A MODEL FOR INFORMATION QUALITY CHANGE
Completed paper

Besiki Stvilia
Florida State University
bstvilia@fsu.edu

Abstract: To manage information quality (IQ) effectively, one needs to know how IQ changes over time, what causes it to change, and whether the changes can be predicted. In this paper we analyze the structure of IQ change in Wikipedia, an open, collaborative general encyclopedia. We found several patterns in Wikipedia's IQ process trajectories and linked them to article types. Drawing on the results of our analysis, we develop a general model of IQ change that can be used for reasoning about IQ dynamics in many different settings, including traditional databases.

Key Words: Information Quality; Information Quality Dynamics, Activity Theory

1. INTRODUCTION

Information is an increasingly critical resource in our modern lives. The quality of outcomes of individual and institutional processes is often determined by the quality of the information that is used. Because of this link between information quality (IQ) and outcomes of decision making and actions, a theory and tools are needed that would allow effective and efficient management of IQ. A number of frameworks and models have been proposed for measuring IQ (e.g., [18]). To manage IQ, however, one also needs to know how IQ changes, what causes it to change, and how and when to intervene effectively. In this paper we develop a general model of IQ change based on an analysis of IQ process data in Wikipedia. Some statistics on the IQ dynamics of Wikipedia articles are also presented.

1.2 Background and Related Research

A substantial number of works have analyzed the problem of IQ measurement and have proposed sets of IQ measurement criteria (see [6], for a review). Research focusing specifically on IQ dynamics, however, has been scarce. Researchers have agreed that IQ is contextual. Studies have found repeatedly that moving information from one context to another changes how its quality is viewed and evaluated (e.g., [18]). [16] developed a basic control theoretic model of IQ dynamics and suggested a connection between information use and IQ. [1] proposed a decision theoretic model of information workflow process optimization consisting of 4 variables: timeliness, data quality, cost, and value. Several works have explored the efficacy of mining time series data for detection of failures and intrusions in telecommunication and banking systems (see [14, 17], for sample discussions). However, these studies have focused mainly on macro, aggregate patterns of quality dynamics and have not investigated the underlying micro IQ-related interactions: activities, roles, and strategies. Consequently, they might have overlooked or deemphasized some important variables and relations of the IQ ecology that could be exhibited only at the local levels.

IQ is often defined as the degree of usefulness of information or its "fitness for use" for a particular task or activity system [11, 23]. Information activities are complex webs of relationships among actions, roles, and tools as well as among activity-specific IQ norms or requirements, dimensions, metrics, problems, and the cost and value of IQ changes. Activity theory [11, 12, 22] allows one to conceptualize these relationships in a holistic, integrated, and systematic way. One of the main tenets of activity theory is the dialectical notion of tool mediation and evolution, which includes a process of continuous development

and learning through individual and collective feedback loops. As such, activity theory can provide a holistic theoretical structure for reasoning about IQ and IQ assurance work.

Successful theories and techniques for quality control have been used in manufacturing [7]. Frameworks and methods such as Total Quality Control, Six Sigma, or Statistical Process Control provide powerful philosophical principles for IQ assurance (e.g., continuous quality improvement, the process approach), but it is not clear whether these techniques would be directly applicable to modeling and controlling IQ. The difficulty of applying manufacturing quality control models and techniques to IQ can be found in the peculiar properties of information: its lack of physical properties, the context dependency–nonlinearity of information content (i.e., that information conveyed by the whole is not just the sum of its components), the lack of stability, and the nonrandomness of information errors [19].

In an earlier work, we began investigating a dynamic IQ model using an agent-based computational simulation [9]. This simulation modeled the process of a collection of agents differentially interacting with a large information base to accomplish tasks that were driven by the agents' strategic goals. The four types of agents—user, environment, malicious, and IQ assurance agents (IQAs) —both use and change individual information units and the relationships among them to execute tasks that achieve strategic goals. Our simulation suggested a nonlinearity of IQ dynamics as agents selectively improved or degraded information through the use of simple strategies. Using empirical data on IQ evaluation and content change in Wikipedia articles, we build on the previous research and develop a holistic model of IQ change, which is illustrated in the following sections.

