Snuggle: Designing for efficient socialization and ideological critique

Aaron Halfaker
Research & Data
Wikimedia Foundation
ahalfaker@wikimedia.org

R. Stuart Geiger
School of Information
University of California, Berkeley
stuart@stuartgeiger.com

Loren Terveen
GroupLens Research
University of Minnesota
terveen@cs.umn.edu

ABSTRACT

Wikipedia, the encyclopedia "anyone can edit", has become increasingly less so. Recent academic research and popular discourse illustrates the often aggressive ways newcomers are treated by veteran Wikipedians. These are complex sociotechnical issues, bound up in infrastructures based on problematic ideologies. In response, we worked with a coalition of Wikipedians to design, develop, and deploy Snuggle, a new user interface that served two critical functions: making the work of newcomer socialization more effective, and bringing visibility to instances in which Wikipedians current practice of gatekeeping socialization breaks down. Snuggle supports positive socialization by helping mentors quickly find newcomers whose good-faith mistakes were reverted as damage. Snuggle also supports ideological critique and reflection by bringing visibility to the consequences of viewing newcomers through a lens of suspiciousness.

Author Keywords

Wikipedia; Activism; Design; Algorithms; Critique

ACM Classification Keywords

H.5.2. Information Interfaces and Presentation: Graphical user interfaces (GUI)

INTRODUCTION

A decade of research on Wikipedia has sought to explain how the self-proclaimed "free encyclopedia anyone can edit" could possibly work. This multidisciplinary research has documented and analyzed the social and technical aspects behind the project's peer production model, which solved a set of difficult problems and led to Wikipedia's massive and unexpected success. These studies include explorations of self-driven newcomer socialization patterns[6], the development of highly-effective quality control algorithms[36], and the robust, distributed workflows and procedures used to enforce the project's rules and norms [15]. However, recent work has identified a set of complex, deep-seated, socio-technical problems around newcomer socialization in Wikipedia. These problems threaten the health of the community and the long-term viability of the project.

In 2005, Wikipedia's volunteer editor community and the size of the encyclopedia began growing exponentially[35]. During this time Wikipedia faced a series of crises in the public sphere over its trustworthiness and legitimacy. Wikipedia's "vandal fighters" came to see Wikipedia as a firehose of edits needing constant surveillance. By 2007, they had developed quality control practices around a suite of standards, discourses, procedures, and roles. To make their work practical, they formalized the practice of reviewing edits around a suite of algorithmically-assisted, semi-automated tools[15].

Consequently, today's vandal fighters see a different Wikipedia than most people do. In one sense, this is a metaphor about "social worlds"[33], where people learn such different ways of interpreting and experiencing that they can be said to inhabit different worlds. Yet these vandal fighters also *literally* see something different: their work often begins not by opening up a web browser, but through algorithmically-assisted external tools. For example, Huggle, a popular counter-vandalism tool, is a desktop application that presents a queue of before-and-after edits to review, each edit ranked by "suspiciousness". With one click, vandal fighters can instantly reject an edit, send its author a prewritten warning, and mark the author as a potential vandal to be blocked.

Tools like Huggle raise practical design challenges and ethical issues for HCI researchers. In previous work, we have critiqued the "professional vision"[17] they enact and the assumptions and values they embody: most tools situate users as police, not mentors, affording rejection and punishment. Newcomers who make low-quality edits are situated as potential vandals or spammers to be thwarted, instead of individuals who may need assistance and guidance in navigating Wikipedia's labyrinth of policies, rules, and norms. These highly-automated tools have become the dominant way in which Wikipedians – established editors with well-defined social roles who make hundreds or even thousands of edits a month – interact with non-Wikipedians[14]. In our previous work, we implicated this ideology of "gatekeeping socialization" and the use of automated counter-vandalism tools in causing several newcomer socialization problems that have

resulted in a substantial drop in the retention of desirable new contributors[19].

As HCI researchers who are also deeply situated in Wikipedia, we have a unique perspective on the context and history of these tools and practices. These tools reflect and perpetuate their designers' situated understandings of the specific problems they were facing as Wikipedia became an increasingly important social institution. As Wikipedia is a complex socio-technical system, its problems and solutions are likely to be just as heterogeneous and multifaceted. So how should we approach designing a solution? We were initially split between solving a specific problem about information systems with effective design, and critiquing the fundamental assumptions that are embedded in the designs of existing systems. However, we found ourselves doing both. We sought to both design a new tool to efficiently support underdeveloped socialization tasks in Wikipedia, as well as give Wikipedians a way to critically reflect on their own practices and assumptions around socialization.

In this paper, we describe the design and evaluation of *Snuggle*, a collaboratively-designed newcomer socialization system for Wikipedia. This paper contributes to HCI in three ways, by presenting Snuggle as (1) a case of an intelligent newcomer socialization tool that increases task efficiency and supports the development of new norms of practice, with implications for systems like MOOCs and citizen science; (2) as a critical HCI project that reverses the assumptions built into an existing, dominant system to enact and support reflexive ideological critique from within; and (3) a case of a highly-participatory design process, in which situated methods were used to not only shape the design of the interface, but the design of the design process itself.

WIKIPEDIA'S SOCIO-TECHNICAL PROBLEMS

Wikipedia has become widely scrutinized for seeming less like a participatory community where "anyone can edit." Academics, journalists, and celebrities use mass and social media to share their struggles in editing articles [11]. Newcomers in general are not staying around as long as they used to [35], and those making good-faith contributions find their work rapidly and unexpectedly deleted at over three times the rate they did in 2006 [19]. This is even more problematic given that Wikipedians are disproportionately young, white, college-educated men in the US and Europe. Wikipedia's coverage has gaps that reflect these systemic biases [25].

Gatekeeping as socialization

Many of these problems are inextricably linked to a fundamental shift Wikipedians made during the 2005-07 growth period to focus on standards, practices, roles, tools, and algorithms. This tool-enabled division of labor afforded ad-hoc quality control that efficiently scaled, leaving most Wikipedians free to do more productive tasks[15]. Today, these counter-vandalism tools are critical to how Wikipedians maintain a decentralized, open system at massive scale. When one of these algorithmic tools went down in 2011, it took Wikipedians almost twice as long to remove vandalism [13].

