Published:

Arazy O., Stroulia E., Ruecker S., Arias C., Fiorentino C., Ganev V., and Yau T., 2010, Recognizing Contributions in Wikis: Authorship Categories, Algorithms, and Visualizations, *Journal of the American Society for Information Science and Technology (JASIST)*, 61(6), pp. 1166-1179.

Recognizing Contributions in Wikis:

Authorship Categories, Algorithms, and Visualizations

Abstract

Wikis are designed to support collaborative editing, without focusing on individual contribution, such that it is not straightforward to determine who contributed to a specific page. However, as wikis are increasingly adopted in settings such as business, government and education, where editors are largely driven by career goals, there is a perceived need to modify wikis so that each editor's contributions are clearly presented. In this paper, we introduce an approach for assessing the contributions of wiki editors along several authorship categories, as well as a variety of information glyphs for visualizing this information. We report on three types of analysis: (a) assessing the accuracy of the algorithms, (b) estimating the understandability of the visualizations, and (c) exploring wiki editors' perceptions regarding the extent to which such an approach is likely to change their behavior. Our findings demonstrate that our proposed automated techniques can estimate fairly accurately the quantity of editors' contributions across various authorship categories, and that the visualizations we introduced can clearly convey this information to users. Moreover, our user study suggests that such tools are likely to change wiki editors' behavior. We discuss both the potential benefits and risks associated with solutions for estimating and visualizing wiki contributions.

Keywords: wiki, editors, contributions, authorship categories, algorithms, visualization

Introduction

Wiki, derived from the Hawaiian word for *fast*, is a web-based collaborative authoring application (Leuf and Cunningham, 2001; Wagner, 2004). In some ways, wikis are similar to discussion forums and blogs; both technologies enable communities to share information on the web. At the same time, there is a distinct difference between these technologies. In forums and blogs, users post distinct entries that are appended sequentially, according to their timestamp, to a single page. As a result, each individual post is directly associated with its contributor. In wikis, on the other hand, users can edit any part of the content of a wiki page; as a result, at any point in time, the most recent page version reflects the cumulative contributions of all users that have edited the page until then, without maintaining explicit references between content parts and editors' contributions on the article page (although such information could be derived from the log of wiki activity – the 'History' tab).

Wikis have already had a profound impact on the Internet at large, fundamentally changing the ways communities of practice collaborate online, with Wikipedia being the most prominent example (Nov, 2007; Fallis, 2008; Lim 2009;). In addition, recent evidence suggests that wikis are quickly penetrating into corporate settings (Majchrzak et al., 2006; Patterson et al., 2007; Arazy et al., 2009), government (Wagner et al., 2006), and education (Kane and Fichman, 2009). While wikis were designed for the kinds of openness and democracy that are typical of Internet culture (Stvilia, 2008), some of wikis' design features may not be suitable for organizational work processes. In this paper, we focus our attention on a specific design feature: authorship not being foregounded. Wikis were originally designed to hide the association between a wiki page and the editors who have produced it (Arazy and Stroulia, 2009). Two advantages are associated with this design, and explain why it is appropriate for Internet wikis. First, the lack of explicit authorship annotations around wiki pages eliminates the social biases associated with group deliberation, thus contributing to the diversity of opinions and to the collective intelligence of the group (Sunstein, 2006). Second, by not emphasizing authorship, editors are directed towards group goals, rather than individual benefits. Notwithstanding these advantages, this design is less suitable when users are motivated primarily by self-promotion and career-advancement and when traceability and provenance are of key importance. Thus, the main limitation of not highlighting authorship in organizational work processes is that it hampers accountability and reduces the motivations of wiki users to contribute content. For example, there is evidence

suggesting that users will contribute more when they know that their contribution is visible and is valued (Rashid et al., 2006).

While in some settings it would be sufficient to calculate and present editors' overall contribution, we believe that in many cases it would be important to distinguish between *the types of contributions* editors make. There are various ways in which editors can make contributions to a wiki page, e.g. one editor may add new content, another editor may re-organize the text, and a third may remove redundant text to make the page flow better. We believe that a method for estimating editors' contributions should capture these various authorship categories. Building on earlier works (Pfeil et al., 2008; Ehmann et al., 2008), we propose the following primary categories of wiki authorship tasks: (a) adding content, (b) enabling navigation within the page, (c) deleting content, (d) linking to other pages, and (e) proofreading.

Our study had two objectives. The first was to design and validate a set of algorithms for capturing wiki editors' contributions across the various authorship categories mentioned above, as well as an overall contribution score. It should be stressed that our focus is on the *quantity* of an editor's contributions; issues related to estimation of contributions' *quality* are beyond the scope of our study and are discussed at the conclusion of this paper. The second objective was to explore several visualization techniques for presenting information regarding editors' contributions on the wiki page, and to test the understandability of these visualizations, as well as the extent to which they are likely to impact wiki users' participation. We expect that if an estimate of editor contributions was presented on the wiki, wiki users would be encouraged to participate in the collaborative authoring process. Since the extent to which such an attribution tool impacts users' motivations and participation depends on the setting where the wiki is deployed, our user study focuses on two alternative wiki usage contexts. The first is a classroom setting, where students work together towards a group project. The second setting is a collaborative research environment. We expect that the impacts – both positive and negative – of an attribution tool on users' motivations for participation would be stronger in the classroom setting, where users see a potential impact on their evaluation (i.e. they suspect that the contribution score would affect, either directly or indirectly, their grade).

The remainder of the paper is organized as follows. Next we review related research; we proceed to describe our methods for estimating editors' contribution and for visualizing this information; the section that follows reports the results of the evaluation of the algorithms and visualizations; we then discuss the results, and conclude with our reflection on the findings, discussion of this study's limitations, and an outline of our plans for future work.

Related Work

Several recent studies have discussed wiki authorship tasks and proposed extensions to wikis that would automatically calculate users' contributions and associate a wiki page with its contributing editors. In the following sections, we review related works on wiki authorship tasks, algorithms for calculating editors' contributions, and visualizations of contribution data.

Wiki Authorship Tasks

Users can make various contributions to a wiki: add new content, remove redundant text, re-structure existing content, etc. There seems to be little theoretical guidance that would help in categorizing these various wiki editing types. While extant theory on professional editorial work describes editing strategies used by experienced editors (Bisaillon, 2007), the context of professional editorial work, where the roles of author and editor are well defined and quite distinct, is substantially different from wiki collaborative authoring, where a document is coauthored by its editors. Thus, existing theoretical frameworks of professional editing cannot be directly applied to the context of wiki authoring.

Recent studies, however, have worked in a more emergent manner to classify the wiki editorial tasks. Pfeil et al. (2008) employed a grounded theory approach to elicit a categorization of Wikipedia contribution types, which was later adopted by Ehmann et al. (2008). Their framework includes the following categories: add link, add information, format, spelling, style/typography, clarify information, fix link, vandalism, delete information, reversion, grammar, mark–up language, and delete link. While this detailed categorization is useful for understanding wiki work processes, it was not intended to guide the design of automatic attribution algorithms. For example, it would be difficult to automatically determine that a specific contribution entails "clarify information". Tools that wish to distinguish between various types of wiki contributions are likely to require a higher-level categorization. However, to the best of our knowledge, prior works on wiki attribution have not addressed this issue.

Estimating Users' Relative Contributions

Previous works on author contribution have proposed some preliminary methods which are based on metrics that are directly calculated by most wiki platforms: the total text created and the total number of edits performed.

