Accessible Online Content Creation By End Users

Katie Kuksenok, Michael Brooks

University Of Washington
Seattle, WA 98105
kuksenok@cs.washington.edu, mjbrooks@uw.edu

Jennifer Mankoff

Carnegie Mellon University
5000 Forbes Ave, Pittsburgh, PA 15206
jmankoff@cs.cmu.edu

ABSTRACT
Like most online content, user-generated content (UGC) poses accessibility barriers to users with disabilities. However, the accessibility difficulties pervasive in UGC warrant discussion and analysis distinct from other kinds of online content. Content authors, community culture, and the authoring tool itself all affect UGC accessibility. The choices, resources available, and strategies in use to ensure accessibility are different than for other types of online content. We contribute case studies of two UGC communities with accessible content: Wikipedia, where authors focus on access to visual materials and navigation, and an online health support forum where users moderate the cognitive accessibility of posts. Our data demonstrate real world moderation strategies and illuminate factors affecting success, such as community culture. We conclude with recommended strategies for creating a culture of accessibility around UGC.

Author Keywords: Accessibility, User-generated content

ACM Classification Keywords: H.1.2 **[User/Machine Systems]:** Human factors, Human Information Processing

INTRODUCTION
The ease with which end users can generate content is a defining feature of the World Wide Web. The ease of authoring HTML initially encouraged end user content creation. As more complex web technologies like CSS and Ajax have become ubiquitous, content editors provided by social media sites such as blogs and wikis have made it even easier to create content. Although the diversity and amount of user-generated content is increasing, much of it suffers from accessibility problems.

Web accessibility has been a rich research area for decades, including identification of accessibility problems (*e.g.,* [1] [2]), accessibility metrics [3], automatic improvement of website accessibility through content extraction [4] and

crowd-sourcing [5], and both server and client side content modification (*e.g.*, [6]). This work has tended to focus on content producers who have the resources to generate their own HTML/CSS and/or run studies of web accessibility. In contrast, end users generating content must deal with third party tools (such as WordPress.com or forum sites), and end users consuming content are dependent on both third party tool creators and end user content creators for accessible content. Studies of end-user content generation tools (*e.g.*, [7], [8]) point to the need for accessible authoring tools (*e.g.,* [8], [9]) but much of the work in this area has focused on the end user experience rather than the author experience.

Our contribution to this literature is a qualitative study of end user moderation of accessibility based on online data drawn from two existing communities. Our analysis, based on data drawn from Wikipedia's *Accessibility* project and from an online health forum, LymeSpace (a pseudonym), demonstrates the existence of a variety of mediation strategies that arise when authors are asked to make content more accessible. Our findings demonstrate that good authoring tools alone cannot ensure accessibility. The culture of a website and the training and personal circumstances of the author producing content impacts the accessibility of the content created. For example, users who have difficulty accessing content may have an intrinsic incentive to produce accessible content.

We demonstrate that accessible end user content creation requires structural support that goes beyond rules in the existing World Wide Web consortium (W3C) accessible content creation guidelines [10]. These needs include better tools, organizational support, support for author advocacy for other authors or readers, and support for the fluid way in which accessibility may be defined across different communities.

Next, we present a literature survey highlighting what is known about the accessibility issues and solutions currently being used across four main types of UGC: Wikis, Blogs, Forums and Social Networks. Following that we present our two case studies (Wikipedia and LymeSpace). We end with implications for design, proposing goals for UGC platform developers that focus on creating tools that can support a culture of producing accessible content.
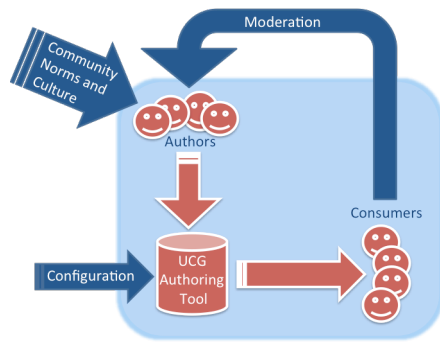
Figure 1: Content is generated by end users (*authors*) using a UCG authoring tool (such as a wiki or blog) and read by end users (*consumers*) who may also be authors. Existing guidelines focus primarily on the UCG authoring tool, and most studies of this domain focus on the consumer experience. Our studies of authors show that additional factors, such as moderation strategies and community norms and culture, can affect the accessibility of content.

AUTHORING ACCESSIBLE UGC

Breakdowns in accessibility involve three primary factors: UGC authoring tools, content *authors*, and content *consumers* (who may also be authors). Figure 1 summarizes these (in red). As our studies show, factors such as community norms and culture may also influence the accessibility of content. However the focus of existing literature and guidelines has primarily been on the consumer experience and the features of authoring tools.

Authoring tools that support accessible content are of broad interest, evidenced by the existence of the W3C's authoring tool accessibility guidelines (www.w3.org/TR/ATAG10/). These guidelines focus on both authoring tool accessibility and how authoring tools can support the production of accessible content. They emphasize the need to guide users in creating accessible content, assist authors in checking for accessibility problems, and promote accessibility features.

For example, developers of a wiki or a blog may hide the underlying HTML from a user with a simplified markup language or a WYSWYG editor. As a result the user may never see whether an image has an ALT attribute—a piece of metadata optionally associated with images in webpages, and used by screen readers. Additionally, structural elements of a site, like the navigation bar, are often pre-specified. The placement or existence of accessibility features may be completely outside of the control of the author of the content itself. For example, skip links, which allow users of screen readers to skip over navigation and reach the content of a page, are typically in the navigation area of a page, which UGC authors do not control. Authoring software may allow content authors to make use of headings appropriately, caption images using the ALT attribute, or edit the raw HTML. However, even then, authors may not be comfortable doing so, or may not know how to use these options to improve content accessibility.