As fully-automated bots and tool-assisted "cyborgs" did more gatekeeping work, they became the first point of contact for 75% of newcomers by 2010 [14]. Almost all of these systems were primarily built to support counter-vandalism, seeing individual edits as potential threats and supporting fast-paced reversion, warning, deletion, and blocking. While newcomers are not making lower quality edits than before, recent research shows that the rate at which newcomers recieve warnings for their "vandalism" has grown substantially [19].

Self-directed socialization

While vandal fighting and 'gatekeeping socialization' has been well-supported in Wikipedia, traditional socialization practices like mentoring have been far less organized. Newcomers are often expected to be proactive and self-directed. In 2005, Bryant et al. studied how newcomers move from peripheral readers to prolific participants: "becoming Wikipedian"[6] typically involves gaining new skills as needed when performing increasingly complicated tasks. This process of self-directed socialization worked for many newcomers when Wikipedia was young; they became today's veteran Wikipedians. However, this self-directed socialization is increasingly not working for many desirable newcomers, who quickly get frustrated and leave.

The reasons for this include but go beyond counter-vandalism tools. Since Bryant's work in 2005, Wikipedians have developed hundreds of policies, procedures, and guidelines that new editors are expected to understand. This has lead to "literacy asymmetries" [11]. To assist newcomers in navigating this complexity, Wikipedians have developed specialized mentoring spaces and practices to assist new editors, but these efforts have not been as numerous or successful as those for vandal fighting. Researchers have critiqued how these programs require that newcomers seek out help: newcomers most in need often do not know where[1] and how to ask for help [27] and newcomers often leave the project before mentors are able to intervene [29].

DESIGN STRATEGY

As we mentioned in the introduction, our initial formulation of Snuggle was based on our concern with how the existing Huggle counter-vandalism tool framed newcomers' activities as problems to be dealt with. Our previous work suggests that the widespread use of Huggle has systemic, long-term implications in that viewing newcomers through lenses of quality control and counter-vandalism situates newcomers as inherently suspicious, rather than people who may make well-intentioned mistakes in the course of learning how to be a part of a community[19].

We knew of many Wikipedians who were interested in mentoring and socialization, so we saw an opportunity to design a tool that would support their practices just as Huggle supports vandal fighters. However, if vandal fighters continued their gatekeeping unabated, Snuggle might only represent a stop-gap attempt to retroactively respond to newcomers who

[1] `http://meta.wikimedia.org/wiki/Research:New_user_help_requests`

already had been bitten[2]. To further complicate the picture, we knew that Huggle served a real need: vandalism was and is a real threat to Wikipedia. In fact, some of our collaborators and beta testers were also users of "competing" tools, and emphasized the need for Snuggle to support tasks like requesting an admin block a problematic newcomer. We empathize with these needs, and we do not see vandal fighters as "the enemy" and Snuggle as a tool to equip our own army of "vandal fighter fighters."

Reflecting on our roles as design researchers and Wikipedians and inspired by critical HCI research advocating reflexivity in design [32], we began to see Snuggle as more than an instrumental tool. Even before our first prototype was sketched, other Wikipedians began talking about Snuggle in broader conversations about participation, representation, and inclusion in Wikipedia. Snuggle became aligned with emerging newcomer socialization projects like the Teahouse, first in discourse, then later in code. We realized Snuggle could add missing voices to the conversations about Wikipedia's discontents, and found that looking closely at these backgrounded processes of gatekeeping socialization gave us many compelling stories to tell.

As HCI researchers situated in Wikipedia, we didn't just have a better sense about how to support work practices with software; we also had a better sense about what kinds of conversations were taking place and what was missing from them. We wanted to give all Wikipedians – not just dedicated mentors – a tool for finding, exploring, and reflecting on cases where newcomers were making good faith efforts to contribute, but had their mistakes flagged as vandalism. Thus, Snuggle is intended to both (a) support early and positive mentoring and (b) show where Wikipedians' current practice of gatekeeping socialization breaks down in order to enable reflection and critique. Put crudely, Snuggle will (we hope) reduce the biting of newcomers both immediately (by enabling support of bitten newcomers) and systemically (by changing how experienced Wikipedians view newcomers).

HOW DO WE EFFECT CHANGE?

The previous two sections gave two narratives of our motivation for building Snuggle. First, empirical research helped us design an effective tool supporting the practices and activities of Wikipedians who wanted to find and help newcomers in need. Second, situated reflection helped us design a tool that would reveal and critique the broader ideological assumptions embedded in the design of dominant systems. HCI literature was useful for both these goals.

Classic user-centered design emphasizes three factors: iterative design, empirical measurement, and a focus on users and tasks [18]. It seeks to design systems in ways that align with how people actually approach situations. Later research emphasized the co-evolution of systems and practices based on the insight that introducing a new tool transforms the user's tasks and context [31], and saw formal task descriptions as resources that people dynamically use and interpret in the course of performing situated actions [34]. Frameworks such as activity theory[30] and distributed cognition[23] take into account how action or cognition is situated in a diverse set of technological and social contexts. Design approaches such as ethnographically-informed design[3] and participatory design[28] view people not as users to be designed for, but as collaborators to design *with*. We took these lessons to heart as we worked with Wikipedians to understand existing mentoring practices, design prototypes for supporting unmet needs, iterating designs based on feedback and evaluation, and reflecting on how the new tool was being deployed in an existing socio-technical system.

We were also inspired by approaches from so-called "third-wave" [21] and critical HCI: ideological critiques of dominant systems resonated with the problems we saw in vandal fighting tools. We were critical of how newcomers in Wikipedia are often seen through one particular lens representing the vandal fighter's perspective, which is far from universal. Yet instead of de-legitimizing this ideology through discourse, we wanted to also build a better lens though which Wikipedians could view newcomers. HCI researchers have long been blending and iterating between "second-wave" and "third-wave" approaches [4], and we found a variety of literatures indispensable in situating Snuggle as both a user-centered design project and a strategic, ideological intervention.