2. RESEARCH DESIGN AND METHODOLOGY

This study used Wikipedia, a wiki-based, open encyclopedia, to identify the general sources of IQ variance and the patterns of IQ dynamics. Wikipedia is a general-purpose encyclopedia that not only allows anyone to edit its articles, but that also maintains and provides public access to the logs of some of its quality assurance processes (see [20], for a detailed analysis of Wikipedia's information processes). All these make Wikipedia an excellent environment for studying and analyzing IQ.

The data set for this study comprised edit histories and images of Featured Articles[1] (FA; n = 715), Former Featured Articles[2] (FFA; n = 375), and a random sample of 1,000 articles from the 2006/11/30 copy of the Wikipedia database. We used these data to generate a time series of monthly data points for the number of article edits, the number of article editors, and article length. Several studies have suggested that there may be a connection between the number of edits and the quality of Wikipedia articles (e.g., [21, 24]). One needs to remember, however, that although the number of edits can often serve as an indirect indicator of high quality, for controversial articles a high number of edits can also mean edit wars and vandalism. In addition, we looked at the logs of Featured Article Review (FAR) and Featured Article Removal Candidate[3] (FARC) process logs for the FFAs.

The research method used in this study consisted of a combination of (1) a conceptual modeling of Wikipedia information processes, (2) a grounded analysis of quality evaluation discussions, and (3) an analysis of time series data on article attributes in Wikipedia.

Activity theory provided us with a conceptual framework for reasoning systematically about the general context of IQ in Wikipedia: the hierarchical nature of goal-oriented activities, the integration points of different sociocultural aspects of the activity system, and the roles and dynamics of the quality and quality assurance infrastructure of Wikipedia articles as a whole.

---

[1] http://en.wikipedia.org/wiki/Wikipedia:Featured_articles
[2] http://en.wikipedia.org/wiki/Wikipedia:Former_featured_articles
[3] http://en.wikipedia.org/wiki/Wikipedia:Featured_article_review

The logs of quality discussions and votes provided valuable insight into context-specific quality assurance decision-making processes. The time series analysis of article attributes (the number of edits, the number of editors, and article length), on the other hand, helped to identify structural trends and patterns in the articles' IQ. Graphical data analysis techniques such as run sequence and autocorrelation plots were used to test the data for non-randomness and identify the trends [2, 3]. The results of the structural analysis informed and guided our analysis of local context-specific quality assurance practices and decision-making processes, and allowed us to identify sources of quality variance.

3. ANALYSIS

The activity theory framework of the human activity system (see Figure 1) allowed us to dissect and reason about Wikipedia's quality assurance work in a conceptually systematic way. It suggests where and how variance can be introduced in an article's IQ as well as the relationships and points of integration among those variances.

To be more precise, there are three levels of human activity:

1. Activity: Activity is the atomic unit, collective in nature and driven by a complex motive of which the individual actors are seldom aware.
2. Actions: Activity manifests itself in the form of goal-oriented individual actions in which the subject is consciously aware of what he or she is trying to accomplish.
3. Automatic operations: Actions in turn rely on automatic, routinized operations, dependent on the conditions at hand.
4. There are continuous two-way transformations between these levels: actions are internalized and become automatic operations through repeated practice, on the other hand, actions may also be expanded into novel collective activities [5].

Note, that the hierarchical definition of activity emphasizes the collective, socially distributed nature of work, which itself implies *division of labor*. In addition to the hierarchical structure of activity, activity theory has a number of other principles:

- Object-Orientedness – which states that activities occur in objective, socially and culturally shaped reality, and therefore have objective properties.
- Process of Internalization/Externalization – which states that there is constant transformation and feedback between internal (mental) and external (physical) activities and one cannot be understood without another.
- Tool Mediation – which emphasizes the social and developmental aspect of labor. Tools shape the way human agents interact with external reality, and, hence shape external activities. The constant feedback between external and internal activities, however, causes the transformation of internal activities as well. Hence, over time tool mediation leads to both cognitive and social development and learning. Tools become carriers of historically accumulated collective knowledge embedded in their structure and rules of use.
- Development – which states activity and activity systems can only be studied from the perspective of continuous historical development and evolution [13].

Thus, the activity theory framework of analysis helps not only in reasoning effectively and systematically about the social structure of information work, but it also can guide the identification and modeling of the structure of variance/variability present in the work, including IQ variance.