Hess et al. (2006) propose a utility that calculates the extent of a user's revision by comparing the current version to the previous one, such that the overall contribution of a user is based on the sum of all her revisions. They propose that user statistics be used to influence participants' behavior and to trigger automatic actions (e.g. send alerts to administrators). Sabel (2007) uses a similar approach and presents a tool for calculating the differences between wiki page versions. He proposes that the difference be used as a 'rating' of the author's contribution, which then feeds into a reputation system, and her reputation could then serve as an estimate of her overall contribution to the wiki. Korfiatis et al. (2006) estimate a user's authority in the wiki based on social network analysis, creating a graph of authors, and calculating the centrality of each author node. They propose that authors' centrality could serve as a proxy for authors' reputation and could be used to help identify experts in an organization. Ding et al. (2007) used a simple count of the edits made to wiki pages. Hoisl et al. (2007) implemented an add-on that measures the relative contribution of authors and displays these user statistics. They calculate the importance of each wiki version (based on various metrics), and estimate the impact of a user's revision by comparing the importance of the current version to the importance of the previous one. The impact of an individual author is calculated by summing up the impact for each revision made by that author. Adler et al. (2008) propose the 'total edit longevity' metric that is based on how long an author's contributed text remains on the wiki before it is changed by others. They argue that the longevity metric is resistant to manipulations and demonstrate that the metric performs better than simple counts of total edits or words contributed.

These recent works provide some interesting solutions to the problem of wiki attribution. However, they suffer from several limitations. First, social network approaches (Korfiatis et al., 2006) are good at estimating the distribution of one's efforts across wiki pages, but do not provide an estimation of the extent of contributions to a specific wiki page. Second, some of the proposed algorithms can easily be manipulated by users seeking to boost their contribution score artificially. For example, by performing a long series of single character changes, someone could "game" the algorithms using a simple count of page edits (Ding et al., 2007). Third, most existing methods (Adler et al. 2008 is the exception) do not take into account the duration that author's contributions remained on the wiki page (Hess et al., 2006; Sabel, 2007). For example, if an author adds content that is deleted immediately, he still receives the same credit as if the contribution stayed on the wiki for months. Assuming that more relevant contributions are likely to persist (Anthony et al., 2009), algorithms should consider the duration a contribution stays on the wiki page. Fourth, it is not clear what the appropriate unit of analysis is for estimating the extent of one's

contribution. Many of the algorithms estimate a contribution based on the number of words (Hess et al., 2006; Sabel, 2007) or simply count the number of contributions (Ding et al., 2007). These metrics, we believe, may not necessarily distinguish between meaningful and trivial contributions, since their unit of analysis is not a semantic element. Lastly, existing approaches do not make the distinction between various types of contributions.

Visualizing Users' Relative Contributions

Considerable research has been conducted on visualizing the collaborative wiki process. Viégas et al. (2004) introduced history flow diagrams that provide a visual overview of editing history, clearly showing patterns of edits such as the tendency for wiki authors to add or delete material rather than reorganizing it, as well as revealing episodes of phenomena such as "edit wars." Ekstrand and Riedl (2009) extended this work and developed a tree-like graphical representation of the relationships among a Wiki page's revisions. This visualization is useful for making salient some of the key properties of article history, especially during times of high conflict. Holloway et al. (2007) developed visualizations that present the topical organization of Wikipedia. Ding et al. (2007) created a wiki visualization they call the Cherry Tree, which shows an author's contribution based on a simple count of edits; it also emphasizes the corporate directory structure. In a similar context of collaborative open source software development, Cheng and Kerr (2009) have been adapting bloom diagrams for the purpose of providing statistics on contributors' activity history and for visualizing these metrics, and have applied this technique to two open source projects: Sourceforge.net and IIOSB (the IBM Internal Open Source Bazaar). In summary, work remains to be done on visualizations that are intended to encourage collaborative authorship (as opposed to studying the history of authorship after the fact). One effort in this direction is WikiDashboard (Suh et al., 2008), a project by Palo Alto Research Center (PARC) that provides tools for supporting collaborative wiki work. WikiDashboard is an add-on to Wikipedia, which visualizes editors' activity patterns over time and is intended to increase transparency and accountability of Wikipedia's collaborative authoring. Although its impact on users has not been studied to date, preliminary evidence suggests that readers find it an effective tool. WikiDashboard uses a simple metric for estimating editors' contributions – number of edits – and does not distinguish between different authorship categories.

Our Proposed Approach for Calculating Wiki Editors' Contributions

In this study, we set out to address the limitations discussed above. We build on our preliminary work (Arazy and Stroulia, 2009) and propose a set of algorithms for calculating editors' contributions. Our proposed approach acknowledges that wiki editors can make contributions in various ways and that a method for estimating users' contributions should capture these various authorship categories. We build on earlier work on the categorization of wiki contribution types (Pfeil et al., 2008; Ehmann et al., 2008), and propose a simpler categorization that should lend itself to automatic estimation of authors' relative contribution. Our proposed categorization includes the following contribution types:

1. *Add*: adding complete new sections or changing existing information
2. *Improve Navigation*: helping users read the page by changing the structure and producing anchors (i.e. internal links)
3. *Delete*: deleting content
4. *Proofread*: making minor corrections and refinements to text and hyperlinks
5. *Link*: connecting to other web pages (either in the same wiki application) or outside

We take this framework as our starting point and introduce a set of algorithms that are intended to capture these various authorship categories, one algorithm per category.

The unit of analysis for each algorithm depends of the corresponding authorship category: algorithms that correspond to categories of content changes (i.e. 'Add' and 'Delete') employ the sentence as the fundamental unit of analysis; the 'Proofreading' algorithm is based on word-level changes, and the algorithms to analyze connections to other content (and correspond to the 'Improve Navigation' and 'Link' categories) use the hyperlink as the basic analysis unit.

Although estimating the quality of contributions is very difficult and is beyond the scope of this study, we do consider an indirect way of estimating contributions' relevancy. Calculating editors' relative contributions to a wiki page could be based on their cumulative work over the page's history, or alternatively on the portions of their work that remain on the most current version of the page. The latter approach assumes that as wiki pages evolve, low quality content is removed and high quality content persists. In this study, we explore both approaches. We apply these two approaches to the algorithms designed for each of the authorship categories, with the exception of

the 'Delete' category (for which the notion of 'content that persists' is irrelevant). Please see Table 1 for a summary of the proposed algorithms' key features. In the following sections we present our method and describe an empirical study that evaluated the extent to which the algorithms indeed capture the authorship categories they intend to.

Insert Table 1 here

Description of the Proposed Algorithms for each of the Authorship Categories

'Add' category. Our method for capturing the contributions of newly added content adopts a sentence as the natural conceptual unit, and employs user's sentence ownership as a metric for new content contributions. Focusing on the addition of "complete sentences" as opposed to characters is motivated by both our intent to focus on more conceptually meaningful edits and also to discourage the manipulation of the metric by users who want to artificially inflate their contribution score. We view a series of edits made by the same author as one continuous editing effort and we define a wiki page "release" as the last of these sequential revisions. Our algorithm calculates the sentence ownership of wiki page authors for each release. The algorithm cleans the data and determines sentence boundaries using the UIUC Cognitive Computation Group sentence-segmentation tool[1] Once sentence boundaries have been established for both the current and previous release, the algorithm identifies which of the sentences of the earlier release have remained essentially the same in the new release, building on the Munkres (1957) method to minimize the total distance between the paired sentences. If a sentence in the current release does not have a match in the earlier release, it is considered a newly added sentence with the author of the current release as the sentence owner. If, conversely, a matching sentence is found in the previous release, we distinguish between a minor change[2]and a major change[3]. For minor changes, if after the changes the original owner is still responsible for creating more than 50% of the words in the sentence, then he still retains ownership of that sentence. On the other hand, if after the changes the original owner is now responsible for creating less than 50% of the words in the new sentence, the original owner loses ownership and the sentence becomes public. For major changes, the contributor responsible for these changes becomes the new owner of the sentence.