We saw many similarities with the "values in design" literature, where designers explicitly acknowledge principles they value and seek to uphold [10]. We also saw alignment with "action research"[22], which aims to bring about large-scale social, cultural, political, economic, or environmental benefits. Finally, our approach resonates with the goals of "critical technical practice"[1], which seeks to reverse the assumptions built into dominant systems to provoke reflection and critique. Finally, we found feminist theory and standpoint epistemology[20] particularly useful in thinking about how systems tend to universalize a single way of seeing the world, and we are indebted to Bardzell & Bardzell's[2] commentary on this literature as it applies to HCI.

These critical and feminist approaches helped us situate the design of Snuggle in relation to other HCI projects that share a similar kind of overarching design strategy. For example, Hollaback[7] represents a critique of the widespread institutional ignorance of street harassment in two related ways: it provides a safe space for victims of street harassment to assemble as a networked public, and it provides an infrastructure for building better accounts of the world, ones that make often-ignored experiences of street harassment visible at a variety of scales. Turkopticon[24] similarly represents a critique of the way Amazon Mechanical Turk turns human workers into an invisible, de-individuated infrastructure, ripe for exploitation with little to no recourse. As design activism, Turkopticon affords workers the ability to rate employers, building a better account of the world for two purposes: to "not only hold employers accountable, but induce better behavior."

[2]'Bite' is Wikipedian jargon for when an experienced editor acts aggressively towards a newcomer.

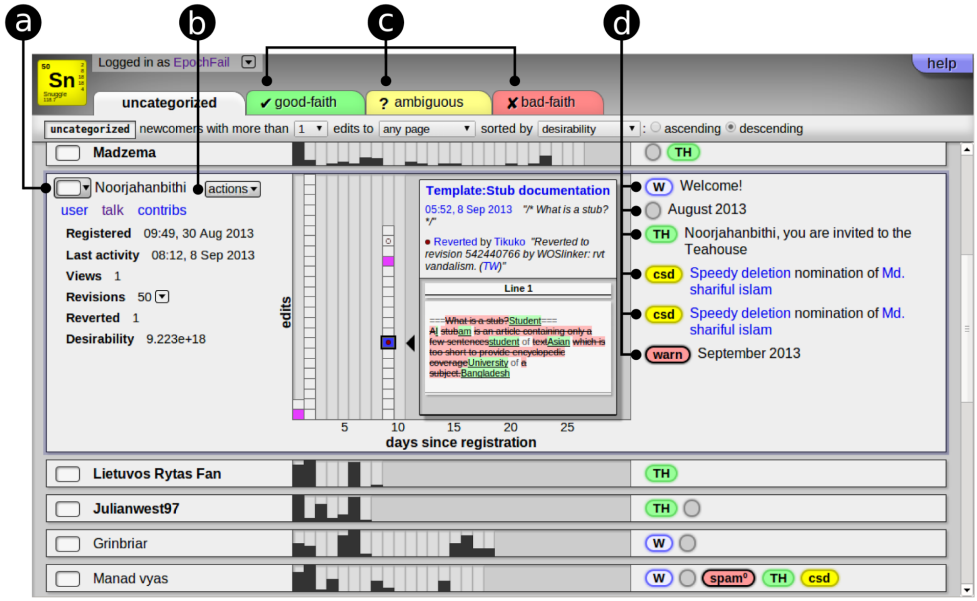

Figure 1: **Snuggle's user browser.** A screenshot of the Snuggle user browser is presented with UI elements called out. The user dossier for "Noorjahanbithi" is selected. An edit in the *interactive graph* is selected and information about the edit is presented. (a) The (unexpanded) categorization menu, (b) The (unexpanded) wiki action menu (see Trace production), (c) Tabs for accessing lists of categorized users, and (d) Talk page icons representing socially relevant traces (see Trace consumption).

DESIGN OF SNUGGLE

System overview

Snuggle is design to afford Wikipedian mentors a set of features that will allow them to identify newcomers in need of help, share their assessments of newcomer activities with each other and perform timely interventions when good-faith newcomers experience harsh treatment.

Snuggle collects information about new users into a "dossier" by tracking activity in Wikipedia. Dossiers include statistics about page editing activity, an interactive graph of edits (including notes on which edits were reverted[3]), and a visual summary of traces (representing communications with other editors) extracted from their talk page (see Social literacy via traces).

We collect this information from the English Wikipedia by reading "recent changes"[4] from the website's API[5]. Using this feed of activity, Snuggle displays user dossiers on all editors who have registered within the last 30 days.

Using Snuggle, Wikipedians can observe the activities of newcomers by viewing their dossier, and they can share their assessment of newcomers' activity with other mentors by moving newcomers' user dossiers between four lists: uncategorized, good-faith, ambiguous, and bad-faith (figure 1c) using the categorization menu (figure 1a). Snuggle users can also interact directly with newcomers and other Wikipedians though the wiki actions menu (figures 1b and 4, e.g. sending a message or reporting abuse to an administrator).

Desirability sorting

Every day, about 1,000 people register an account and make at least one edit to English Wikipedia. Mentors can't wade through that many newcomers unless they devote several hours a day to the work and abandon encyclopedia writing entirely. Snuggle needed to efficiently support identifying desirable newcomers.

Desirability in concept. Wikipedians refer to the desirable behavior of others as "good-faith". In Wikipedia, the concept of "good-faith" is based on the intention of a user as opposed to the effects of their actions. When discussing newcomers, the Assume Good Faith guideline states[6]:

> A newcomer's behavior probably seems appropriate to him/her, and a problem in that regard usually indicates unawareness or misunderstanding of Wikipedian culture.

The guidelines stresses the importance of seeing damaging edits as mistakes rather than as intentional.