- Culture (Language …)
  - Community (Roles, Norms, Rules)
    - Activities
      - Actions & Goals
        » Operations
      - Artifacts (Objects or tools)
      - Agents

Figure 1: The structure of human activity system

In manufacturing, quality can be improved either by improving the production process (reducing the variance and moving the mean toward the target value through a better process), or by adopting stricter quality control of the ready products (reducing the variance and moving the mean toward the target value through scrap and rework) or increasing the robustness to parameter deviations [4]. For digital information products, the line between the production and maintenance processes is generally blurred. This also means that boundaries and attributes of digital information products can be transient. It is easier to modify and recycle a digital information product than a manufacturing product.

An examination of Wikipedia's activity system revealed two major IQ evaluation processes: the process of identifying high-quality or exemplary articles, and the process of identifying low-quality or irrelevant articles and removing them from the collection. Because the logs of deleted articles were not accessible to us, in this study we analyzed only the process of identifying high-quality articles and their characteristics. In particular, we looked at FA selection processes in this work.

FAs are considered to be Wikipedia's best. According to the FA policy, articles are promoted to FA status by the Wikipedia Featured Article Director after the community achieves a consensus that the article meets the FA criteria.[4] Another set of processes (FAR and FARC) is used to demote FAs that no longer meet the FA quality requirements. The FA criteria include both general quality dimensions that are grounded in cultural and social conventions for quality, and characteristics that are specific to the genre of the article and to the Wikipedia community [20]. FFAs can be renominated and regain FA status. It is expected, however, that the quality problems identified in the past FARC discussion will be addressed before the article is renominated. Wikipedia maintains a list of the FFAs, along with a list of the FFAs that have been renominated and reinstated into the FA collection. These lists, along with the archives of the related FAR and FARC discussions, provide valuable empirical data for identifying the sources of quality change.

In an earlier study we modeled the ecology of a large-scale open information collection by using a multi-agent simulation [9]. All four kinds of information agents modeled in that simulation were found in the Wikipedia context as well: (1) editors—agents that contribute or add new content to the article; (2) IQAs—agents that manage the article and collection quality; (3) malicious agents—agents that purposefully degrade article quality; (4) environmental agents—agents that change the IQ of articles through changes in real-world states. Although mostly degrading IQ, in a few instances changes produced by environmental agents can lead to better alignment of the article's IQ with the real-world state.

---

[4] http://en.wikipedia.org/wiki/Wikipedia:Featured_article_criteria

In the open Wiki environment, where the concept of article ownership may not apply and anyone can modify articles at any time, the formal distinction between production and maintenance actions disappears. It may remain in the editors' perception only and be based on how they view their own and each other's edits. As a result, Wikipedia articles, even at the FA state, can be treated as "works in progress."

Each edit action in Wikipedia can be accompanied by the following kinds of actions carried out by IQAs:

- *Identifying* or *locating* the contribution
- *Checking* the validity and quality of contributions
- *Achieving consensus* contributions through sense making, discussion, and negotiation
- *Editing* the contribution or the article to better integrate or align the contribution with existing content

Hence, for each contribution, IQAs may need to perform one to four actions, on average, per contribution in the quality control activity. An extreme case would be a complete reversal or discarding of the contribution. Because the underlying reality described by the article changes over time, the community may need to perform regular *maintenance* or *update* actions to align the article with either the changed underlying entity or the changed general context of article use. Finally, IQAs improve the process quality by building and maintaining its infrastructure (i.e., developing and maintaining policies and procedures; developing templates, guides, automatic maintenance tools, etc.), including editorial groups, by blocking vandals, resolving disputes, identifying qualified editors, and aligning editors with tasks.

Thus, IQ assurance actions comprise a significant portion of the community's overall information work. This is particularly true for mature articles, which have achieved relative completeness content-wise and for which the task of IQ maintenance increases in priority. In fact, although the editing processes of the FA and FFA articles exhibit similar centrality characteristics and variance, the community evaluates the quality of these sets differently, suggesting the presence of additional variance not captured by these measures. Both processes, however, are sharply different from those of the random sample (see Table 1).