Existing research has focused on the accessibility of specific types of UGC (*e.g.*, Wikis [8] and Facebook [11]). We summarize existing work in the context of four popular types of end user generated content: Wikis, blogs, discussion forums, and social network sites.

Wikis

Wikis, which range from personal places such as a Google Sites home page to vast public websites such as Wikipedia, are comprised entirely of content generated by users through iterative edits over time. The variation in wiki software is huge. Taras *et al.* [8] separate accessibility concerns into content viewing; content authoring (including accessibility of authoring tools and their capacity to support the production of accessible content); and actions that are specific to Wikis such as comparing versions of content, navigating recent changes, and so on. They point to the benefits of *group authoring* in helping to build accessibility "step by step."

Valtolina *et al.* describe a wiki, VisualPedia, designed to be accessible to content consumers [12]. It provides features such as image simplification, and was designed to meet as many accessibility guidelines as possible. Taras *et al.* propose a set of goals for Wiki accessibility that focus accessibility as the "conjunction of web authoring and viewing" [8]. An explicit focus on accessibility of authoring – and the capacity to support authoring of accessible content is an important goal that few have focused on.

Blogs

Blogs typically have only one or a few authors, and posts are usually organized chronologically and may be read over an RSS reader, by email, or on the website itself. Some blog sites also double as wikis of a sort.

The American Foundation for the Blind (AFB) studied two popular blogging platforms and 4 popular blogs [9]. Their survey identified accessibility issues in blog creation, most prominently inaccessible CAPTCHAs that affected account creation. Once an account is created, authors can create blog posts fairly easily. However, in the case of more structured, complex content, such as tables, authors typically must edit HTML directly to add table captions, and documentation for this is not consistently available.

Reading a blog presents far fewer problems than authoring a blog. However getting a sense of a blog can be difficult due to the large amount of content found on many blogs, and the AFB recommends allowing readers to receive content through alternative means (such as email) [9].

Discussion Forums

A discussion forum facilitates threaded discussions between members. Interestingly, although we found studies of forum content generated by individuals with disabilities (*e.g.* [13] [14] [15]), we found none that address the question of content accessibility. Rather, these studies contrast online forums with offline support groups in terms

of accessibility (online forums win here). Perhaps the accessibility of end user generated content is not an issue, because of the relative simplicity of websites such as forums and blogs in comparison to shopping sites, email sites, and so on. Or perhaps accessibility is seen as an inherent property of the site and is therefore not a topic of discussion for end users. However, in our case study, some forum users are concerned about accessible content generation, and forum software may present barriers to creating accessible content.

Social Network Sites

Social network sites refer to sites that allow a user to create a profile and share it with others. Most also support a variety of other activities (often based around sharing information with others). In a survey of visually impaired users, Leahy *et al*. [7] found that the accessibility of social websites was a major challenge for 80% of their participants. Some of the major accessibility challenges include lack of descriptions of non-textual content (a major issue for 65% of visually impaired respondents); misuse of tables (62%); badly designed online forms (70%); use of an inaccessible CAPTCHA for sign up (95%) and difficulty controlling video/audio players (63%) [7].

End users may not think about social network participation as authoring in the same way that blog or wiki authors do. As a result, *indirect* support for accessibility is important. For example, simplified mobile versions of online platforms may implicitly address accessibility problems. In a study comparing Facebook and Facebook Mobile, blind participants using the mobile interface with a screen reader had higher completion rates and shorter completion times on sample tasks [11]. Despite possible accessibility advantages of mobile UIs, Wentz *et al.* urge application developers not to rely on separate "accessible" and "inaccessible" interfaces as a solution [11]. Separate UIs provide distinct functionality, and the accessible version may not be kept up to date. An Application Programming Interface (API) can also support accessibility. For example, using Twitter's API an independent developer can replicate Twitter in an accessible fashion (*e.g.,* easychirp.com) [16]. The intentional minimalism of Twitter's features and an accepted diversity of wrappers for Twitter facilitate this.

Survey Conclusions

End user content generation tools are neither accessible to authors nor do they easily facilitate the production of accessible UGC. To meet basic accessibility guidelines, authors may have to edit HTML directly, which is overly burdensome for most authors. Despite the difficulty of authoring accessible UGC, consumers may still be able to read UGC by leveraging existing alternative interfaces such as email updates from blogs and mobile interfaces to social network sites, or by using a third party interface that builds on an API that encapsulates key features of a site, as in the case of Twitter. However, these work-around solutions

raise questions about a dual standard for access, and do not address many of the key recommendations in the W3C's guidelines for authoring accessible UGC.

Thus, accessibility must be addressed during content authorship. However, improved authoring tools is not sufficient. Accessibility can be hard for non-expert UGC authors to evaluate [2]. Like usability, it is best achieved using an iterative process, involving multiple people. This may require additional time and effort from authors who may or may not value accessibility. As a result, community norms and expectations, although not a focus of the work we surveyed, are likely to play a critical role in real world attempts to create accessible UGC.

The studies presented next show how community culture may encourage or discourage accessible content creation. To explore this issue we turn to examples where end users themselves encourage accessible content creation. We call this process *accessibility moderation*. Although end-user moderation of accessibility may be unusual, studying it can provide valuable insights into what it takes to generate accessible UGC. The next two sections explore two exemplary communities of practice we uncovered.