Desirability in practice. Classifying the intentions of newcomers as "good faith" or "bad faith" is a core part of socialization in Wikipedia, used to direct further efforts. To help mentors efficiently prioritize newcomers to assist, we sought to rank newcomers using the same technique as counter-vandalism tools, but with the opposite valence. We built a

[3] When an edit is "reverted" it's changes are removed from an article.
[4] a chronologically sorted list of most activities that take place on Wikipedia from edits to new user registrations
[5] `http://en.wikipedia.org/w/api.php`
[6] `http://enwp.org/WP:AGF`

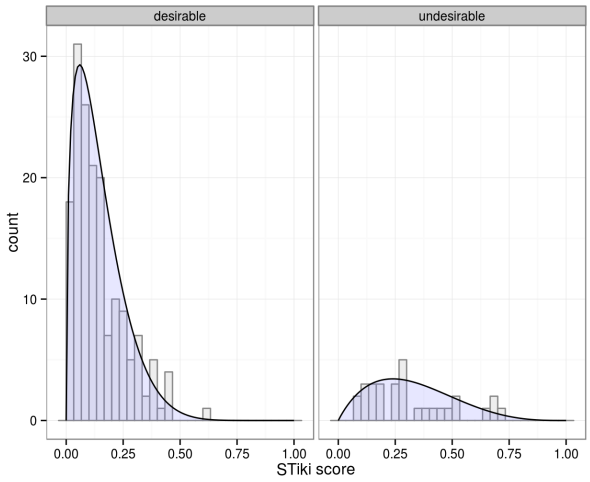

Figure 2: **Desirability scores.** Histograms of the frequency of STiki scores are plotted for the training set newcomers' edits with expectation maximization fits of beta distributions overlayed.

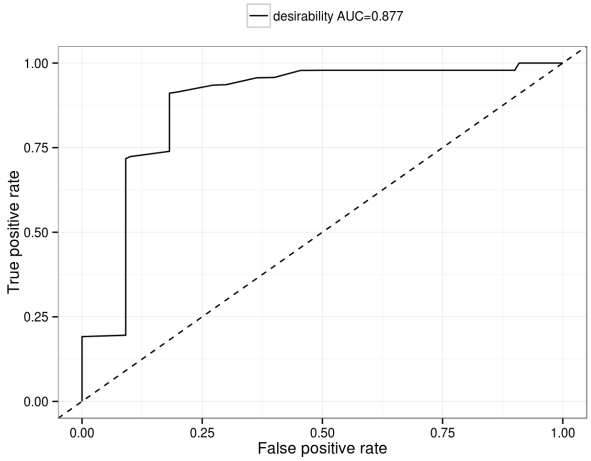

Figure 3: **Desirability ROC.** The receiver operating characteristic of the desirability ratio of newcomers from the test set is plotted.

model of newcomer activities that assigns a likelihood that the editor is good-faith. In order for this model to be useful in Snuggle, the modeling strategy needed to accurately classify newcomers who have made few edits without accounting for negative reactions received by newcomers (e.g. reverts, warnings, and blocks) since those are cues for mentorship opportunities.

Information source. One approach is to sort newcomers by the proportion of their edits that have been reverted, but this defeats the broader goals of Snuggle. If only newcomers who are least reverted are determined to be working in good-faith, then Snuggle would not be a useful tool for identifying good-faith newcomers who are reverted due to mistakes or misunderstandings.

In order to avoid considering vandal fighters' reactions to newcomers, we strategically take advantage of sophisticated models used to assess newcomer behavior in Wikipedia: counter-vandal bots. Many of these bots publish scores of individual edits, based on the probability that the edit is vandalism. We suspected such scores would be useful for differentiating the activities of good-faith newcomers from bad-faith newcomers, independent of whether or not the edits were eventually reverted.

Modeling desirability. We constructed a Bayesian model by intersecting a dataset of newcomers hand-coded as "desirable" and "undesirable" from [19] with scores retrieved from STiki's[7] API to arrive at 152 hand-coded newcomers and 377 scored "first session"[8] edits.

[7]A tool that's very similar to design and use as Huggle. See `http://enwp.org/WP:STiki` and [36]

[8]An edit session is a concept formalized in [12] that temporally clusters edits together into "sessions". A users "first session" represents their first editing experience as a registered editor.

We randomly split the set of newcomers in half to create a pair of training and test sets (76 users/set). We then used an expectation maximization approach to fit two beta distributions to the training set scores for desirable and undesirable users. Using these two distributions as models for STiki scores attributable to desirable and undesirable editors (see figure 2), we use the following function to generate the odds ratio:

$$\text{desirability ratio} = \frac{p(\text{scores}|\text{desirable})p(\text{desirable})}{p(\text{scores}|\text{undesirable})p(\text{undesirable})} \quad (1)$$

With this approach, we were able to attain a relatively high AUC (0.877) using only scores from edits that newcomers performed in their first session.

Social literacy via traces

To visualize socially relevant activity and act in Wikipedia, Snuggle took into account the various metadata, log entries, revision histories, template markers, page categories, and other standardized records that Wikipedians rely on to coordinate related tasks and share information. These structured documentary traces are a core component of social and organizational interaction in not just Wikipedia, but a variety of 'virtual' and traditional co-located organizations. These traces "not only document events but are also used by participants themselves to coordinate and render accountable many activities." [16] For example, vandal fighters use template markers and page categories to track newcomers, using their public talk pages as a kind of shared database to determine how close they are to being blocked from editing [15]. Understanding traces is part of what it means to be a Wikipedian, and traces are followed and left in performing many socially relevant actions. Ford & Geiger argue that newcomers who are not yet "trace literate" suffer from power imbalances by

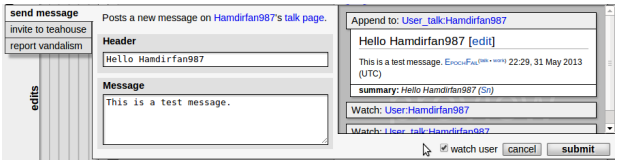

Figure 4: **Wiki actions menu.** A screenshot of the "wiki action menu" is presented with the message sending functionality selected and a test message written. Note the preview on the right side specifies which page the message will be appended to.

being unable to participate effectively or know how they are being tracked [11].