Article type		Mean	Median	Std. dev.
Featured (715 articles, 23,744 time points)	No. of edits	30	7	70
	No. of editors	14	5	29
	Article length	19,481	15,912	16,202
Former featured (375 articles, 14,260 time points)	No. of edits	28	8	64
	No. of editors	15	5	31
	Article length	19,836	16,408	15,500
Random (1,000 articles, 6,012 time points)	No. of edits	3	1	8
	No. of editors	2	1	4
	Article length	3,442	1,764	5,407

Table 1: Centrality measures for the average monthly number of edits, number of editors, and article length

A content analysis of FAR and FARC discussions and votes for FFAs identified three main reasons for changes in their IQ evaluations and their loss of FA quality status (see Table 2). The analysis showed that 86% of FFAs were demoted because of continuously increasing FA quality requirements. The first consistent set of IQ criteria was developed in early 2004 and has been redefined several times since then [20]. The article trajectories reflected these changes by showing surges of editorial activity matching the timing of the criteria changes, and the FAR and FARC review processes triggered by those changes.

It is important to note that the ultimate goal of the FAR and FARC processes is to encourage the existing FA articles to evolve and improve in quality as Wikipedia grows and the supply of FA candidates

increases. Most of the time, an article loses FA status if the review process finds that it does not have the potential to be improved and meet the current quality requirements in a reasonable time frame due to either an inactive or a misaligned editorial group:

> This article was a featured in November 2004, but currently seems to be in a state of stagnation.

In some cases, editors might disagree with the community consensus about the criteria and simply refuse to make necessary changes:

> I am the only editor of this article, and as un-Wiki as it sounds, I wrote it (check the contributions). If I'm not here, it's going to end up out of date (it is already, as it happens). Secondly, I'm very, very annoyed about the requirement for inline citations. When it was made an FA, it wasn't required. Seems like they are now. Well, I know I'm not going to do that.

The second most significant cause of FAs losing their status was modifications of the articles themselves, which degraded their quality instead of improving it. These could be caused by malicious attacks and vandalism as well by unintentional degradation caused by incompetent or irrelevant edits:

> If this wasn't a FA, it would get less attention from well-meaning folks trying to "improve" it by adding a link to their favorite fractal gallery.

Finally, changes in the article's underlying entity too could lead to its demotion:

> With the introduction of the current S-197 Mustang, and the addition of information and models that was not included in the original FA . . . I think the article has been severely compromised.

Cause of FA status removal	No. of articles	%
Criteria change	285	86
Article change	47	14
Underlying entity change	3	1

Table 2: Causes of FA status removal (332 articles total; more than one reason for losing FA status could be applicable)

The continuous "work in progress" approach was reflected in the time series data on the number of edits and editors for the articles. Most of the sample articles had a non-zero number of monthly edits. In addition, the time series trajectories exhibited several interesting patterns, which too pointed to a connection between the dynamics of an article's quality and the changes in context or the life cycles of the entities they represented (see Table 3).

Like many other information objects, Wikipedia articles describe or are about different kinds of entities: *concepts*, *people*, *places*, *events*, or *things*. The analysis of FAR and FARC discussions showed changes in these entities that triggered changes in their representative objects—articles. Figure 2 shows how the periodicity of a reoccurring event such as a religious holiday may affect the number of edits an article about the holiday may receive. Both the number of edits and the number of editors for the Christmas article exhibit cyclic surges at Christmastime, even though the article length shows less cyclical regularity, suggesting that a substantial number of these edits were vandalism or irrelevant additions later reverted by IQAs.
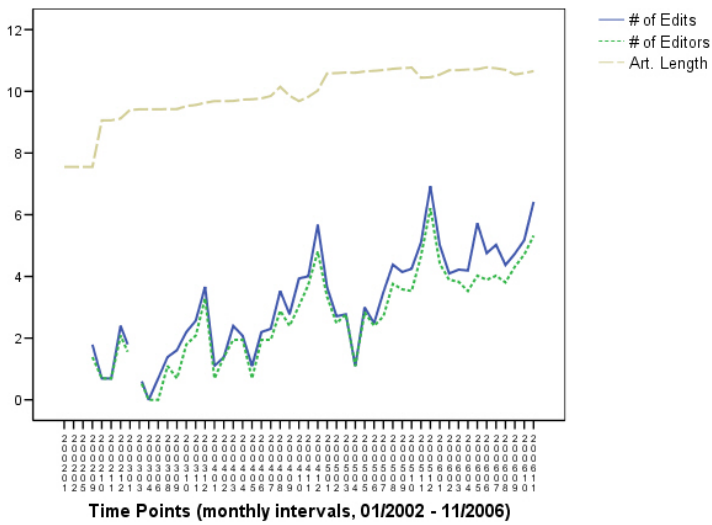
Figure 2: Christianity (natural log transformation was used for Y axis)