CASE STUDY ONE: WIKIPEDIA

Wikipedia, an online encyclopedia written by volunteer authors, has a complex social structure and mechanisms for coordination among editors [17]. All Wikipedia pages have a *Talk* page where discussions relevant to the content of that page take place. Our analysis focused on *Talk* pages that included discussions of accessibility issues. We will refer to the specific author of a page as the *author* or (A#), other editors involved in accessibility discussions as *moderator* (M#)*,* and content consumers as *readers.*

Wikipedia has a formal channel for handling accessibility problems. The *AccessibilityDispute* template can be added to a content page by any moderator when they believe that there is an accessibility problem with a page. This adds a message at the top of the content page declaring that "this article's accessibility is in question." The specific problems are discussed on the content's associated *Talk* page, where moderators or authors can work out a resolution for the problem. Wikipedia also has an *Accessibility* Project, where meta issues about how to handle accessibility are discussed, and individuals can post accessibility-related questions.

In addition to content itself, authors can create templates, which function as a sort of intermediary between authors and the authoring tool. These templates are at times themselves the subject of accessibility discussions.

Wikipedia Method

We used a snowball sample to collect data for our analysis of Wikipedia, starting with the *AccessibilityDispute* template page and the *Accessibility* project and talk pages. These three pages are central to the work being done around accessibility on Wikipedia. Wikipedia's *WhatLinksHere* feature lists all of the links to a page. Using

WhatLinksHere, we downloaded the text from 68 pages linking to the *AccessibilityDispute* template and 50 pages linking to the *Accessibility* project and talk pages as of September 2012. After removing duplicates and pages whose content was not accessibility related, we were left with 29 pages in our final sample. Some had a single paragraph of text relating to accessibility, others had many pages of text containing many different accessibility discussions. Across all 29 pages, we identified 179 accessibility discussions involving 82 contributors. Of these, ten contributors currently have their name listed on the main Accessibility project page (these ten participated in 33% of the discussions, and are referenced as moderators M1-M10 when quoted). The remaining participants are referred to as authors A11-A82 when quoted.

The data was coded by the third author. Through open coding, we generated 69 codes, which were then grouped into 22 categories using selective coding. Based on these categories, we identified central themes such as the importance of authority. We present our data using representative quotes, modified to preserve anonymity.

Wikipedia Case Study Results
Wikipedia's Accessibility project, run by 32 moderators, centralizes efforts to make Wikipedia more accessible. The scope of accessibility problems as defined by the project is based on the WCAG 2.0 guidelines, aiming for level AA compliance. Although accessibility in Wikipedia is defined according to accepted national standards, we will show that the definition of accessibility was still disputed.

In practice, most problems discussed in our data set had either to do with visual elements (9 discussions about color, font size, animations, ALT tags, *etc.*); tables (5 discussions); or internationalization (2 discussions).

We identified several roles assumed by moderators within the Accessibility project. When moderating *content accessibility,* a moderator provides advice, guidance, or critique relating to an author's content. When moderating *Wikipedia accessibility*, the moderator is influencing meta-issues such as Wikipedia policy. Guideline design and tool configuration were also issues of concern for moderators.

Moderation Styles and Outcomes
A moderator officially initiates moderation of content by adding an *AccessibilityDispute* template to a page. However, other kinds of moderation are also common, such as educating authors (by restating or explaining guidelines; encouraging authors to download a screen reader for testing; or linking to articles on accessibility); and directly editing content. Sometimes an accessibility dispute arises without moderation and is then brought to the attention of the *Accessibility* project *via* a post to a project page.

Even moderators who were already oriented toward accessibility issues struggled to accomplish it ("didn't realise this accessibility stuff would be so difficult" [M15]).

As a result, moderators or authors sometimes requested experts to step in and help with a discussion. For example:

> "Could an expert explain why it is important not to skip heading levels at [link to discussion]….I feel I cannot do it justice" [A11].

> "After [another author reverted my change] I was wondering if you could provide advice about placement [of visual element] because I could not find helpful information." [A15]

More experienced moderators can provide evidence and authority to support an author's position and help make a change successful.

The Role of Authority. Moderators tended to have more authority than other authors, but their background also affected their authority. For example, one moderator pointed out that "since [a person with a disability] began to help with improving accessibility, he has been listened to more than all of us" [M10]. M10 actively encouraged other users with disabilities to participate in moderation.

Authority is important for moderators, because sometimes authors did not want to comply with moderator requests. Although moderators sometimes idealized the process ("It isn't that hard to make sure a page is accessible - just use conventional formatting and heading standards" [M2])*,* in eight discussions, moderators and authors spent significant time on tensions between accessibility goals and other editing goals. For example, authors/moderators worried that accessibility would require them to remove valued moving images; increase markup (in violation of another Wikipedia policy); and change page titles (affecting google search results). Additionally, some authors simply disagreed with or reverted changes made by moderators.

Tools as Moderators. Moderators occasionally discussed editing intermediate representations and tools through the provision of CSS, templates, Flash plugins, and so on. Thus, discussions turned at times to the need for bots (automated software scripts) that could fix some accessibility errors, the use of CSS to provide unified templates solving problems for many authors at once, and other technical fixes that are relatively sophisticated compared to what one might find in other UGC settings.

Moderating Wikipedia Accessibility
Wikipedia moderators were concerned not only with the accessibility of specific content pages, but also with fostering a culture of accessibility across Wikipedia, making Wikipedia more accessible to disabled authors.