In order for Snuggle users to take into account the actions taken against newcomers – as well as to have their own actions affect Wikipedia – Snuggle will have to consume and produce traces.

Trace consumption

For Snuggle's trace consumption, we focused on newcomers' talk pages. As mentioned previously, a user's talk page is used both to capture one-on-one conversations as well as to document the interactions that the user has had with Wikipedia's quality control system.

Luckily, the structure of these traces is highly consistent because Wikipedians mostly use templated messages when interacting with newcomers. This consistency lends itself to detectability, so we were able to define whole classes of traces with a simple set of regular expressions. Snuggle represents traces with icons on the right side of the user dossier (see figure 1d).

Trace production

In order to support Wikipedians' work practices, some actions will need to be performed back in the wiki. Snuggle supports relevant newcomer mentorship and routing actions: send a message, invite to the Teahouse[9] and report vandalism [10].

In Wikipedia most traces are preserved via edits to pages, so we developed a configurable trace production system capable of previewing and producing page edits. Figure 4 shows the "wiki actions" menu with a form describing the action to be performed on the left hand side and a preview of resulting page edits on the right.

Social translucence

Snuggle is intended to support and extend a work practice in a socio-technical system. In order to support the *social processes* surrounding the development of new practices, we took inspiration from Erickson & Kellogg's work describing the design of socially translucent systems [8]. They argue for three characteristics of social translucence: visibility, awareness, and accountability.

[9] a newcomer socialization space discussed in [27]

[10] posts to a forum for bringing bad-faith editors to the attention of administrators

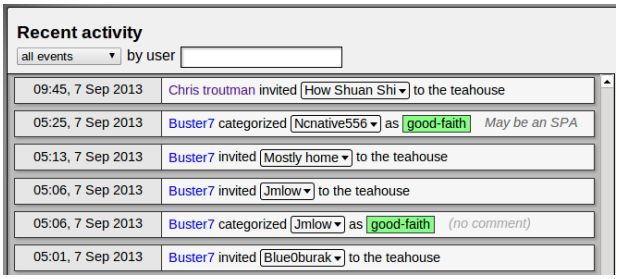

Figure 5: **The recent activity feed.** A screenshot of Snuggle's recent activity list is presented.

Snuggle makes the activities of Snuggle users both visible and prominent. Snuggle's welcome screen displays a list of recent activities performed by Snuggle users (see figure 5). Clicking on the username of the newcomer acted upon will open the user dossier for that newcomer complete with categorizer and wiki action menu. Most critically, this visibility is made apparent to Snuggle users before they've had an opportunity to log in and begin using Snuggle.

This social translucence supports the development of practices and norms around Snuggle by enabling mentors to observe each other's behavior. This visibility may also help allay concerns about bad behavior by encouraging feelings of accountability that will make our users think carefully about their actions and by supporting peer-policing in the case of troublesome users.

Participatory design process

Wikipedians have built their own infrastructure and processes for creating new tools, extensions, and bots, which we used to design Snuggle. Our participatory design took place in spaces routinely used to document other tools used by Wikipedians[11]. There, we maintained an evolving wiki page where we described the project, published prototypes, and recruited collaborators and testers. The talk page was active and successful, with 23 distinct editors who sent 107 messages. Throughout the prototyping stage, we used standard wiki talk pages to introduce various features and affordances that a mentoring tool could have, prompting open discussions about what mentoring in Wikipedia was and could be. We used talk pages as a forum throughout the design process to bring individual concerns and conversations about the design to a wider audience. We also added new design elements as probes to intentionally provoke discussion and reflection about mentoring norms, particularly the privacy of mentor-mentee interactions (see Social translucence below).

One example of how Wikipedians participated in the design process took place in the early stages of our design process: some beta testers complained that they needed to copy and paste the username of a newcomer in Snuggle to look at their activity on Wikipedia through the web browser, a practice we did not anticipate. This lead to a conversation about the established practice of including three links, commonly abbreviated as UTC, anytime a Wikiepdia user is referenced: the

[11] see `https://en.wikipedia.org/wiki/WP:Snuggle`

[U]ser's profile page, to get information about the editor; the user's [T]alk page, to send them a message; and the user's [C]ontributions, to see what other edits they had made. This simple change allowed Snuggle to become much more effectively coupled with the workflows that Wikipedians had already become familiar with. By mimicking the naming and format of these links as they are commonly represented the in Wikipedia's web interface, Snuggle users who joined the project after this feature were able to use it intuitively.

Our collaborators influenced not only the design of Snuggle, but also the design of the design process by aligning our approach with Wikipedia's norms. Practices regarding releasing updates and changelogs had to be mutually negotiated. Our collaborators worked to recruit other Wikipedians, facilitated discussions, and created some of the spaces in which we did participatory design. One even created an IRC channel for Snuggle users and configured a bot to post messages to the channel when a new update or design was posted. We see similarities between this process and other highly participatory design efforts such as in children-led cooperative inquiry sessions[37], social movements[7], and activism[24].

DESIGN REFLECTIONS

Snuggle as a newcomer socialization tool

As discussed in the previous section, Snuggle is designed to support a specific newcomer socialization task, the detection of desirable newcomers in need of help. In a way, this manifests similarly to anomaly detection. Snuggle's desirability ratio is designed to bring attention to good-faith newcomers and the trace consumption system is designed to bring attention the negative attention that these good-faith newcomers receive. In a perfect system, good-faith newcomers would not receive such negative attention.

Support for the detection of these anomalous situations is intended to make the process of identifying desirable newcomers who need support faster and more effective than the passive model currently employed by Wikipedians. Snuggle users can use the talk page trace visualization to identify warnings, deletion notifications, and other negative reactions that newcomers received and compare this response with the edits presented in the interactive graph in order to target newcomers in need of support. By providing a means for mentors to selectively intervene when desirable newcomers are treated badly, we intend to enable Wikipedian mentors to more efficiently deal with concerns we raised in [19].