The Munkres algorithm was designed to address the so-called "assignment problem", namely to identify an assignment of n jobs to n workers, such that the overall time to completion is minimized. It takes as input a n x n matrix C, where the value of $C(i,j)$ corresponds to the time it will take for the worker i to complete the job j. Let us consider the problem in terms of a bipartite graph $G=(S, T; E)$ with n worker vertices (S) and n job vertices (T),

where each edge *(E)* has a nonnegative cost *c(i,j)*. We are trying to find a perfect matching, i.e. assign each of the vertices in $s \in S$ to a distinct vertex in $t \in T$, with minimum overall cost, $\sum c(s,t)$. A function $y:(S \cup T) \rightarrow Q$ is called a *potential* if $y(i)+y(j) \leq c(i,j)$, for each $i \in S$, $j \in T$. The value of potential y is $\sum_{\upsilon \in (S \cup T)} y(\upsilon)$. It can be seen that the cost of each perfect matching is at least the value of each potential. The Munkres algorithm finds a perfect matching and a potential with equal cost/value, which proves the optimality of both. In fact, it finds a perfect matching of *tight* edges: an edge *ij* is called tight for a potential y if $y(i) + y(j) = c(i,j)$. In our case, the lists of sentences in the earlier and subsequent releases correspond to the jobs and the workers. Since the number of sentences in the two releases may not be equal, we use an extension to the original Munkres algorithm for rectangular matrices.

In order to study the effect of longevity, we calculate the amount of newly added content using three alternative approaches. First, we calculate the cumulative additions of an editor over all revisions. Second, we consider only the content that remained on the most recent page revision, assuming that high-quality content persists. The main limitation of this latter approach is that it fails to recognize contributions that persisted on the wiki for a long time if they were later overwritten. Thus, if a user provided the skeleton for an entire page, and over the course of collaboration his initial text was refined and enhanced, this instrumental user may not receive credit for his work. The third alternative is a compromise of the previous two approaches, and borrows from the ideas of Adler et al. (2008): it considers all additions made by an author and the *longevity of these contributions*, such that a user's score is based on the number of revisions that her sentences persisted on the wiki page (one credit for each sentence persisting on each 'release').

'Improve Navigation' category. One way to improve the logical flow is by organizing the structure of the page. However, such changes are very difficult to recognize automatically, and previous works do not provide any guidance (i.e. they focus only on newly added content). Adding anchors (i.e. internal hyperlinks) is another method for improving the flow and enabling easier navigation within a page, which is much easier to estimate. Thus, our method for estimating contributions that are intended to improve the logical flow of the wiki page uses as a proxy the number of internal hyperlinks introduced by the editor, under the assumption that these links establish an organization structure over the content of an individual page as well as the entire wiki content. Our unit of analysis here is the internal hyperlink, and in order to investigate the effect of longevity we provided two alternative algorithms: one based on the total number of internal links added and the other based on only the links that remain in the most recent wiki page version.

'Proofread' category. In order to capture contributions of proofreading, we counted the number of word-level changes (i.e. the unit of analysis is a word). As previously, we investigated the effect of longevity and looked at both the cumulative changes as well as the changes that persisted on the most current page version.

'Delete' category. We employed two alternative methods for estimating the scale of contributions through text deletion. First, using the notion of sentence as a unit of meaning, we calculated the number of sentences deleted by the user over all revisions. Second, we counted the number of revisions where the user deleted content, independent of the scope of deletion. For this category, the notion of longevity is irrelevant.

'Link' category. In order to capture contributions of adding links to other web pages, we counted the number of external hyperlinks introduced by a user (i.e. the unit of analysis is the external link). As earlier, we studied the effect of longevity by comparing two alternative algorithms: one based on the total number of external links added and the other based on only the links that remain in the most recent page version.

The set of algorithms described above were implemented on Annoki (Espiritu et al., 2006), our own extension of MediaWiki, as a service parallel, and complementary to the MediaWiki *history*. The MediaWiki history simply lists the revisions of a page, while our contribution-analysis service provides a summative overview of the nature of the page's evolution.

Method for Evaluating the Algorithms

In order to assess the accuracy with which the proposed algorithm captures contributions, we compared it against manual assessments. We analyzed nine random Wikipedia articles. The average article was evaluated for a period of 3.5 years starting from the time the article was created, over 130 distinct contributions, made by 105 unique authors. Please see Table 2 for details.

Insert Table 2 here

Two research assistants independently analyzed each edit made to each article, by reviewing the 'History' section of articles (where Wikipedia keeps a log of all changes to a page) and comparing subsequent versions. First, in order to get a score for editors' overall contribution to a page, the assessors provided their perceptions of the top 5 contributors, and ranked the editors in order of their contribution (the number one contributor received a score of 5,

the second a score of 4, and so on). Next, in order to assess the extent of authors' contributions for each of the authorship categories, the assessors examined each of the edits made to the article by recording the editor's name (or IP address) and classifying the contribution type into one or more of the categories previously mentioned: (a) Add, (b) Improve Navigation, (c) Delete, (d) Proofread, and (e) Link. The extent (i.e. quantity) of each contribution was then ranked on a 1-5 scale, from Minor to Major.

After completing the analysis of a page, we've summed up the contributions (accounting for the number and extent of contributions), giving each editor an overall page score for each of the contribution categories. All manual analyses were performed independently by the two assessors, who exhibited excellent inter-rater agreement (above 0.9 for all categories), and we used the average of the two assessors in our evaluation. The output for this process was a list of the editors for the set of Wikipedia articles, with a contribution score across the five authorship categories for each editor.

Results for Study Comparing Algorithms to Manual Assessments

For each of the authorship categories, we compared 2-3 versions of our algorithm, as described in Table 1. Below we present the results comparing the algorithms to the manual assessments.

'Add' category. We compared editors' manually-calculated scores to the three variations of our proposed sentence ownership algorithm. As a baseline, we've used Wikipedia's calculation for the number of characters added by each user. The results are presented in Table 3, and demonstrate that the four algorithms – baseline and our set of three sentence-ownership algorithms – exhibit statistically significant (at $p < 0.001$) correlations with manual evaluations. The best performing algorithm was based on the sentence ownership across all revisions, where the alternative approach that was based on sentences owned in the most recent page version performed slightly worse. Both these methods performed substantially better than the baseline. Interestingly, though, the approach based on the duration sentences persisted on the page performed much lower than the other algorithms (and worse than the baseline).

Insert Table 3 here

'Improve Navigation' category. When comparing the manual evaluations to our proposed internal hyperlink metric, we found that the correlation between the manually-produced value and our algorithms' output was rather high: the correlation for the metric calculating the total number of internal links was 0.566 ($p < 0.001$), while the correlation for the metric based on the links remaining on the current page version was 0.426 ($p < 0.001$).

'Proofread' category. When comparing the manual evaluations of proofreading changes to the output of our proposed algorithms, we found that the correlation between the manually-produced value and the metrics calculating the total number of word-level changes was 0.313 ($p < 0.001$), while the correlation for algorithm based on the changes that persisted on the current page version was 0.266 ($p < 0.001$).