Creating Appropriate Structures for Accessibility. Some moderators advocated the adoption of a Universal Access policy, an issue that lies at the intersection between the goals of the *Accessibility* project and structural aspects of Wikipedia that influence authorship. Between 2007 and 2011, these moderators discussed and drafted a proposal to the WikiMedia foundation to modify the Manual of Style. The Manual of Style currently has an entire page on accessibility, an apparent outcome of these efforts. They

also discussed marking articles as accessible, but the idea was discarded because, "It would be as if people went and marked templates as 'This is an article with a Neutral POV' and 'This is an article that is well referenced' [A31, discussing with M5]. Finally, moderators discussed creating a "Cleanup Template for accessibility" by [M3] that would be available on Wikipedia:Template messages/Cleanup. Currently, the AccessibilityDispute template discussed earlier is used.

Defining Accessibility. The creation of accessibility guidelines was itself a topic of discussion. A combination of changing standards (such as the impact of the introduction of W3C's accessibility guidelines 2.0) and a wish to support a variety of impairments led to dispute over what exactly should go in guidelines. For example, M16 comments that "[M10]'s view is that providing ALT tags is all you can do…. I would prefer to [ask for more] because without a mouse some people can't access the alt text." There was no agreement on how accessibility is to be verified, aside from several mentions of using tools such as a screen reader to find problems or check solutions.

Moderator as Advocate. Moderators also discussed participating in advocacy on behalf of specific disabled authors. Some of this was moderator initiated. For example, moderators heard about and discussed advocating for a blind user who could not create an account because of an inaccessible CAPTCHA, and a user with a disability affecting their participation style who was asked to stop editing pages. Moderators were also concerned by a hostile environment created by unfriendly discussions on a talk page. Sometimes authors asked moderators to help. For example, A81 posted on a moderator's *Talk* page: "I feel hounded by some administrators because of my [disability]" [A81]. However, moderators were concerned that not everyone would speak up to them, and discussed pro-actively addressing situations. This raises questions of autonomy from a disability studies perspective [2]. Moderators seemed to understand this and commented on the difficulty of advocating for other authors:

> I could be [an author advocate] but what exactly would I say? "Don't ban this person, they have [disability]."? That might not work. [M1]

> No, how about "Hi, this user has identified herself as being [disease], which is probably why [this problem arose]. Before we ban her … she seems to have her heart in the right place and has a lot to offer our community. Obviously, this would not apply to users who are violent or angry…" [M4]

Discussions of moderator-initiated author advocacy in the data lacked any reference to whether it had actually been done successfully or whether authors would want this.

Summary
In Wikipedia, moderation of accessibility occurs along formal channels with respect to content, as well as through informal channels and advocacy efforts. Our findings illustrate the importance of subtle issues such as moderator

authority, automation, and structural support for accessibility. Advocacy, though not directly involved in accessible content generation, contributes to a culture of accessibility, an important prerequisite for successful moderation.

CASE STUDY TWO: HEALTH FORUM
Our second case study focuses on a very different domain: A discussion forum, LymeSpace (pseudonym), used by individuals with Lyme disease. LymeSpace is one of the most established and high volume social support sites for individuals with Lyme disease. It includes a large and diverse set of patients and caregivers, a subset of whom have a cognitive impairment. Thus, this group of people is motivated to consider the needs of individuals with cognitive impairments when creating content, but authors may not all have a cognitive impairment themselves. LymeSpace had a total of over 70,000 threads with 600,000 replies and about 8,500 users at the time of our analysis. A small population of users post frequently and respond to many new user posts.

We selected LymeSpace because its users actively moderate content to improve cognitive accessibility. Lyme disease is the most common vector-born disease in the United States today [18] and may cause a complex set of persistent physical and neurological symptoms [19]. One in particular, typically referred to as "brain fog" by patients, led participants in our study to modify how they wrote forum posts. Brain fog describes cognitive impairments that may encompass: fatigue; short-term memory impairment; poor concentration; difficulties in formulating ideas; difficulties in word finding; difficulties in judgment; and difficulties in naming objects [19]. Additional issues found in our case study that affected the use of online content include: easy distraction when faced with visuals ranging from animated advertisements to bold text; difficulty remembering the start of a long passage by the end of the same passage; and getting lost among multiple pages of content. Cognitive impairments can be extremely heterogeneous, and can affect web accessibility [20].

LymeSpace Qualitative Method
We collected all of the posts made on LymeSpace over the period from 2001 through 2009 (almost 150,000 posts and 67,000 threads). Using this data, we conducted an initial, exploratory qualitative study, where we discovered several types of end-user moderation. Examples of such moderation, as well as some responses to it, demonstrate the diversity and culture around accessibility moderation in this forum. Additionally, we analysed the incidence and effect of one particular kind of moderation we found across 300 discussion threads using automatically extracted quantitative metrics.

For the qualitative portion of our analysis, we randomly selected six forum members from this data set. We specifically sampled from forum members who produced

few (1 to 5) or many (6 or more) posts which led to further discussion (which we operationalized as any post having at least one response). Once the six forum members were selected, we randomly selected a maximum of 50 threads started by each of them. We then opportunistically added any moderators who participated in these threads to our user set, and added further examples of moderation from each moderator we identified. Finally, we removed discussions from the qualitative data set that had not been moderated. At the end, our data set had 97 discussions, involving three different moderators.