Article process trajectories about a nonrecurring event or a specific instance of a recurring event, on the other hand, may exhibit a downward trend as more time passes from the time of the event. Figure 3 shows how the monthly rates of edits and editors steadily declined after the event (2004 Democratic National Convention) had occurred.
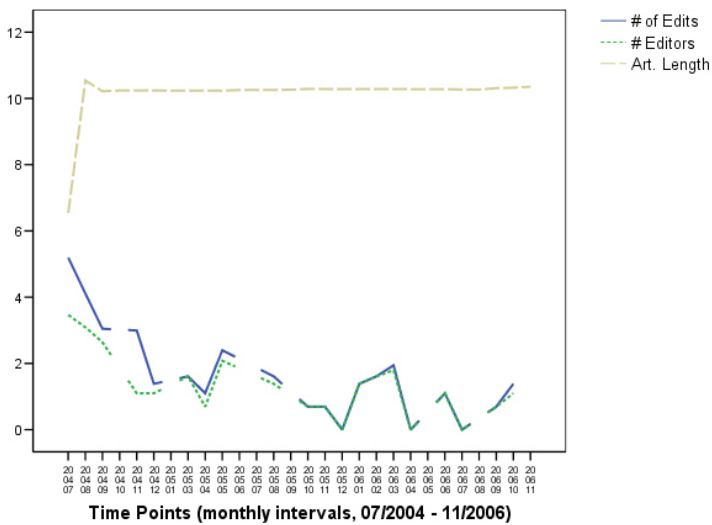

Figure 3: The 2004 Democratic National Convention (natural log transformation was used for Y axis)

The trajectories for articles about persons represent similar trends. Articles about persons who become increasingly famous and influential may exhibit an upward trend for the number of edits and editors (see Figure 4). Interestingly, one of the picks (October 2005) in the trajectory for an article on Nicolas Sarkozy coincides with the riots in the immigrant communities of Paris, which Sarkozy played a significant role in quelling as Interior Minister. Articles dedicated to persons who ended their career or passed away, on the other hand, may receive less attention from editors unless a person has become a cultural or political symbol whose life is celebrated as a regular event, as in the case of Martin Luther King.
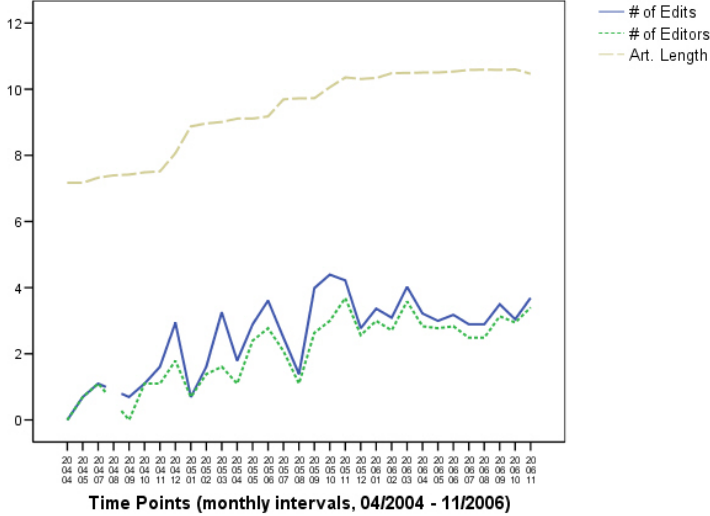
Figure 4: Nicolas Sarkozy (natural log transformation was used for Y axis)

Trajectories for articles about concepts, theories, and places did not represent any significant regularities, with the exception of spikes of activity related to FA or FARC processes. In general, religious concepts and theories appeared to attract editors and edits at a higher rate than scholarly concepts.

Trend	Concepts	People	Places	Events	Things
Upward	x	x	x	x	x
Downward	x	x	x	x	x
Cyclic				x	
Flat	x	x	x	x	x

Table 3: Process patterns for different kinds of articles

4. DISCUSSION

The analysis of Wikipedia process logs showed that changes in the IQ of an article were caused by *changes in the article*, *changes in its underlying entity*, and *changes in its activity system's context* (see Table 4). Note that the context could change both in time and in space:

> This is en.wikipedia.org (English-speaking world), not usa.wikipedia.org. The article should be re-written to include a WORLD-WIDE view, or it should be de-listed as a featured article.

