, Cummings and Pickering 2006). The control variable of process complexity is measured independently of the survey[2]. The user version includes measurement items for knowledge quality – user dimensions, and knowledge usage. This version also includes measures of two control variables: system quality and service quality, as they

[2] We control for the complexity of the refinement process by measuring the number of days that the refinement process lasted, and the number of revisions made. These data are available from system logs.

are known to impact knowledge usage (DeLone et al. 2003; Qian et al. 2005). Please see the measurement items in Appendix 1 at www.pitt.edu/~ting/ICIS07RefinementAppendix.pdf.

The survey study will consist of the following steps. First, a list of 500 validated tips will be randomly chosen from the pool of tips validated in 2006 from one or more Xerox product families. Second, authors and validators of these tips will be invited to fill out the producer version of the online questionnaire. For each tip we would try to receive responses from both the author and the refiner. Meanwhile, users will be invited to respond to the user version of the questionnaire with respect to the same tips.

References

Alavi, M. "Managing organizational knowledge," in: *Framing the Domains of IT Management: Projecting the Future.Through the Past,* R.W. Zumd (ed.), Pinnaflex Educational Resources, Cincinnati, Ohio, 2000.

Bobrow, D.G., and Whalen, J. "Community Knowledge Sharing in Practice: The Eureka Story," *Journal of the Society for Organizational Learning* (4:2) 2002.

Boucher, M. "Growing community knowledge through a socio-technical tip sharing system," Xerox, Montreal, Canada.

Cho, K., Chung, T.R., King, W.R., and Schunn, C. "Peer-based computer-supported knowledge refinement: An empirical investigation," *Communications of the ACM*) in press.

De Dreu, C.K.W. "Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis," *Journal of Applied Psychology* (88:4) 2003, pp 741-749.

DeLone, W.H., and McLean, E.R. "The DeLone and McLean model of information systems success: a ten-year update," *Journal Of Management Information Systems* (19:4), Spr 2003, pp 9-30.

Dixon, N. *Common Knowledge: How Companies Thrive by Sharing What They Know* Harvard Business School Press, Cambridge, MA, 2000.

Espinosa, J.A., Cummings, J.N., and Pickering, C. "Your time zone or mine? Geographic configurations, global team coordination, and project outcomes," in: *Academy of Management Conference*, Atlanta, Georgia, 2006.

Garud, R., and Kumaraswamy, A. "Vicious and virtuous circles in the management of knoweldge: The case of Infosys Technologies," *MIS Quarterly* (29:1) 2005, pp 9-33.

Gerwin, D., and Moffat, L. "Withdrawal of team autonomy during concurrent engineering," *Management Science* (43:9) 1997, pp 1275-1287.

Goodman, P.S., and Darr, E.D. "Computer-Aided Systems and Communities: Mechanisms for Organizational Learning in Distributed Environments," *MIS Quarterly* (22:4) 1998, pp 417-440.

Gray, P.H., and Durcikova, A. "The role of knowledge repositories in technical support environments: Speed versus learning in user performance," *Journal Of Management Information Systems* (22:3) 2005, pp 159-190.

Hackman, R. "The design of work teams," in: *Handbook of Organizational Behavior*, J. Lorsch (ed.), Prentice Hall., Englewood Cliffs, NJ, 1987.

Hickins, M. "Xerox Shares its Knowledge," *Management Review* (88:8) 1999, pp 40-45.

Kahn, B.K., Strong, D.M., and Wang, R.Y. "Information quality benchmarks: Product and service performance," *Communications of the ACM* (45:4) 2002, pp 184-192.

Kane, A.A., Argote, L., and Levine, J. "Knowledge transfer between groups via personnel rotation: effects of social identity and knowledge quality," *Organizational Behavior and Human Decision Processes* (96:1) 2005, pp 56-71.

Kankanhalli, A., Tan, B.C.Y., and Wei, K.K. "Contributing Knowledge to Electronic Knowledge Repositories: An Empirical Investigation," *MIS Quarterly* (29:Issue 1) 2005, pp 113-143.

King, W.R., Marks, P.V., and McCoy, S. "The most important issues in knowledge management," *Communications Of The Acm* (45:9), Sep 2002, pp 93-97.

Ko, D., Kirsch, L.J., and King, W.R. "Antecedents of knowledge transfer from consultants to cleints in enterprise system implementations," *MIS Quarterly* (29:1) 2005, pp 59-85.

Majchrzak, A., Cooper, L.P., and Neece, O.E. "Knowledge reuse for innovation," *Management Science* (50:2), Feb 2004, pp 174-188.

Markus, M.L. "Toward a theory of knowledge reuse: Types of knowledge reuse situations and factors in reuse success," *Journal Of Management Information Systems* (18:1), Sum 2001, pp 57-93.

Markus, M.L., Majchrzak, A., and Gasser, L. "A design theory for systems that support emergent knowledge processes," *Mis Quarterly* (26:3), Sep 2002, pp 179-212.

McGrath, J.E. *Groups: Interaction and Performance* Prentice Hall, Englewood Cliffs, NJ, 1984.

Nelson, K., and Cooprider, J. "The contribution of shared knowledge to IS group performance," *MIS Quarterly* (20:4) 1996, pp 409-429.

Powell, T.C., Lovallo, D., and Caringal, C. "Causal ambiguity, management perception, and firm performance," *Academy of Management Review* (31:1) 2006, pp 175-196.

Qian, Z., and Bock, G.W. "An empirical study on measuring the success of knowledge repository systems," Proceedings of Hawaii International Conference on System Sciences, Hawaii, 2005.

Steiner, I.D. *Group Process and Productivity* Academic, New York, NY, 1972.
Sussman, S.W., and Siegal, W.S. "Informational influence in organizations: An integrated approach to knowledge adoption," *Information Systems Research* (14:1), Mar 2003, pp 47-65.
Szulanski, G. "Exploring internal stickiness: Impediments to the transfer of best practice within the firm," *Strategic Management Journal* (17:Special Issue: Knowledge and the Firm) 1996, pp 27-43.
Szulanski, G., Cappetta, R., and Hensen, R. "When and how trustworthiness matters: Knowledge transfer and the moderating effect of causal ambiguity," *Organization Science* (15:5) 2004, pp 600-613.
Von Krogh, G., Ichijo, K., and Nonaka, I. *Enabling Knowledge Creation: How to Unlock the Mystery of Tacit Knowledge and Release the Power of Innovation* Oxford University Press, New York, NY, 2000.
Voss, J. "Measuring Wikipedia," Proceedings of 10th International Conference of the International Society for Scientometrics and Informetrics, Stockholm, Sweden, 2005.
Wang, R.Y., and Strong, D.M. "Beyond accuracy: What data quality means to data consumers," *Journal of Management Information Systems* (12:4) 1996, pp 5-34.
Zack, M.H. "Managing codified knowledge," *Sloan Management Review* (40:4), Sum 1999, pp 45-58.