For the analysis of this data, two authors independantly read all of the data, and grouped the data into two sets of categories: styles of moderation and moderation outcomes. Within these top level categories, sub-categories were identified in a bottom-up fashion. Both authors discussed the resulting sub-categories until they were in agreement.

For the quantitative portion of our analysis, we automatically identified 300 threads where an author had been moderated (asked to make a post more accessible) by matching against several key phrases present in one prolific moderator's moderating posts (*eg,* "please break up your wordy sentences into one sentence paragraphs"). In each case, after one or more posts, the moderator requested that content be made more accessible, following which 240 of the 300 authors responded and 60 did not. Next, we randomly selected 300 other threads. This selection procedure was weighted by thread size, in number of posts, so longer threads were more likely to be selected. Thus, our final sample contained 240 threads where the author *responded to moderation*, 60 where the author *did not respond to moderation,* and 300 *random* threads.

For the analysis of this data, we used easily calculated metrics that corresponded to specific accessibility issues defined by the community as important through moderating posts. Our focus was on metrics that would be easily visible to content consumers including the number of characters per block (paragraph), and the frequency with which all caps, punctuation, and bold text were used. For clarity, we reiterate the metrics with precise formulations when we present the results of this analysis.

Moderation Approaches and Outcomes

Accessibility in LymeSpace was defined in terms of the cognitive impairments experienced by its members. The community defined the meaning of accessibility through its moderation posts, and we used their definition in our analysis. To them, accessibility meant addressing the impact of brain fog, using: (1) short paragraphs to avoid big blocks of text; (2) double-spacing between paragraphs to provide more visual space around blocks of text; and (3) generally making text simpler and easier to read.

LymeSpace has a less formal, more varied, set of strategies for managing accessibility than Wikipedia. As summarized in Table 1, two of the three moderators in our data set were

Table 1: Summary of moderators we analyzed and their moderation styles

Name	Style	Comments
M1	Educational; Positive Re-enforcement; Reformatting	Power user and moderator
M2	Educational	Casual user; Less than 50 posts total
M3	Reformatting; Teaching by example	Casual user; Taught by example (reposted other' posts with correct formatting)

primarily content consumers with no particular authority or role beyond contributor in the forum. The third, M1, was a p ower user who responded to almost every new user on the forum and cared deeply about setting community standards for accessibility. We explore these moderators' different styles of moderation and users' responses to moderation.

Styles of Moderation

Moderators employed several strategies when intervening to moderate the accessibility of content. M1 sometimes used a post containing guidelines to educate new posters about the standards of the community, including guidelines for accessible content. Interestingly, the guidelines themselves were not always accessible. For example, M1 sometimes used ALL CAPS for emphasis, a violation of her own guidelines.

M1 also used positive reinforcement to encourage compliance. For example, one of her welcome posts mentioned that she liked how the original poster visually separated text into individual thoughts and made it simple.

Non-power-users also engaged in moderation. An example is this post by M2, who had less than 50 posts overall:

> I wasn't able to read your entire post.

> Maybe you could go back and break up your paragraphs so they are easier to read? [M2]

As opposed to the sanctioned guidelines provided by power users like M1, these posts were more casual and reactive in nature.

Moderation was not always explicit. For example, M3 reformatted an inaccessible post and reposted it "with added ••• at a couple points for emphasis" [M3]. M3's approach of reformatting and reposting content represents a more direct form of help than M1's critiques and advice.

M3 also wrote posts that referred to her own experiences with Lyme disease in support of M1's guidelines. For example, in one post M3 included a medical text that M1 had reformatted in an earlier thread. M3 gave credit to M1 for making the text easier to read. Although this was not a moderating post, it helped to reinforce the ethics of the community with regard to cognitive accessibility.

Responses to Moderation

Moderation met with varying success among the posts that we examined. In the best case, we observed *uptake*, where a poster adapted her style after receiving a moderating prompt of some kind. For example, the first thread that A2 started began with a fifteen hundred word text. Three different moderators responded (two were educational, a third was a reformatted version of the original post). A2 then indicated approval by replacing the original post body with the reformatted version.

In another example, A1's posts included ellipses, double punctuation, and rogue capitalization. After M1 suggested adherence to the posting guidelines, A1's posting habits improved. Most notably, she had split up her thoughts visually into smaller blocks of text. For example, a portion of her later post read:

> Thanks everyone for your help. I know my sx are lyme and ms symptoms. Hopefully being that I've had two Normal MRI's it isn't MS.
>
> I ordered my Igenex test yesterday.
>
> Do most people not test positive on reg WB? Like I said I only had Reactive bands at IGG 31,41 on Quest. [A1]

In some cases, a poster may explicitly respond to moderation. For example, in response to moderation A3 edited a post to be more accessible and wrote:

> I apologize for my original post. I thought it was just me… losing my place and reading the same lines over and over…I'm still learning…hope this is better [A3]

Sometimes, users responded negatively to moderation. For example, A5 was a new member who tended to write one paragraph posts without capitalization, and with non-standard punctuation (*i.e.,* an excess of ellipses). A5 posted three times on LymeSpace. M1 responded to one of A5's posts by welcoming her, without a moderating section to the post, and directed her to check her private messages for the response. A5 continued to request information with increasing desperation in 10 more posts, including one similar to her initial question, which M1 believed she had answered. Eventually, A5 commented on M1's moderation:

> ...please be more supportive. Thanks. [A5].