'Delete' category. We compared editors' manual score to the two variations of our proposed 'delete' algorithm: (a) based on the total number of deleted sentences and (b) based on the number of revisions where the editor deleted content. As a baseline, we used Wikipedia's calculation for the number of characters deleted by each user. The results are presented in Table 4, and demonstrate that the three algorithms – baseline and our two deletion algorithms – exhibit statistically significant (at $p < 0.001$) correlations with manual evaluations. The best performing algorithm was based on the number of revisions where content was deleted (substantially better than the baseline), while the worst performance was for the algorithm based on the number of sentences deleted (substantially worse than the baseline).

Insert Table 4 here

'Link' category. When comparing the manual evaluations of connections made to external web pages to our proposed algorithms, we found that the correlation between the manually-produced value and the metric for the total number of added links was 0.391 ($p < 0.001$), while the correlation for the metric based on the links remaining on the current page version was 0.338 ($p < 0.001$).

In addition to the per-category analysis, we also wanted to test whether our proposed set of algorithms is able to produce an aggregated score of an editor's overall contribution to a page. To analyze the optimal combination, we performed a regression analysis, using the manually-identified 'top contributor' score, and the results are presented in Table 5 below. The total variance explained for the outcome variable (i.e. adjusted R^2) was 0.252. If we used the number of edits made by a user as a baseline metric of overall contribution, it would produce an adjusted R^2 of

0.230, below the score produced by our combination. The metrics with largest impact were internal hyperlinks (Beta = 0.37; $p < 0.001$) and revisions with deletions (Beta = 0.19; $p < 0.001$). A regression analysis with only these two variables produced an adjusted R^2 of 0.249. This result is surprising, as we anticipated the metric for newly added content (i.e. sentence ownership) to be the most salient, and did not expect the deletion of content to substantially influence assessors' perceptions of top contributors.

Insert Table 5 here

Visualizing the Contribution Scores of Wiki Editors

Our proposed visualization utilizes the information calculated by the various algorithms described above. We assume that we would be able to synthesize from these various algorithms a user's score for each of the authorship categories (this is validated in our first study). The visualization also assumes that an overall user score could be calculated by aggregating the contributions across the various categories.

The Proposed Visualization Approach

In a preliminary work (Ruecker et al., 2008), we introduced two glyphs that in various ways represented the relative contributions of authors across the different authorship categories. These glyphs are intended to reside on a wiki page, so that editors' contributions to that page are always visible. Here we extended this approach and introduced two additional glyphs. We were interested in finding out more about how people perceived these four glyphs, if there were a combination of glyphs that serve users best, or if there were one that was clearly superior.

The situated visibility of an editor's relative contribution to the wiki is an innovation that contradicts the principle of unattributed wiki work; we contend that in certain situations, such as the classroom or research settings, this modification is potentially beneficial in increasing motivation to contribute. Figure 1 shows the summary form of the glyph, which we have nicknamed "sunword." It provides the names of wiki editors around a central aggregate score for all contributors; the fontsize of the names corresponds to their relative scores, which are also included as a numerical value.

Insert Figure 1 here

The detailed form of the glyph (Figure 2) allows the user to see the components of the score that is summarized in sunword. The detailed glyph combines pie-shaped sections for each editor that show the relative contribution across each of the five authorship categories, indicated by five concentric circles that make up the wedge. When the name of a specific editor is selected ("wikidiver" in the figure), the details for that editor are highlighted.

Insert Figure 2 here

Our third glyph (Figure 3) was an attempt to provide the user with a sense of editors' contribution not only to the current wiki page, but also to the larger wiki site (assuming the site consists of many pages). Using two overlapping circles, one for the current page and one for the total wiki application, we display the same editor in both contexts. We display an editor's contribution score, as well as a percentage of all editors' contributions. The continuity of the contribution across the two different contexts is provided by using both a band of color and the numeric scores set in the same relative position on both circles. The two circles are also redundantly labeled, which helps make their relationship clear at first glance, although logically they can only occur in one order (except in the boundary condition where the current page is the only page in the whole wiki application).

Insert Figure 3 here

Based on the results of a user study (the results are presented in the following sections), we have implemented a detailed version of the information glyph (Figure 4) within MediaWiki. This new glyph provides information similar to what we showed in Figure 2, except all authors are differentiated with color. The glyph is also more interactive, allowing the user to turn the section labels off and on, although with the labels off, the glyph responds to rollover by swelling the appropriate section of the rings and temporarily displaying the label.

Insert Figure 4 here

In principle, the glyphs can take as an input the metrics calculated by any algorithm. In our implementation – for each authorship category – the glyphs take the calculations of the best performing algorithm. For example, for

the ‘Add’ category, the visualizations use cumulative additions (rather than the longevity variation). These glyphs were implemented on our MediaWiki extension – Annoki – on the *history* tab.

The Method for Evaluating our Visualization

In order to evaluate the clarity of our proposed visualization, we set out to address two primary sets of research questions. First, to what extent can users understand the information visualized in the glyphs? Are there specific variations of the visualizations that users can more easily understand? What level of detail do users require? Second, how would visualizations of the attribution metrics impact users' attitudes towards contributing to the wiki? What is the expected impact of the visualization on users’ wiki behavior? Are there differences between the various glyphs?

We were interested in examining the possible effects of our information glyphs in two different settings and seeing if there were any differences between the settings. The first setting consisted of students who had been asked to contribute to a wiki as part of their course work. These people would be explicitly evaluated on their contribution through a grade for the assignment, and would also be subject to peer pressure through their participation in a group project. The second type of wiki editors consisted of collaborative researchers who were using a wiki as part of their research activity. In this case, there would be no explicit evaluation; instead, the presumed motivation would be to act appropriately within the context of the dynamics of the research group.

We therefore conducted a user study involving semi-structured interviews with ten participants: five students and five researchers. Each of the participants was shown the full set of information glyphs, embedded separately on wiki pages to provide a context that would be similar to how they may appear in use, although for the purposes of our study the pages were printed. Participants were allowed to examine all the pages at once, in order to get a perspective that would allow comparative comments, but the interviewer drew attention to one at a time while asking the questions.

Each participant was provided with one of two scenarios, depending on their previous experience as students working on a collaborative writing assignment or researchers using a wiki.

Scenario 1: Classroom. “As part of your assignments in your undergrad course_____(a real course they’ve taken) you are asked to work collaboratively with your classmates in a wiki project. The project consisted in writing an article of quality about _____ (specific interest of the participant). “

Scenario 2: Research team. "As a member of a collaborative research team, you have been asked to contribute to the wiki project. As a researcher, you won't be graded, but you hope that your work on the project will eventually lead to conference presentations and publication in peer-reviewed journals."

The interview questions fell into two categories. The first set dealt primarily with issues of comprehension and explanation: what were the participants seeing in the different kinds of visualizations, and how did they interpret what they saw? Participants were asked to: (a) try to explain the data displayed, (b) describe the strengths of each visualization, (c) explain how they would identify specific data they are interested in (e.g. their contributions), and (d) try to recognize the relative contribution of each editor. In order to assess the expected impact of visualizing contribution data, we performed a second component of the user study, using the same subject pool described above. In this case, our questions focused on the possible effects such visualizations might have on the motivation of students in a class or researchers on a collaborative team to contribute to a wiki. Participants were asked to: (a) describe how they would feel if a graphic showing contributions were to be implemented as an add-on in a wiki technology, (b) discuss how having a visualization tool inserted in a wiki would change their wiki participation, and (c) describe what their strategy for becoming the top ranked editor would be. We kept an audio recording of the sessions and also took notes of the answers, then carried out a qualitative analysis that consisted of extracting the statements about each glyph that related to our questions. Since our participant pool was small, we did not concern ourselves with frequency of statements, but instead considered as potentially informative each individual utterance by each participant.