Furthermore, the process of an information object's IQ change could be *passive* or *indirect*, caused by changes in the underlying entity and context—culture, sociotechnical structures, and domain knowledge. In general, these changes were not intended to affect the IQ of the object. In the case of Wikipedia, for instance, these changes could a particular editor leaving Wikipedia or an article's editorial group, changes in the FA criteria, or removal or modification of the articles that a given article references or is referred to by. The context could also be changed *actively* to affect the quality of the information object. New sources could be introduced or planted in or the existing ones modified with the intention of supporting or refuting the information presented by the object [8, 10]. A qualified editor(s) could be invited to help with improving the IQ of an article. There could be *active* or *direct* quality degradation through malicious

corruption or removal of the information entity (see Figure 5). Quality degradation actions may not necessarily be malicious, however. We observed in Wikipedia how administrators often had to remove edit access to an article (reduce its accessibility) to protect it from greater quality degradation caused by edit wars or frequent vandalism. Information also could be abridged (reduced in completeness) to meet the needs of certain audiences or uses or, alternatively, the conventions of a certain genre.

Context or Underlying Entity Change
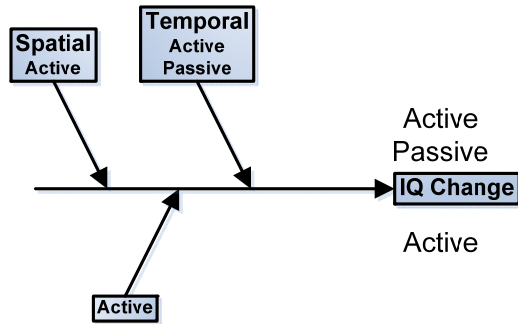
Information Entity Change

Figure 5: Sources of dynamic IQ problems

As mentioned earlier, quality in manufacturing can be improved either by reducing the process variance and improving its mean, or by imposing stricter quality control. In Wikipedia's context process improvement can mean formalizing the policies and procedures for article construction, standardizing the style and structure of articles, and better aligning editorial groups with article topics through better communication and selection. The stricter control of the final product would mean more frequent quality review and stricter enforcement of quality criteria. Interestingly, this study found that one of the ways Wikipedia improved the quality of its collection was to *continuously increase* the quality requirements for articles to remain in the collection. This not only reduced the collection's quality variance at the low end, but also increased its mean characteristics without actually changing the production process.

Agents	Activities or events	Knowledge and technology/tools	Community	Culture
Changes occur in editorial groups—existing editors leave or become inactive; new editors arrive who may not be aligned with the group	New activities are introduced that may generate new needs and uses for the information object. Alternatively, some of the existing activities in which the information object was used may become obsolete, making the related information needs obsolete as well. New events may occur that may affect the information object directly (e.g., initiation of a peer-reviewing or quality assessment process) or indirectly through its underlying entity (e.g. a country has elected a new president)	The current state of knowledge changes—what was considered to be accurate in the past may not be accurate now. New technologies are developed that may change the cost structure for activities, including quality assurance activities—activities that were prohibitively expensive in the past becomes affordable now	The community makeup as a whole changes—it can become smaller, larger, more aligned, or less aligned	The culture changes—what was admissible and aligned with the value system of the previous culture may not be admissible or interpreted in the same way in the current culture

Table 4: Sources of IQ change

4. CONCLUSION

In this study we analyzed time series data on the edit processes of FAs and FFAs. Although the time series data exhibited different trajectories for different articles, we observed a number of stable patterns in the trajectories. The patterns appeared to follow the life cycles of the underlying entities.

An analysis of FAR and FARC discussions on FFAs showed that IQ could be changed not only actively by editors, malicious agents, or IQAs editing the article, but also passively by changes in the article's underlying entity or the context of its evaluation and use. The IQ of the majority of FFAs had been reevaluated as lower, and these FFAs lost their high-quality status after the community decided to increase IQ requirements.

We believe that this study of the patterns and sources of IQ variance in Wikipedia can contribute to a better understanding of IQ dynamics, and that it has useful implications for optimizing IQ assurance in traditional databases. In particular, the activity theoretic model of IQ change and information type specific edit process patterns identified in this study can serve as a reusable knowledge resource for predicting IQ changes and guiding IQ maintenance actions and resource allocation. The model can also inform the design of software architecture and tools for automatic IQ assurance. Future work will include investigating the cost structure of IQ and linking it to IQ decision making.