Not all attempts at moderation succeed. Users may be unsure how to interpret the welcome message and advice about posting guidelines, or how to create a post that is likely to get an answer. This may be further confounded by the overall stress associated with, possibly, having just received a difficult diagnosis. We cannot know why a user does not respond to moderation. However, when the only response to a plea for help is a request to increase accessibility, this might drive a user away from the site.

Influence of the Authoring Tool on Post Accessibility

LymeSpace suffers from usability problems. Methods for posting new threads and retrieving private messages are unintuitive. As a result, some posts asking users to edit their message or change formatting include detailed and lengthy instructions. For example, A6 provides a method for using Firefox scripting to disable bolding. Without his script, quoted text is displayed in bold, which is inaccessible for readers and goes against group guidelines.

LymeSpace authors can edit their posts after they are posted. For example, if another reader points out that a post is difficult to read, some posters may edit that original post. However when something is reposted by a moderator, the original poster loses control over her own potentially sensitive content, as no user without special administrative privileges can delete or edit another user's post.

Summary

To summarize, our qualitative data indicates that cognitive accessibility on LymeSpace is something that even casual users take seriously. Authors may respond positively and at times even apologetically to requests to improve their posting habits. Even occassional posters re-inforce these goals, perhaps in part because they themselves benefit when authors post more accessible text. However, a limitation of our analysis is the small number of authors and moderators that we studied. The next section addresses this by using quantitative methods to explore community norms acrosss 600 separate threads.

Incidence and Effect of Moderation

To further explore moderation incidence and its impact on author behavior, we used 600 threads (240 moderated by M1 with further responses, 60 moderated by M1 without response, and 300 random threads not moderated by M1). We calculated metrics that characterized accessibility based on the types of changes found in our qualitative data and the specific changes asked for in M1's moderating post:

Characters per block: Number of alphanumeric or punctuation characters, in each visually distinct block (defined as piece of text two or more newlines away from each adjacent piece of text).

Capitalization: Percent capitalized characters per block
Punctuation: Percent punctuation characters per block
Bold: Percent characters per block enclosed in bold tags

For authors who responded to moderation, we calculated separate mean scores for their posts before and after moderation. For authors who did not respond, and in randomly selected threads, we calculated a single score. We had two hypotheses, based on our qualitative data:

H1—Moderation improves compliance. *Each metric would decrease from before to after moderation.*

H2—Random posts are in between. *Random posts would score worse (higher) on each metric than post-moderation content. Random posts would score better on each metric than pre-moderation content.*

Because the distributions for each measure were not normal, we used non-parametric tests in our analysis. A
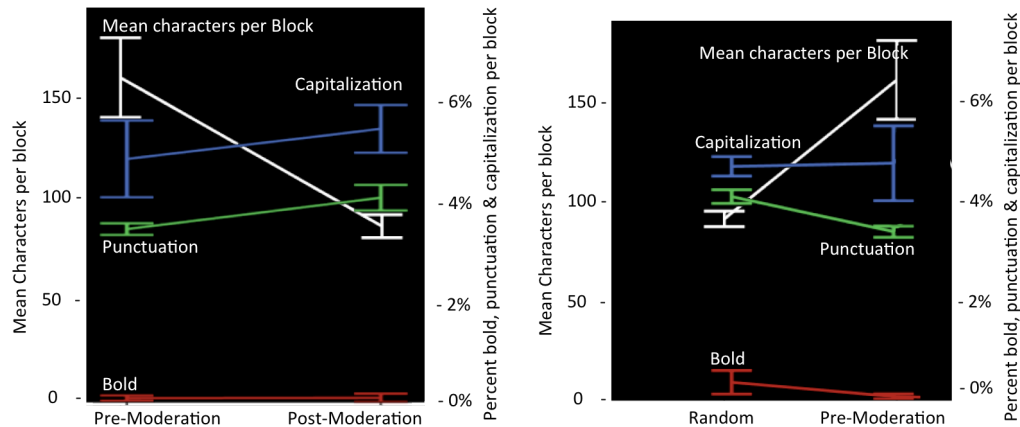
Figure 3: (LEFT) Mean characters per block were twice as high prior to moderation as after moderation (white line). Other lines: Percentage of characters that are bold did not change, but punctuation and capitalization increased. **(RIGHT)** Mean characters per block in randomly selected, unmoderated posts (representing the group norm) is half that of posts about to be moderated. Other lines: Percentage of characters that are bold and punctuation is slightly higher in random posts than posts that are about to be moderated. Percentage capitaized characters is about equal.

Wilcoxon Signed Ranks test was used to compare matched pairs of measurements (before and after moderatio n). This test was done over the 300 moderated samples, of which 240 had *before* and *after* measures, and 60 only had *before* measures. For comparing random threads to moderated threads, a Mann-Whitney U test was used. We corrected for multiple comparisons using the Bonferroni method.

Moderation improves compliance, but not in the way we expected. As expected, responses to moderation modified their characters per block, reducing characters per block to half that of posts about to be moderated (white line, Fig. 2 left; Z=-7.797, p<0.001). Randomly selected threads also had almost half as many characters per block as posts that were about to be moderated (white line, Fig. 2 right; Z=-7.797, p<0.001). In other words, after moderation, authors conformed to the group norm, as represented by the random threads. For characters per block, *H1—Moderation improves compliance* is supported, but *H2—Random posts are in between* is rejected.

This effect is reversed for punctuation: The group norm is to have more punctuation than found in threads just prior to moderation. The use of bolded characters was less salient (though the random posts used bold significantly more than the pre-moderatino posts). Capitalization increased slightly as a result of moderation.