Results on the Understandability of the Visualizations

Our first research question for the information glyphs dealt with how well our participants were able to interpret what they were seeing: "to what extent can users understand the information visualized in the glyphs?" Our user studies indicate that by and large our glyphs were understood by the users. Perhaps not too surprisingly, from the perspective of the two different settings (classroom and collaborative research), there were no differences in their ability to understand the visualizations. We have therefore merged the two sets of comments in the following results.

Responses to the summary glyph (Figure 1) included various indications that our participants found it to be easy to understand that they were seeing a visualization of relative contribution, e.g.:

> *"…the names are clear as to who's contributing the most."*

"you can see exactly at a glance the amount of contributions"

In addition, some participants noted the comparative aspect of the visualization, i.e. including all editors in a single visualization provided a context for understanding the contributions of a specific editor. In fact, the intention of the glyph is to provide scores for only the top 25 editors, but there was no way for participants to recognize that there was a limit. Some example statements include:

"…it just doesn't isolate you as a user…is more specific, more easy to locate in context. It helps me understanding the whole spectrum of users."

"[I see]…the relative rank of everyone and the total number of contributions"

There was, however, some confusion about the precise meaning of the numbers, e.g.:

"And this little number on top… at first sight I don't know what it is, and… because it doesn't correspond to the percentage I think…yeah because Alex has three percent of contributions and four something… and wikidiver has 11, it doesn't make sense that this percentages are relative to these numbers, I don't know what it is."

The detailed glyph (Figure 2) showed two states, one with only the details about a particular wiki contributor, and the other with details about all contributors with one highlighted. Responses suggested that participants readily identified the breakdown into more details.

"…it's nice to have it broken down like that."

"Provides a really good breakdown of the type of contributions made."

As with the summary glyph, participants also readily identified that information was being shown of a comparative nature.

"…it shows not only one user but also different users. So… it presents more data and that you can compare different user's contributions…"

"…it's great because you can see your relative percentage…"

However, as with the summary glyph, guesses about the actual nature of the details were not correct.

"And hmm, I don't really understand like what the numbers "2" and "11%" mean. Like… 11% of what? I don't know what that is, and "2" I don't know what that is either…"

The third glyph (Figure 3) showed contributions by a single author for the current wiki page against a total contribution to the entire project wiki. Many of the participants were able to correctly describe in general terms what they were seeing.

> *"…it does only the other ones don't do, which is compares to all of the wiki pages."*
>
> *"…your involvement in the whole wiki…"*

However, here too some confusion was expressed, e.g.:

> *"I don't know why they are superimposed, I don't know if that's supposed to have a meaning or something."*

Results on the Expected Impact of the Visualizations

Our second research question related to the potential impact of the information glyphs: "How would visualizations of the attribution metrics impact users' attitudes towards contributing to the wiki? What is the expected impact of the visualization on users' wiki behavior?" Overall, the visualizations contribute to a sense of competitiveness, to which some subjects responded positively, while others did not think this is appropriate in the wiki context. While some participants recognized that the glyphs were not necessarily adding a radical new kind of information:

> *"you know, you can always tell in a group, it's not like you don't know! so… this just makes it easier to track I think."*

Many others recognized the potentially competitive nature of this aspect of the visualizations, as the first two visualizations ("sunword" and detailed) contained information that could be used for ranking oneself against others. The responses to this competitive pressure differed across settings.

Classroom Setting. For the students, the visual presence of a score attached to the information glyphs seemed to resonate with a number of existing concerns about being under the competitive microscope, while others responded by indicating that they felt competition can be motivating, and thought that they would take into consideration what the other members of their group thought about their contributions. Several of the student participants mentioned that if their peers had access to the information in the visualization, they would try to work harder on the quality of their edits.

> *"I think when people can see themselves stacked up against other people then they're going to try harder right away."*
>
> *"As far as how it would improve the information they were inputting, I think it makes a little bit of competition and it's always good, just to see who is doing more effort, just to try to add a little bit more. And I think it creates a sense of pride in your work, just to see how much effort you put into it and how much it matters in the end."*

Others, however, expressed reservations about the potential effects of the visualization on the quality of the resulting wiki content, and some felt that if an explicit connection were made between a wiki author's ranking and grade, the results could be negative

> *"…try not to abuse the system just to get on top, people may try just to put more data just to get more ranking."*
>
> *"… I also think that it would become an endless competition. Um… because, people are going to assume that you are graded on what part or size of the chart you have, and like we had happen before, everyone wanted to edit the last minute,".*
>
> "*Students would try and undermine each other just to get the largest piece of the pie. I would! If my mark was on the line and I had 11% [emphasis] I'd like do as much as I could to get 50% of the pie or bigger than anybody else, 'cause we are on a curve right?"*

On the other hand, one of the researcher participants suggested that the glyph could be used by students, not in a competitive mode, but instead as a means of ensuring that all group members contributed equally.

> *"I think… that… If I was in a class and I was working with fellow students on a project I think it would act as an equalizer, it'd make sure that the work that I'm doing is consistent with the work that*

> *other folks are doing in a class, and I would hope that it would act as an equalizer for everyone else as well. Umm… particularly if you use it as a set up as an expectation that everyone would be in a certain range of contributions it would be a useful tool to make sure that everyone is contributing to the group project."*

A very interesting comment was sent by email by one of the student participants after the interview. This student saw the contradiction between a collaborative assignment and a competitive ranking of individuals.

> *"I think that the entire concept of wikis revolves around collaboration and the concept behind the assignments we're given is competition. Collaboration and competition don't really mix well together and I think it's ultimately flawed if the two are trying to be used together in an assignment."*

For the detailed glyph (Figure 2), one of the student participants mentioned the possibility of using the information about the other editors as the basis for assigning roles for the next group assignment:

> *"I think it's nice to have access to the percent of other people as well… just to see, you know, if you wanted to do it again with a different article, maybe somebody is really good at proofreading and you can just say… "Why don't you just do all the proofreading edits?" and then kind of assign things and make things a little more easy."*

In a similar vein, another participant commented on the possible use of the detailed glyph as a management tool.