Snuggle as an ideological critique

Snuggle also serves as a critique of more dominant software systems designed to support efficient interaction with newcomers: counter-vandalism tools. In the design of Snuggle, we aim our critique at three characteristics of counter-vandalism tools: edits as the unit of assessment, sorting by undesirable only, and supporting only negative reactions.

Edits as the unit of assessment. Counter-vandalism tools like Huggle only show their users a very narrow view of an editors. An individual edit, taken out of context, is a very limited frame by which to view a newcomer's activities. Given this limited amount of information, we suspect that it is much easier to ascribe simple *good* or *bad* to complicated individuals. In other words, it's easy confuse mistakes or even good work that just happens to look suspicious[12] as the work of vandals (false positives).

In Snuggle, we sought to provide our users with a more complete view of the newcomers they interact with. The first component of Snuggle that we conceptualized was the user dossier that brings together as much information as we could about a user's activities and the interactions they have had with other Wikipedians.

Sorting by undesirable only. Categorization systems and sorting practices are ubiquitous and inevitable, but we are reminded of Goodwin's analysis of the "professional vision"[17] of policing explicated in the Rodney King trial. Counter-vandalism tools like Huggle constantly show vandal fighters the worst parts of Wikipedia, which do not just include errors, spam, and nonsense, but also hate speech, shock images, and aggressive trolling.

Snuggle reverses this strategy by setting the default sort order to bring attention to good newcomers and their activities first. With this, we both enable and encourage our users to see the value that many newcomers bring to Wikipedia. Incidentally, this sort order also encourages our users to critique the practices encouraged by counter-vandalism tools by juxtaposing the activities of desirable newcomers with the negative reactions they receive.

Primarily supporting negative reactions. Huggle users are primarily afforded two responses to the edits that the user interface presents: pass or reject and send a warning. There's no affordance for saying "thank you" for good edits or even re-writing edits that would be good contributions if they were only formatted correctly. Our previous work brings attention to this problem of affordances directing user behavior and shows a dramatic growth in the rate of both rejection of desirable newcomers' edits and posting of warnings to their user talk pages [19].

In Snuggle the available actions that affect a newcomer are ordered purposefully. First, users can send a personalized message – a practice which has been declining since the introduction of counter-vandalism tools [14]. Next, users can invite newcomers to the Teahouse, a question and answer space for newcomers [27]. Finally, we still do provide the means to report vandals to the administrators of Wikipedia so that they can be banned from editing if necessary.

INTERVIEW STUDY

To evaluate Snuggle, we performed a study that combined guided use of Snuggle and semi-structured interviews. We recruited Wikipedians who were recently active in projects aimed toward positive interactions between experienced editors and newcomers (e.g., "Adopt-a-user", the Teahouse and

[12]For example, edits related to Mohamed Homos, an Egyptian midfielder who usually goes by "Homos" are often mistaken for vandalism by Huggle users

Articles for Creation[13]). We posted 70 invitations, received 25 responses, and conducted 14 interviews (at which point we saw a convergence of themes). Three of the participants had used Snuggle before and and 11 had not.

We performed the interviews using Google Hangout and Skype; we used their screen sharing features to virtually look over participants' shoulders. Our semi-structured interview and guided use session occurred in three phases. First, we asked a set of questions designed to check our assumptions (e.g., "How often do newcomers run into trouble and not know where to go for help?") and extend our understanding of current mentoring practices (e.g., "Where do you interact with newcomers?"). Next, we had participants load the Snuggle interface and gave them a high level overview of the system using a standard script that described Snuggle's user dossier lists, but did not instruct the participants about how to use the system. Then, we asked them to perform a task: identify a desirable newcomer in need of help. Finally, we concluded with a discussion of their strategy for performing the task, their opinion of Snuggle's user interface, how they felt about categorizing newcomers, and when they might consider performing each of the wiki actions Snuggle affords.

Results & discussion

When presenting our results in this section, we include quotations both from the interviews and from conversations started by Wikipedians on the Snuggle discussion forum.

Wikipedians recognize the problems newcomers face

Our participants discussed how new editors regularly run into trouble and don't know where to go to get help. When asked how common an experience this was for newcomers:

> #6: "I think it's very common. If they start doing anything, they're going to run into trouble. [...] It could be making mistakes. It could be an editor exerting ownership. They might put their talk post at the top of the page – these little rules that no one knows. Eventually you're going to do something wrong."

Most participants told us that there was no good way to find these newcomers, although some had developed intricate strategies. When asked how he finds these newcomers, one of our participants explained how he used STiki, a counter-vandalism tool, to find newcomers who are making mistakes. As STiki has no built-in support for mentoring actions, he would then return to the wiki to offer support.

Snuggle supports the identification task

Participants competently used the Snuggle UI to perform mentoring tasks without explicit guidance by the interviewer. All participants successfully used the interface to identify a newcomer they thought needed assistance within seconds of being tasked to do so, and all of them made use of the talk trace summary (see figure 1d) to identify the newcomer. Without fail, every Snuggle user used the trace icons when looking for a newcomer in need of help. Most looked for desirable newcomers with warnings. Some looked for prolific newcomers with no messages at all. Many also de-prioritized newcomers who already had a social interaction (e.g., an invitation to the Teahouse).

Surfacing these traces was essential to Snuggle's use as a means to find newcomers in need of help. When discussing talk icons, interviewee #6 commented that, "Welcome is obvious. Vandal is obvious. Warning! He got a warning. It just gives me information. I know she was welcomed. I know she was invited to the Teahouse. Here's a warning. It gives me something to work with."

Bowker and Star[5] discuss the invisibility of successful classification systems and information infrastructures, which work so well precisely because they disappear into the background and become routine. Snuggle seems to afford what they call an 'infrastructural inversion' that calls attention to Wikipedia's vandal fighting system, removing warnings from the contexts where they are taken for granted and positions them in a way that invites reflection and discussion about whether they have been appropriately applied.