Thus, *H1—Moderation improves compliance* is strongly supported for characters per block, and the opposite effect (though smaller) is seen for punctuation and capitalization. Although ellipses and all-caps are a problem, unpunctuated run-ons may potentially be a worse or more common one. Also, M1 was seen to use all caps to signpost lengthy posts, and this behavior may be emulated by modarated authors, even though it is not explicitly requested. *H2—Random posts are in between* is rejected, but this seems to be because group norms are *better* than we expected. If

random posts are *already* compliant with accessible content goals this could mean that most unmoderated posts are not moderated *because they do not need moderation.*

This analysis is limited in its focus on a single moderator. Our qualitative findings show that there are many other kinds of moderation, such as teaching by example, that were not addressed in this analysis. In future work, we would like to explore their impact as well.

DISCUSSION AND IMPLICATIONS FOR DESIGN

The case studies presented focus on two very different situations. In the first case, we explored a very large and established Wiki (Wikipedia) with a complex management structure and many authors that work together to create content for a larger audience. Moderators made use of accepted standards drawn from W3C guidelines, and operated within a variety of structural and cultural forces in their attempts to improve accessiblity.

In the second case, we explored a small community forum with minimal management structure, creating content pimarily for themselves. Accessibility was defined by the community based on the typical cognitive impairments experienced by its members.

Despite the differences in these communities, both include end users who are motivated to make content more accessible, yet limited by the tools available to them. In both cases, we can see that standard definitions of accessiblity do not suffice; the meaning and practice of accessible content creation is negotiated. Technology could support this practice in several ways:

Support for accessibility should be proactive: Existing literature as well as our case study indicate that barriers to accessible content persist despite the best of intentions, so the software itself should do what it can to automatically meet accessibility guidelines. As we saw in Wikipedia accessibility moderation practices, many problems may be

addressed by automated, technical solutions that can often be articulated by some of the end-users involved. Although the templates and bots used on Wikipedia are helpful, we believe that more could be achieved by explicitly engaging users in the process. For example, Burke *et al.* created a tool that helps users to compose posts more likely to receive a response by suggesting modifications before the post was finalized [21]. Our quantitative analysis of LymeSpace suggests that a similar tool could be created to help authors improve the accessibility of their text before it is posted. A similar approach could also be used to prompt users for *alt* text, table descriptions, appropriate use of headers, readable text, and so on.

The definition of accessibility should be negotiable: The accessibility issues that matter to end users may vary with each community and may include specific issues not covered by more generic accessibility guidelines [10]. As we saw in the Wikipedia study, although moderators leveraged existing standards, even these standards were not static. Furthermore, accomodations for some disabilities (such as cognitive impairments) are not well defined in standard guidelines. Despite this, as on LymeSpace, participants knew what would address their specific needs. Clear and easy control over parameters that relate to accessibility, such as formatting properties, including the use of bold for quoted text, could help to support accessibility. Additionally, UGC-driven websites would benefit from allowing users to have a formal channel for discussion of accessibility concerns, such as a designated thread on a forum, or the talk page we saw on Wikipedia. Although many users may not know how to adjust the presentation of information to make it more accessible, bringing such options to the forefront and enabling active collaboration could make moderation more effective.

UGC platforms should leverage varying expertise and ability: The diversity of experience and motivation of end users presents an opportunity for UGC sites. The lengths to which users will go to achieve accessibility is impressive. In our study, this included switching browsers; writing scripts and tutorials to fix browsers; and writing new Wikipedia Templates. Developing custom code to share with other users or a custom interface using a platform's API represents an extreme case of the negotiated standards mentioned above. UGC platforms could scaffold this sort of involvement more explicitly by adding API support, providing access to CSS, and enabling the creation of forms to intermediate between authors and UGC plafforms.

Allow users to take on a diversity of roles: Wikis and blogs, which may have multiple authors, can leverage the varying motivation and expertise of groups of authors to achieve greater accessibility than would be possible by any one author alone. UGC platforms can encourage readers to flag problematic content, and authors to indicate their areas of expertise for fixing problems. This can allow matching authors with problems, or provide data that could be used

in the long run to automate the detection or correction of inaccessible content. Support for emergent standards, such as in LymeSpace, and emergent roles, such as the advocacy role in Wikipedia, is also important. This could help sites customize their definition of accessibility to match that of the community of users being supported.

Connect formal channels with informal practice: A formal channel for discussing accessibility problems and solutions exists in Wikipedia, but not LymeSpace. The informal moderation on LymeSpace was epehmeral and decentralized, but nevertheless effective in setting group norms. The changes that are called for in more formalized channels on Wikipedia talk pages are more persistent, but arguably less integrated into authoring culture. Strategies such as involving authors with disabilities in the discussion and educating authors when they first begin to produce content, are important complements to formal channels.

CONCLUSIONS AND FUTURE WORK

User generated content is an important, growing source of information online. Unlike most other web content, UGC authors and consumers often overlap, and typically have little web design or accessibility experience. We cannot expect end users to become experts in accessibility, make use of complex authoring tools, or investigate the accessibility of their content. Instead, a combination of automated and active support for accessibilty can enable end users to meet accessibility goals, as they define them.

Basic standards of documentation and usability are not met by many UGC platforms. It is not enough for UGC platforms to document the importance of accessible content creation. Such documentation exists, yet the literature demonstrates that it does not result in accessible content (*e.g.,* [7]*,* [9]). Users are unlikely to seek and read documentation without better support and incentives. Providing direct support in the authoring interface for accessibility is also not enough. Support must at a minimum be useable (*i.e.* not require authors to directly edit the HTML of their content) and discoverable (*i.e.* define or explain of terms such as "Alternative Text" on image dialogues). Authoring tools for some of today's most popular sites fail to meet this standard.