> *"I think it would actually be a little bit better just as a bit of a… more information on how things are going and where people strengths are, in terms of their edits. So we can say "hey, we are doing a good job!" or maybe if we had a full bunch of links, "you guys are putting links everywhere but we don't have any additional content". It could be used as a directional thing… it's like a… to give you a snapshot of where you are…"*

Research Setting. Some participants recognized that a little bit of competition can be a positive factor even in a collaborative setting:

> *" I think that it [Glyph 2] would motivate the person to contribute more because you are not the only one that is going to be seeing it; there's other people that are going to be seeing it. It is kind of an ego booster in a way too, and you know ok… other people know when I'm contributing to the article."*

However, other wiki users working in research settings expressed some anxieties about the presence of a score and how it might give a false impression about the collaboration. Specifically, when introduced to the two variations for calculating contribution – based on the total contributions or based on the contributions that remained on the current version, one of the participants responded:

> *"I think the aggregated score [based on what content remained on the page], um it could be a bit misleading by… the problem I would have with having this score at all is that it may not tell the entire truth as to the quality of contributions that someone's made. So I think the aggregated score [total score] would be perhaps a more honest reflection of the work someone's put into it, I think is the time they've spent with the material, the time they spend working on it… not the portion of their work that makes it through to the end."*

Discussion

Recent studies suggest that certain design features that make wikis attractive on the Internet may not be desired in organizational settings. One such design feature is not highlighting authorship: while on the Internet it is useful in promoting democratic deliberation, it prevents users from gaining recognition for their wiki work, which may be important in settings such as business, government, academia, and education. Some researchers have proposed software utilities that would automatically attribute a wiki user with a score representing his contribution. However, these proposed algorithms suffer from several drawbacks: as they often use coarse measures (such as an edit count) and do not distinguish between the types of contributions made. In this paper, we've tried to address these gaps by proposing a set of algorithms for estimating wiki editors' contributions. In terms of algorithms, the novelty was (1) in the sentence-ownership algorithm, (2) in considering the effect of longevity and distinguishing between metrics based on total contributions and those that persist on the most current page, and (3) by providing a comprehensive mapping between authorship categories and algorithms. We conducted an empirical study to assess the extent to which the proposed algorithms indeed capture contributions across the various authorship categories. In addition, we also introduced a series of visualizations for communicating these contribution metrics to users, and have conducted user studies to investigate the clarity and potential impact of our proposed visualizations.

The comparison of our proposed algorithms to the manual evaluations found that the algorithms are correlated with the manual assessments, and these correlations were higher than the baselines, namely for the 'Add' and 'Delete' categories (we note that the baseline used a character as the unit of analysis, and thus it is possible that some of the improvements in performance may be attributed to these differences). Overall, the correlations between algorithms and manual scores were in the 0.3-0.5 range and statistically significant, indicating that we are able to automatically calculate proxies for editors' wiki contribution across the various authorship categories. Correlations were higher for the internal links metric estimating 'Improve Navigation' contributions, approximately 0.5, while for the other authorship categories the correlations were in the 0.30 region (although still highly significant statistically). This is somewhat counter-intuitive, as we expected the complex sentence ownership algorithm to be highly correlated with assessors' perceptions of the extent to which editors add new content. In addition, we were concerned that a simple proxy such as the number of internal hyperlinks would not well capture contributions intended to improve navigation in the page, since such contributions include the restructuring of the page which we were not able to estimate through our algorithm; to our surprise the internal link metric was highly correlated with the corresponding authorship category.

We investigated the effect of longevity and proposed two alternative approaches for estimating contributions: one based on the total number of changes made by the editor and the other based on the contributions that remain on the most recent wiki page version; we found that in all cases the algorithm based on the total number of contributions demonstrated higher correlation with the manually-produced score, suggesting that this approach is superior. This was unforeseen, as we expected the metrics for the current version to indirectly capture the quality of contributions and thus to be more correlated with the manual assessments. A possible explanation might be that our procedure for manual assessment inadvertently preferred the cumulative score, as the assessors analyzed the series of edits sequentially, from start to finish. Our proposed implementation of the sentence ownership algorithm that was weighted by the duration that the content persisted on the page (inspired by Adler et al., 2008), which we expected to outperform the other algorithms, performed poorly: worse than both the total contributions and the persisting contributions algorithms, and even worse than the baseline metric. One possible explanation for this may be that the contributions that survive the collaborative refinement process are not necessarily of higher quality, as proposed by (Luyt et al., 2009).

When trying to combine the various algorithms in order to produce an overall score of an editor's contribution, the results from the regression analysis indicate that this aggregated score matches fairly well the manual assessments of top contributor (R^2 was 0.25). While we expected the assessor's ranking to be highly influenced by content addition type contributions, our results show that this type of contribution makes little impact on the top contributor score. Interestingly, the algorithms having the largest effect on the manual score were the internal links (associated with the "Improve Navigation' category) and the number of revisions where content was deleted (i.e. the 'Delete' category). We suspect that is not necessarily an indication that these two types of contributions are more important than others; rather we explain this result by the fact that the editors adding links are active across a variety of categories, and this is why they are perceived as top contributors.

The accuracy of the algorithms is definitely important; other factors, however, such as the robustness in face of manipulation, should also play a role when making the choice of algorithms. While the internal links metric yielded the highest accuracy scores (both in terms of correspondence to the 'Improve Navigation' category and in its impact on the top contributors ranking), this metric is easy to manipulate. I.e. if users were aware that this is how contributions are estimated, they would create links even if they are irrelevant. In contrast, our proposed sentence-ownership algorithm is much more difficult to manipulate (especially the variation based on the sentences remaining in the current version), and even if users knew how the score is calculated it would be difficult to artificially inflate the score.

In terms of communicating the contribution scores, we proposed some novel visualizations. Our approach to visualization is similar in spirit to WikiDashboard, yet there are some important differences. First, WikiDashboard in interested in exposing collaboration patterns over time, while our interest is in editors' cumulative contributions. Second, we recognize the various types of contributions and visualize editors' contributions across these various authorship categories. Third, WikiDashboard uses a simple metric that could be easily manipulated for estimating contributions, while we developed more complex algorithms (i.e. sentence ownership). Finally, our graphical; representation is quite different from WikiDashboard, and we use color, shape, and font size to communicate our more complex data. Our user studies indicate that by and large our glyphs were understood by the users, and we did not notice any real differences between the classroom and collaborative research settings. It seems that a good approach would be to provide access to both a summary version (where an overall contribution score is presented) and a more-detailed version of the information glyphs (where contributions per authorship category are presented).

However, some of the details were not entirely understood by some users; this is not an entirely negative finding, in that it suggests that it would not be immediately obvious how someone could "game" the algorithm to get a better score.

We expected the visualization of contribution scores to impact the motivations for participation, and our findings support this conjecture. Results for the user study suggest that adding the visual representation of relative contribution would in a sense "up the ante" for both students in a classroom setting and researchers on collaborative projects. Some of our participants felt that the affordance for increased surveillance from others in the group, combined with their own enhanced knowledge of relative contribution, would be beneficial, while others expressed some anxieties about the possible consequences, either in concerns about the meaning of the relative contribution score and how those might be subject to misinterpretation, or in questioning at a more fundamental level the value of an attempt to combine collaboration and competition. Interestingly, some participants also saw potential for uses of the detailed information glyph that we had not foreseen, such as in subsequently assigning tasks based on previous high scores for one kind of editing activity vs. another, or in managing the current wiki task by monitoring where a lot of activity was taking place while another activity was being neglected. The findings for the collaborative research settings were also two-sided: some researchers felt that a visualization of contributions could motivate group members to increase both the quantity and quality of their participation, while others expressed concerns about the ability of an automatic tool to accurately capture peoples' contribution towards the group effort.

Conclusion

Wiki is a promising technology that has the potential to transform knowledge management by eliminating the bottlenecks associated with traditional knowledge acquisition (Wagner, 2004; 2006). However, for such a decentralized collaborative technology to succeed in settings where users are largely driven by career goals, e.g. corporate, government, research, or education settings, incentives for user contribution are required. The application of such incentives demands that we are able to estimate user's relative contribution towards the group effort. Recently, several studies have begun exploring ways to automatically estimate contributions and visualize them. Our work extends these previous studies by: (a) acknowledging that wiki work includes various authorship categories and proposing a set of algorithms intended to capture contributions across all these categories, (b) developing some

novel algorithms for estimating contributions, namely the sentence-ownership set of algorithms, (c) proposing some novel visualizations of wiki editors' contributions, and (d) empirically testing both the algorithms and the visualizations. The scope of this paper is very broad; it includes: a discussion of wiki authorship categories, development of algorithms and visualizations, and the empirical evaluation of both. We acknowledge that in deciding to include this breadth in a single paper, we might to some extent have sacrificed depth.