Some volunteers shy away from 1:1 interaction.

Many users were comfortable performing actions with Snuggle. Some used Snuggle to identify newcomers, but they preferred to go back to Wikipedia to send messages. We were surprised to find that many preferred not to interact with newcomers at all – yet they were happy to help categorize.

Studies of prosocial behavior in organizations found that empathy correlated strongly with citizenship behaviors directed towards specific individuals [26]. However, Finkelstein et al. observed no correlation between empathy and time spent volunteering [9]. In other words, this prosocial orientation predicts whether a volunteer will favor 1:1 interactions, but not how much time and effort they will spend volunteering.

An efficient volunteer-based system should be able to take advantage of the time and effort of both prosocial and antisocial volunteers. Thus, to take advantage of all potential mentors, newcomer socialization systems like Snuggle should support work that does not require 1:1 interaction. There are a number of such socialization tasks, e.g., manually classifying newcomers as good-faith or bad-faith and flagging good-faith newcomers in need of help.

Evidence of reflection

Strong reactions to undue warnings. Mentors were able to use the Snuggle's user dossier to identify false positives of counter-vandalism tools and direct their support to the user. Many participants felt the need to act immediately. For example, during the task evaluation, interviewee #10 remarked, "I don't see why this guy was reverted. [...] I don't see how this is vandalism. This is a false positive. I'm going to go ahead and categorize him as good faith." He then sent a message to the newcomer discussing warnings the newcomer received, offering his help and finishing his message with, "Keep up the good work." This example demonstrates how Snuggle brings visibility to a destructive part of Wikipedian's current socialization practices and the strong reaction that some Wikipedians have when they see an example of it.

[13] a space in Wikipedia designed to help new editors create encyclopedic articles

Complexity in categorizing. Some participants were uncomfortable coming to a conclusion about the value of another person without substantial interaction. #7: "I will not do that [categorizing] very fast to someone. Judging people or categorizing them before I've interacted with them or just based on a limited history is very hard." Other users were less concerned about the practice. #1: "I have no personal issues with that [categorization]. [...] You're going to form an opinion anyway." Some still saw the user dossier as a collection of actions. #12: "I suppose I felt like I was categorizing edit patterns rather than people."

Visibility of actions. We hoped that social translucence via the public and prominent recent activity feed (described in Social translucence) would enable effective peer-policing. However, some participants on the Snuggle forum were worried that this was too much visibility (e.g., "I would rather not have my activity on Snuggle be too accessible."), while others welcomed it (e.g. "Given that Snuggle is a promising new application, I believe that its activities should remain completely public for the time being.").

Who gets to use Snuggle? Participants raised concerns about who would be able to use Snuggle: specifically, could newcomers do so? For example, "I think that whatever decision we come to, the biggest thing is to not advertise the existence of Snuggle to newbies." and "One issue that I see a lot lately is that we have helpers and mentors in the various help spaces [...] who are not sufficiently experienced." Some interview participants brought up similar suggestions: #9 "I think that users should be autoconfirmed at least."[14]

CONCLUSION

In this paper, we discussed the design and evaluation of a novel user interface with the goal of solving a complex socio-technical problem in Wikipedia's counter-vandalism and newcomer socialization processes. To achieve this goal, we strategically redeployed concepts of quantification, formalization, and information processing used by Wikipedia's problematic counter-vandalism tools in ways that helped support better socialization and provoke critical reflection. These aspects worked together to build a system that insisted on an alternative account of newcomers than existing systems enacted. We also thought critically about assumptions embedded in existing systems, deeply situated ourselves in Wikipedia, and positioned our users as fellow collaborators, which made the design of Snuggle as a social information processing system more effective.

Implications for design

As a newcomer socialization support system, Snuggle is designed to solve a specific social information processing problem, which does not universally exist. In this frame, Snuggle-like systems are less useful when the reactions that newcomers receive are consistent with their behavior – but who gets to decide what is and is not "consistent"? This is typically an abstract third-wave question, but we found it useful to both designing a more effective system as well as to critique ideology. While existing counter-vandalism systems maximized efficiency by having one reviewer evaluating one edit, we found that by relaxing this goal and increasing the diversity of reviewers evaluating a holistic dossier revealed inconsistencies that were otherwise obscured.

Given that these issues arose in Wikipedia alongside automated evaluation systems, systems that use algorithms for similar purposes may have similar issues and can learn from Snuggle, both as a warning for designers of new systems and as a reaction to existing systems. For example, in massive open online courses, students number in the thousands. Automation is a common response to assessment, as evaluation by hand can become as impractical as it was in Wikipedia. Recalling the case of Wikipedia's counter-vandalism tools, such automated grading tools ought to:

Perform assessment in context. Give a holistic view of students' work so that graders can judge the current activity (e.g. a single answer) in context of their other work (e.g. the entire test).

Don't focus on negatives. Don't focus the grader's attention only on incorrect answers and mistakes. Bring equal attention to the rest of their work.

Support and scale existing practices. Afford the same types of nuanced feedback currently employed by non-automated graders in well-designed traditional courses.

However, if these automated systems fail in the same way that Wikipedia's counter-vandalism systems have, a Snuggle-like tool that brings attention to inconsistencies between the desirable characteristics of students pre-assessment (e.g. time spent on material) and the assessments they receive (grades) can direct support to good students who get bad grades in the short term and enable the types of visibility necessary to change minds about what good grading practice looks like in the long term.

[14] 'autoconfirmed' refers to a minimum experience threshold applied to newcomers who have registered their accounts at least 4 days ago have made at least 10 edits.

ACKNOWLEDGEMENTS

We gratefully acknowledge the support of the NSF under grants IIS 11-11201 and IIS 08-08692. We'd also like to thanks our Wikipedian collaborators – especially Subham Soni and Technical_13.