However documentation and usability of UGC creation tools are unlikely to be enough. If UGC is to be truly accessible, we also need to develop a better understanding of the role of culture and authorship abilities in the process of content creation. This suggests additional opportunities for UGC tools to amplify the accessibility of UGC that have not been discussed in the literature.

We have presented two case studies of UGC communities with a culture of self-moderating accessibility. We investigated Wikipedia, which deals with the full range of accessibility issues but focuses on access to visual information and navigation. We also studied LymeSpace, whose members cope with cognitive impairments. In both

cases, we showed how a small group of people marshaled accessible content and educated authors. We explored issues such as authority, the definition of accessibility, uptake or rejection of moderation. Wikipedia moderators became engaged in advocacy as well as moderation, demonstrating the need to support a variety of roles. Accessibility on both sites was influenced not only by decisions of posters but also by community culture, the site design, usability issues, and technology constraints.

Further work should examine the full range of UGC platforms, including not only blogs, wikis, social network sites, and discussion forums, but also data collection and sharing platforms (such as FourSquare), content management systems, and email programs. These communities reflect the sustained effort of individuals with a common goal. UGC platforms should channel this effort constructively toward the important goal of making content accessible to a broad audience.

ACKNOWLEDGMENTS

REFERENCES

1. Ivory, M. Y., Mankoff, J. & Le, A. (2003) Using automated tools to improve web site usage by users with diverse ability. *Info. Tech. & Society*, 3(1):195-236

2. Mankoff, J., Fait, H., & Tran, T. (2005) Is your web page accessible? A comparative study of methods for assessing web page accessibility for the blind. In *CHI*, 41-50.

3. Vigo, M. & Brajnik, G. (2011)Automatic web accessibility metrics: where we are and where we can go. *Interacting with Computers*, 23(2):137-155.

4. Gupta, S., Kaiser, G., Neistadt, D., & Grimm, P. (2003) DOM-based content extraction of HTML documents. In *WWW*, 207-214.

5. Takagi, H., Kawenaka, S., Kobayashi, M., Itoh, T., & Asakawa, C. (2008) Social accessibility: Achieving accessibility through collaborative metadata authoring. In *ASSETS*, 193-200.

6. Fairweather, P. G., Hanson, V. L., Detweiler, S. R., and Schwerdtfeger, R. S. (2002) From assistive technology to a Web accessibility service. In *ASSETS*, 4-8.

7. Leahy, D. & Broin, U. Ó. (2009) Social networking sites and equal opportunity: The impact of accessibility. In *22nd Bled eConference: eEnablement: Facilitating an Open, Effective and Representative eSociety*, 41.

8. Taras, C., Siemoneit, O., Weißer, N., Rotard, M., & Ertl, T. (2008) Improving the accessibility of Wikis: A basic analytical framework. In *ICCHP*, 430-437.

9. American Foundation for the Blind. *Is blogging accessible to people with vision loss?* Archived by WebCite® at http://www.webcitation.org/67PvGjdh6.

10. Chrisholm, W., Vanderheiden, G., & Jacobs, I. (1999) *W3C Web Content Accessibility Guidelines 1.0*.

11. Wentz, B. & Lazar, J. (2011) Are separate interfaces inherently unequal? An evaluation with blind users of the usability of two interfaces for a social networking platform. In the *2011 iConference*, 91-97.

12. Mesiti, M., Ribaudo, M., Valtolina, S., Barricelli, B. R., Boccacci, P., & Dini, S. (2011) Collaborative environments: Accessibility and usablity for users with special needs. In Pardede, Eric, ed., *Community-built Databases: Research and Development*. Springer.

13. Braithwaite, D. O., Waldron, V. R., and Finn, J. (2009) Communication of social support in computer-mediated groups for people with disabilities. *Health Communication*, 11(2):123-151.

14. Maloney-Krichmar, D. and Preece, J. (2005) A multilevel analysis of sociability, usability, and community dynamics in an online health community. *TOCHI*, 12(2):201-232.

15. Finn, J. An exploration of helping processes in an online self-help group focusing on issues of disability. *Health Soc. Work*, 24, 3 (1999), 220-231.

16. Ellis, K. & Kent, M. (2010) Community accessibility: Tweeters take responsibility for an accessible Web 2.0. *Fast Capitalism*, 7(1) http://www.uta.edu/huma/ agger/fastcapitalism/7_1/elliskent7_1.html.

17. Kittur, A. & Kraut, R. E. (2008) Harnessing the wisdom of crowds in Wikipedia: Quality through coordination. In *CSCW*, 37-46.

18. Weintraub, P. (2008). *Cure unknown: Inside the Lyme epidemic*. NY: St. Martin's Press.

19. Cairns, V. & Godwin, J. (2005) Post-Lyme borreliosis syndrome: A meta-analysis of reported symptoms. *International J. of Epidemiology*, 34(6):1340-1345.

20. WEBAIM (Web Accessibility in Mind). *Cognitive Disabilities: Introduction*. 1999-2012.

21. Burke, M., Joyce, E, Kim, T., Anand, V., & Kraut, R. E. (2007) Introductions and requests: Rhetorical strategies that elicit response in online communities. In *Intern. Conf. on Communities & Technologies*, 21-40.