REFERENCES

[1]   Ballou, D., Wang, R., Pazer, H., & Tayi, G. (1998). Modeling information manufacturing systems to determine information product quality. *Management Science, 44*(4), 462–484.

[2]   Chatfield, C. (1989). *The Analysis of Time Series: An Introduction*, Fourth Edition, Chapman & Hall, New York, NY.

[3]   Cleveland, William (1993), *Visualizing Data*, Hobart Press.

[4]   Cook, H. (1997). *Product management: Value, quality, cost, price, profits, and organization*. Amsterdam: Chapman & Hall.

[5]   Engestrom, Y. (1990). When is a tool? Multiple meanings of artifacts in human activity. In: Y. Engestrom (Ed.), *Learning, Working and Imagining.* (pp. 171-195). Helsinki, Finland: Orienta - Konsutit Oy.

[6]   Eppler, M. (2003). *Managing information quality: Increasing the value of information in knowledge-intensive products and processes*. Berlin, Germany: Springer-Verlag.

[7]   Evans, J., & Lindsay, W. (2005). *The management and control of quality* (6th ed.). Mason, OH: Thomson.

[8]   Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.

[9]   Gasser, L., & Stvilia, B. (2003). *Using multi-agent models to understand the quality of organizational information bases over time*. Paper presented at the NAACSOS Conference, Carnegie Mellon University, Pittsburgh, PA.

[10]   Gracy, D. (2002). What you get is not what you see: Forgery and the corruption of recordkeeping systems. In R. Cox & D. Wallace (Eds.), *Archives and the public good: Accountability and records in modern society* (pp. 247–264). Westport, CT: Quorum Books.

[11]   Juran, J. (1992). *Juran on quality by design*. New York: The Free Press.

[12]   Leontiev, A. (1978). *Activity, consciousness, personality*. Englewood Cliffs, NJ: Prentice-Hall.

[13]   Kaptelinin, V. (1996). Activity theory: implication for human - computer interaction. In: B. Nardi (Ed.), *Context and Consciousness: Activity Theory and Human Computer Interaction*. (pp. 103-116). Cambridge, MA: The MIT Press.

[14]   Milek, J., Reigrotzki, M., Bosch, H., & Block, F. (2001). Monitoring and data quality control of financial databases from a process control perspective. In *Proceedings of the International Conference on Information Quality—ICIQ 2001*. Cambridge, MA: MIT.

[15]   Nardi, B. (1996). Studying context: A comparison of activity theory, situated action models, and distributed cognition. In B. Nardi (Ed.), *Context and consciousness: Activity theory and human–computer interaction* (pp. 69–102). Cambridge, MA: MIT Press.

[16]   Orr, K. (1998). Data quality and systems theory. *Communications of the ACM, 41*(2), 66–71.

[17]   Pelletier, J., & Dasu, T. (2005). Mining network logs: Information quality challenges. In *Proceedings of the Tenth International Conference on Information Quality—ICIQ-2005* (pp. 327–339). Cambridge, MA: MIT.

[18]   Strong, D., Lee, Y., & Wang, R. (1997). Data quality in context. *Communications of the ACM, 40*(5), 103–110.

[19]   Stvilia, B. (2006). *Measuring information quality.* Unpublished doctoral thesis, University of Illinois at Urbana-Champaign.

[20]   Stvilia, B., Twidale, M. B., Gasser, L., & Smith, L. C. (2005a). Information quality in a community-based encyclopedia. In S. Hawamdeh (Ed.), *Knowledge management: Nurturing culture, innovation, and technology—Proceedings of the 2005 International Conference on Knowledge Management* (pp. 101–113). Charlotte, NC: World Scientific.

[21]   Stvilia, B., Twidale, M. B., Smith, L. C., & Gasser, L. (2005b). Assessing information quality of a community-based encyclopedia. In F. Naumann, M. Gertz, & S. Mednick (Eds.), *Proceedings of the International Conference on Information Quality—ICIQ 2005* (pp. 442–454). Cambridge, MA: MIT.

[22]   Vygotsky, L. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

[23]   Wang, R., & Strong, D. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems, 12*(4), 5–35.

[24]   Wilkinson, D., & Huberman, B. (2007). Assessing the value of cooperation in Wikipedia. *arXiv.org*. Retrieved March 8, 2007, from http://uk.arxiv.org/abs/cs.DL/0702140