In the future, we plan to extend this study in various ways. First, we intend to refine the algorithms to better capture contributions along various authorship categories. The correlations we have obtained (in the 0.3-0.5 range) leave much room for improvement. Some specific directions include modifications to the longevity sentence-ownership algorithms and the development of a new algorithm to estimate structural changes (considered a part of the 'Improve Navigation') category. Second, future research should generalize our results to other settings. In this study, we evaluated the algorithms on a small set of Wikipedia pages. Although we cannot identify a clear reason why the performance of the algorithms would be restricted to the specific pages we have analyzed, it is still necessary to repeat the evaluation on a larger set of Wikipedia pages and on wikis specifically used in other contexts, e.g. corporate, government, research, or education. Third, the visualizations could be improved, based on insights from our user study. Lastly, our findings illustrate both potential risks and benefits that are associated with tools for estimating and visualizing wiki contributions. Such tools should be implemented in real-life settings, so their actual impact could be investigated.

An important aspect that was not addressed in this study is the quality of editors' contributions. We concentrated our efforts on estimating the quantity of contributions, and only indirectly touched on the issue of quality (through the notion of longevity). Contribution quality is an important aspect of online collaborations, yet it is very difficult to capture. Recent years have seen the development of various tools for estimating the quality of the group output (e.g. Wikipedia developed a rating system for articles), the quality of one's contributions (e.g. a rating system, such the one used in the technology news forum Slashdot.org), and the reputation (or authority) of contributors (e.g. using a reputation system, such as the one used in eBay). These approaches represent promising possibilities for addressing the notion of content quality in wikis.

In conclusion, we believe that the gap between the affordances of wiki technology (that was originally designed for peer-based governance) and traditional top-down corporate governance presents an opportunity for extending wiki technology. In this study we have focused on one particular enhancement – highlighting editors'

contributions – but other wiki extensions are possible. For example, wikis might be extended to include peer-based quality assurance tools, such as rating mechanisms. We hope that the current study will encourage further research into this promising area.

References

Adler T.B., de Alfaro L., Pye I., and Raman V. (2008). Measuring Author Contributions to the Wikipedia. P*roceedings of WikiSym'08*, September 8–10, Porto, Portugal.

Anthony, D., Smith, S. W., & Williamson, T. (2009). Reputation and reliability: The case of the online encyclopedia Wikipedia. *Rationality and Society , 21* (3), 283-306.

Arazy O. and Stroulia E. (2009). A Tool for Estimating the Relative Contributions of Wiki Authors. *Proceedings of the 3rd International Conference on Weblogs and Social Media (ICWSM'09)*, May 2009, San Jose, California, USA.

Arazy O., Gellatly I., Jang S., and Patterson R. (2009). Wiki Deployment in Corporate Settings. *IEEE Technology and Society*, Volume 28, Number 2, Summer 2009, pp. 57-64.

Bisaillon J. (2007). Professional editing strategies used by six editors. *Written Communication, 24* (4), pp. 295–322

Cheng, L. and Kerr B. (2009). A Collaborative User Experience Project, *IBM Watson Research Center*, http://domino.watson.ibm.com/cambridge/research.nsf/242252765710c19485256979004d289c/864fa36402234708852570f90079a47a?OpenDocument

Ding, X., Danis, C., Erickson, T., & Kellogg, W. (2007). Visualizing an enterprise wiki. In *Proceedings of ACM CHI '07*, pp. 2189-2194, San Jose.

Ehmann K., Large A., and Beheshti J. (2008). Collaboration in context: Comparing article evolution among subject disciplines in Wikipedia. *First Monday*, 13:10, 6 October 2008

Ekstrand, M., and Riedl, J. (2009). Rv you're dumb: identifying discarded work in Wiki article history. *Proceedings of WikiSym '09*, Orlando, FL. ACM Publishing, New York, NY (USA).

Espiritu, C., Stroulia E., and Tirapat T. (2006). The ViskiMap Toolkit: Extending Mediawiki with Topic Maps. In Y. Manolopoulos et al. (Eds.): ICEIS 2006, LNBIP 3, pp. 420–438, 2008, Springer-Verlag, Berlin Heidelberg.

Fallis D. (2008). Towards an Epistemology of Wikipedia. *Journal of the American Society for Information Science & Technology*, 59:10, 1662-1674.

Hess, M., Kerr, B., & Rickards, L. (2006). *Wiki user statistics for regulating behaviour*, working paper. http://icd.si.umich.edu/684/files/684%20wikistat%20paper%202.pdf

Hoisl, B., Aigner, W., and Miksch, S. (2006). Social rewarding in wiki systems - motivating the community. *Lecture Notes in Computer Science, 4564*, pp. 362-371.

Holloway T., Bozicevic M., and Börner K. (2007) Analyzing and visualizing the semantic coverage of Wikipedia and its authors, *Complexity*. 12:3, 30-40.

Kane, G., and Fichman, R. G. (2009). The shoemaker's children: Using wikis to improve IS research, teaching, and publication. *MIS Quarterly , 33* (1), 1-22.

Korfiatis, N., Poulos, M., & Bokos, G. (2006). Evaluating authoritative sources in collaborative editing environments. *Online Information Review, 30* (3), pp. 252-262.

Leuf, B., & Cunningham, W. (2001). *The Wiki Way: quick collaboration on the web.* Addison-Wesley.

Lim S. (2009). How and Why do College Students Use Wikipedia, *Journal of the American Society for Information Science & Technology*, 60:11, 2189-2202.

Luyt, B., Aaron T., Thian L.H., and Hong C. (2008). Improving Wikipedia's Accuracy: Is Edit Age a Solution? *Journal of the American Society for Information Science & Technology*, 59:2, 318-330.

Majchrzak, A., Wagner, C., and Yates, D. (2006). Corporate wiki users: results of a survey. *Proceedings of the international symposium on Wikis* (pp. 99-104). Odense, Denmark: ACM.

Munkres, J. (1957). Algorithms for the assignment and transportation Problems. *Journal of the Society for Industrial and Applied Mathematics (SIAM)*, 5, March 1957, pp. 32-38

Nov, O. (2007). What Motivates Wikipedians. *Communications of the ACM* 50 (11) 60-64

Patterson, R., Gellatly, I., Arazy, O., and Jang, S. (2007). The effects of wikis characteristics on performance quality. *Proceedings of the 17th Workshop o Information Technologies and Systems (WITS '07).* Montreal.

Pfeil U., Zaphiris P., and Ang C. (2008) Cultural Differences in Collaborative Authoring of Wikipedia. *Journal of Computer-Mediated Communication*, 12:1, pp. 88-113.

Rashid A., Ling K., Tassone R., Resnick P., Kraut R. and Riedl J. (2006). Motivating Participation by Displaying the Value of Contribution. *Conference on Human Factors in Computing Systems*. ACM New York, NY, USA. April 22-27, 2006. pp 955-958

Ruecker S., Fiorentino C., Arias C., Bouchard M., Ganev V., Arazy O., and Stroulia E. (2008). Visualizing Relative Wiki Contributions. *The Third Annual Chicago Colloquium on Digital Humanities and Computer Science; DHCS 2008*, November 1-3, 2008, Chicago, USA.