REFERENCES

1. Agre, P. *Computation and Human Experience*. Cambridge University Press, 1997.
2. Bardzell, S., and Bardzell, J. Towards a feminist hci methodology: Social science, feminism, and hci. In *CHI '11*, ACM (2011), 675–684.
3. Bentley, R., Hughes, J. A., Randall, D., Rodden, T., Sawyer, P., Shapiro, D., and Sommerville, I. Ethnographically-informed systems design for air traffic control. In *CSCW '92*, ACM (1992), 123–129.
4. Bødker, S. When second wave hci meets third wave challenges. In *NordiCHI '06*, ACM (2006), 1–8.

5. Bowker, G. C., and Star, S. L. *Sorting Things Out: Classification and its consequences*. The MIT Press, 2000.
6. Bryant, S. L., Forte, A., and Bruckman, A. Becoming wikipedian: Transformation of participation in a collaborative online encyclopedia. In *GROUP '05*, ACM (2005), 1–10.
7. Dimond, J. P., Dye, M., Larose, D., and Bruckman, A. S. Hollaback!: The role of storytelling online in a social movement organization. In *CSCW '13*, ACM (2013), 477–490.
8. Erickson, T., and Kellogg, W. A. Social translucence: An approach to designing systems that support social processes. *ACM Trans. Comput.-Hum. Interact. 7*, 1 (2000), 59–83.
9. Finkelstein, M. A., and Brannick, M. T. Applying theories of institutional helping to informal volunteering: Motives, role identity, and prosocial personality. *Social Behavior and Personality 35*, 1 (2007), 101–114.
10. Flanagan, M., Howe, D. C., and Nissenbaum, H. Values at play: Design tradeoffs in socially-oriented game design. In *CHI '05*, ACM (2005), 751–760.
11. Ford, H., and Geiger, R. S. Writing up rather than writing down: Becoming wikipedia literate. In *WikiSym '12*, ACM (2012), 16.
12. Geiger, R. S., and Halfaker, A. Using edit sessions to measure participation in wikipedia. In *CSCW '13*, ACM (2013).
13. Geiger, R. S., and Halfaker, A. When the levee breaks: Without bots, what happens to Wikipedia's quality control processes? In *WikiSym '13*, ACM (2013).
14. Geiger, R. S., Halfaker, A., Pinchuk, M., and Walling, S. Defense mechanism or socialization tactic? In *ICWSM '12*, AAAI (Palo Alto, CA, USA, 2012).
15. Geiger, R. S., and Ribes, D. The work of sustaining order in Wikipedia: The banning of a vandal. In *CSCW '10*, ACM (2010), 117–126.
16. Geiger, R. S., and Ribes, D. Trace ethnography: Following coordination through documentary practices. In *HICSS '11*, IEEE (2011), 1–10.
17. Goodwin, C. Professional vision. *American Anthropologist 96*, 3 (1994), 606–633.
18. Gould, J. D., and Lewis, C. Designing for usability: Key principles and what designers think. *Communications of the ACM 28*, 3 (1985), 300–311.
19. Halfaker, A., Geiger, R. S., Morgan, J., and Riedl, J. The rise and decline of an open collaboration system. *American Behavioral Scientist 57*, 5 (2013), 664–688.
20. Haraway, D. Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist studies 14*, 3 (1988), 575–599.
21. Harrison, S., Tatar, D., and Sengers, P. The three paradigms of hci. In *Alt. CHI '07* (2007), 1–18.
22. Hayes, G. R. The relationship of action research to human-computer interaction. *ToCHI '11 18*, 3 (2011), 15.
23. Hutchins, E. *Cognition in the Wild*, vol. 262082314. MIT press Cambridge, MA, 1995.
24. Irani, L. C., and Silberman, M. Turkopticon: Interrupting worker invisibility in amazon mechanical turk. In *CHI '13*, ACM (2013), 611–620.
25. Lam, S. T. K., Uduwage, A., Dong, Z., Sen, S., Musicant, D. R., Terveen, L., and Riedl, J. Wp:clubhouse?: An exploration of Wikipedia's gender imbalance. In *WikiSym '11*, ACM (2011), 1–10.
26. McNeely, B. L., and Meglino, B. M. The role of dispositional and situational antecedents in prosocial organizational behavior: An examination of the intended beneficiaries of prosocial behavior. *Journal of applied psychology 79*, 6 (1994), 836.
27. Morgan, J. T., Bouterse, S., Walls, H., and Stierch, S. Tea and sympathy: crafting positive new user experiences on wikipedia. In *CSCW '13*, ACM (2013), 839–848.
28. Muller, M. J. *Human-computer interaction: Development process*. CRC Press, 2003, ch. Participatory design: The Third Space in HCI, 165–185.
29. Musicant, D. R., Ren, Y., Johnson, J. A., and Riedl, J. Mentoring in wikipedia: a clash of cultures. In *WikiSym '11*, ACM (2011), 173–182.
30. Nardi, B. A. *Context and consciousness: Activity theory and human computer interaction*. The MIT Press, 1996.
31. Norman, D. A. *Cognitive artifacts*. Department of Cognitive Science, University of California, San Diego, 1990.
32. Sengers, P., Boehner, K., David, S., and Kaye, J. Reflective design. In *CC '05*, ACM (2005), 49–58.
33. Strauss, A. L. *Negotiations: Varieties, contexts, processes, and social order*. Jossey-Bass San Francisco, 1978.
34. Suchman, L. A. *Plans and Situated Actions: The problem of human-machine communication*. Cambridge university press, 1987.
35. Suh, B., Convertino, G., Chi, E. H., and Pirolli, P. The singularity is not near: Slowing growth of Wikipedia. In *WikiSym '09*, ACM (2009), 1–10.
36. West, A. G., Kannan, S., and Lee, I. Detecting wikipedia vandalism via spatio-temporal analysis of revision metadata? In *EuroSys '10*, ACM (2010), 22–28.
37. Yip, J. C., Foss, E., Bonsignore, E., Guha, M. L., Norooz, L., Rhodes, E., McNally, B., Papadatos, P., Golub, E., and Druin, A. Children initiating and leading cooperative inquiry sessions. In *IDC 13* (2013).