Sabel, M. (2007). Structuring wiki revision history. *Proceedings of the 2007 international symposium on wikis* (pp. 125-130). Montréal: ACM.

Stvilia B., Twidale M., Smith L., and Gasser L. (2008). Information quality Work Organization in Wikipedia, *Journal of the American Society for Information Science & Technology*, 59:6, 983-1001

Suh, B., Chi, E., Kittur, A. and Pendleton, B. (2008). Lifting the veil: improving accountability and social transparency in Wikipedia with wikidashboard. *Proceedings of SIGCHI 2008*, Florence, Italy. ACM Publishing, New York, NY (USA).

Sunstein, C. (2006). *Infotopia*. Oxford University Press.

Viegas, F., Wattenberg, M., & Dave, K. (2004). Studying cooperation and conflict between authors with history flow visualizations. *Proceedings of CHI 2004* .

Wagner, C. (2004). Wiki: A technology for conversational knowledge management and group collaboration. *Communication of the Association for Information Systems*, *13*, 265-289.

Wagner, C. (2006). Breaking the knowledge acquisition bottleneck through conversational knowledge management. *Information Resource Management Journal* , *19* (1), 70-83.

Footnotes

[1] UIUC CCG: http://l2r.cs.uiuc.edu/~cogcomp/atool.php?tkey=SS

[2] We consider a "minor change" if *more than* Y% of the words of $S_{current}$ can also be found in $S_{previous}$; in this experiment we set Y to equal 50 percent.

[3] We consider a "major change" if *less than* Y% of the words of $S_{current}$ can also be found in $S_{previous}$; in this experiment we set Y to equal 50 percent.

Tables

Table 1
Summary of Proposed Algorithms.

Authorship Category	Unit of analysis	Key principle	Algorithms' variations
Add	Sentence	Editors' own' sentences they have created.	a. Cumulative sentence additions b. Sentence additions that persist on the most recent version. c. Sentence additions weighted by the time they persisted.
Improve navigation	Hyperlink	Internal hyperlinks are added to make navigation within the page easier.	a. Cumulative internal link additions. b. Internal link additions that persist on the most recent version.
Delete	Sentence	Removal of irrelevant content is an important contribution.	a. Total number of sentences deleted. b. Number of revisions where editor deleted content.
Proofread	Word	Editors make small proofreading corrections at the word level.	a. Cumulative number of proofreading changes b. Proofreading changes that persist on the most recent version.
Link	Hyperlink	Hyperlinks to other wiki pages link to relevant content.	a. Cumulative external link additions. b. External link additions that persist on the most recent version

Table 2

Summary of details for the Wikipedia articles analyzed

Article title	URL	Date		Duration	Edits	Unique Authors
		Start	End			
Aikido	/Aikido	11/29/2001	06/13/2004	2.5	72	62
Angel	/Angel	11/30/2001	12/09/2005	4.0	341	277
Baryon	/Baryon	02/25/2002	08/28/2005	3.5	73	62
Board Game	/Boardgame	11/04/2001	12/30/2004	3.2	220	155
Buckminster Fuller	/Buckminster_Fuller	12/13/2001	07/14/2004	2.6	65	55
Centres for Disease Control and Prevention	/Centers_for_Disease_Control_and_Prevention	10/16/2001	03/05/2006	4.4	65	58
Classical Mechanics	/Classical_Mechanics	06/06/2002	08/13/2006	4.2	202	165
Dartmouth College	/Dartmouth_College	10/01/2001	08/26/2004	2.9	70	55
Erin Brockovich	/Erin_Brockovich	09/24/2001	02/02/2006	4.4	59	54
Average				3.5	129.7	104.8

Note: All URLs of Wikipedia articles start with en.wikipedia.org/wiki

Table 3

Add category.

	Sentences owned		
Characters added	Total	Current version	Weighted by longevity
0.236***	0.295***	0.290***	0.114***

Note: Pearson correlations of manual assessment with the proposed sentence ownership algorithms. '***' indicates significance of $p < 0.001$ (two-tailed).

Table 4

Delete category.

Characters added	Sentences deleted	Revisions with deleted content
0.197***	0.157***	0.274***

Note: Pearson correlations of manual assessment with the proposed sentence ownership algorithms. '***' indicates significance of $p < 0.001$ (two-tailed).

Table 5

Results for regression of the set of algorithms on the manually-produced 'top contributor' score.

Algorithm	Category	β	t value	Significance (p<)
(Constant)		__	1.029	.304
Sentence Ownership[a]	Add	.055	1.696	.090
Internal hyperlinks[a]	Improve Navigation	.366	9.285	.000
Word-level[a]	Proofread	- .044	- 1.364	.173
Deletions[b]	Delete	.187	5.538	.000
External hyperlinks[a]	Link	.037	1.091	.276

Note: 'a' refers to algorithms' total scores, while 'b' indicates scores based in number of revisions

Figures

Figure. 1. The sunword glyph provides an aggregate score for all contributors (centre) and individual scores indicated both by font size and a number beside each name.

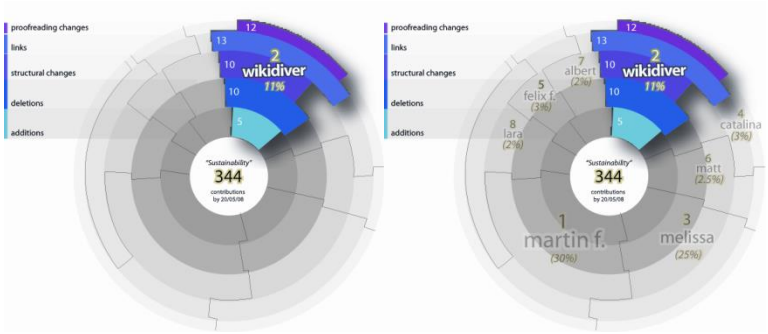

Figure 2. The detailed glyph provides of the selected editor's score for each of the authorship categories. In the version to the left, only the selected author "wikidiver" is shown, while in the right the grey portions also contain details about the other authors.

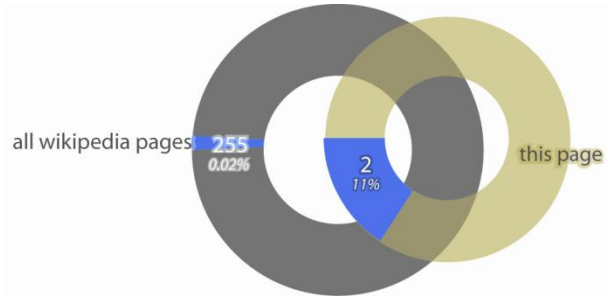

Figure 3. The overlapping doughnuts in our third glyph are an attempt to show a person's contribution not only to the current page (on the right), but also to the larger wiki application (Wikipedia, in this specific visualization).

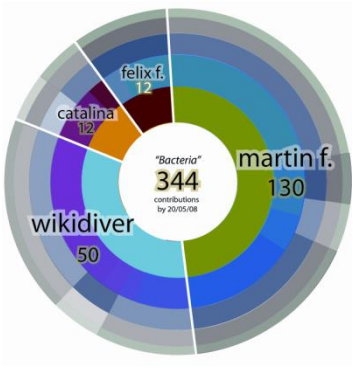

Figure 4. Our detailed prototype, installed within MediaWiki, provides colours for every author, a switch for labels, and an interactive mode where the segments respond to rollover by swelling and displaying their details.