Editorial and Technological Workflow Tools to Promote Website Quality

Emily G. Morton-Owens

Editor's Note: This paper is adapted from a presentation given at the 2010 LITA Forum

Library websites are an increasingly visible representation of the library as an institution, which makes website quality an important way to communicate competence and trustworthiness to users. A website editorial workflow is one way to enforce a process and ensure quality. In a workflow, users receive roles, like author or editor, and content travels through various stages in which grammar, spelling, tone, and format are checked. One library used a workflow system to involve librarians in the creation of content. This system, implemented in Drupal, an open-source content management system, solved problems of coordination, quality, and comprehensiveness that existed on the library's earlier, static website.

Today, libraries can treat their websites as a significant point of user contact and as a way of compensating for decreases in traditional measures of library use, like gate counts and circulation.[1] Websites offer more than just a gateway to journals; librarians also can consider instructional or explanatory webpages as a type of public service interaction.[2] As users flock to the web to access electronic resources and services, a library's website becomes an increasingly prominent representation of the library.

At the New York University Health Sciences Libraries (NYUHSL), for example, statistics for the 2009–10 academic year showed 580,980 in-person visits for all five locations combined. By comparison, the website received 986,922 visits. In other words, the libraries received 70 percent more website visits than in-person visits.

Many libraries conduct usability testing to determine whether their websites meet the functional needs of their users. A concern related to usability is quality: users form an impression of the library partly based on how it presents itself via the website. As several studies outside the library arena have shown, users' experience of a website leads them to attribute characteristics of competence and trustworthiness to the sponsoring organization.

Tseng and Fogg, discussing non-web computer systems, present "surface credibility" as one of the types of credibility affecting users. They suggest that "small computer errors have disproportionately large effects on perceptions of credibility."[3] In another paper by Fogg et al., "amateurism" is one of seven factors in a study of website credibility. The authors recommend that "organizations that care about credibility should be ever vigilant—and perhaps obsessive—to avoid small glitches in their websites. . . . Even one typographical error or a single broken link is damaging."[4]

Everard and Galletta performed an experimental study with 232 university students to discover whether website flaws affected perception of site quality and trust. Their three types of flaws were incompleteness, language errors (such as spelling mistakes), and poor style in terms of "ambiance and aesthetics," including readable formatting of text. They discovered that subjects' perception of flaws influenced their judgment of a site being high-quality and trustworthy. Further, they found that the first perceived error had a greater negative impact than additional problems did, and they described website users as "quite critical, negative, and unforgiving."[5]

Briggs et al. did two studies of users' likelihood of accepting advice presented on a website. Of the three factors they considered—credibility, personalization, and predictability—credibility was the most influential in predicting whether users would accept or reject the advice. "It is clear," they report, "that the look and feel of a web site is paramount in first attracting the attention of a user and signaling the trustworthiness of the site. The site should be . . . free of errors and clutter."[6]

Though none of these studies focuses on libraries or academic websites and though they use various metrics of trustworthiness, together they point to the importance of quality. Text quality and functional usability should be important to library website managers. Libraries ask users to entrust them to choose resources, answer questions, and provide research advice, so projecting competence and trustworthiness is essential.

It is a challenge to balance the concern for quality with the desire to update the website frequently and with librarians' workloads. This paper describes a solution implemented in Drupal that promotes participation while maintaining quality. The editorial system described draws on the author's prior experience working in book publishing at Penguin and Random House, showing how a system that ensures quality in print publishing can be adjusted to fit the needs of websites.

Setting

Editing

Most people think of editing in terms of improving the correctness of a document: fixing spelling or punctuation errors, fact-checking, and so forth. These factors are probably the most salient ones in the sense that they are

Emily G. Morton-Owens (emily.morton-owens@med.nyu.edu) is Web Services Librarian, New York University Health Sciences Libraries, New York.

most noticeable when neglected. Editors, however, have several other important roles. For example, they select what will be published. In book publishing, that involves rejecting the vast majority of material that is submitted. In many professional contexts, however, it means soliciting contributions and encouraging authors. Either way, the editor has a role in deciding what topics are relevant and what authors should be involved. Additionally, editors are often involved in presenting their products to audiences. In book publishing, that can mean weighing in on jacket designs or soliciting blurbs from popular authors. On websites, it might mean choosing templates or fonts. Editors want to make materials attractive and accessible to the right audience. Together, *correctness*, *choice*, and *presentation* are the main concerns of an editor and together contribute to *quality*.

Each of these ideas can be considered in light of library websites. *Correctness* means offering information that is current and free of errors, contradictions, and confusing omissions. It also means representing the organization well by having text that is well written and appropriate for the audience. Writing for the web is a special skill; people reading from screens have a tendency to skim, so text should be edited to be concise and preferably organized into short chunks with "visible structure."[7] There is also good guidance available about using meaningful link words, action phrases, and "layering" to limit the amount of information presented at once.[8] Of course, correctness also means avoiding the kind of obvious spelling and grammar mistakes that users find so detrimental.

Choice probably will not involve rejecting submissions to the website. Instead, in a library context it could mean identifying information that should appear on the website and writing or soliciting content to answer that need.

Presentation may or may not have a marketing aspect. A public library's website may advertise events and emphasize community participation. As an academic medical library, NYUHSL has in some sense a captive audience, but it is still important to communicate to users that librarians understand their unique and high-level information needs and are qualified to partner with them.

Workflow

A workflow is a way to assign responsibility for achieving the goals of correctness, choice, and presentation. It breaks the process down into steps that ensure the appropriate people review the material. It also leaves a paper trail that allows participants to see the history and status of material. Workflow can alleviate the coordination problems that prevent a website from exhibiting the quality it should.

A workflow is composed of *states*, *roles*, and *transitions*. Pages have *states* (like "draft" or "published") and users have *roles* (like "contributor" or "editor"). A *transition* happens when a page moves from one state to another. The very simple workflow in figure 1 shows two roles (author and editor) and three states (draft, approval, and published). There are two transitions with permissions attached to them. Only the author can decide when he or she is done working and make the transition from draft to approval. Only the editor can decide when the page is ready and make the transition from approval to published. (In these figures, dotted borders indicate states in which the content is not visible to the public.)

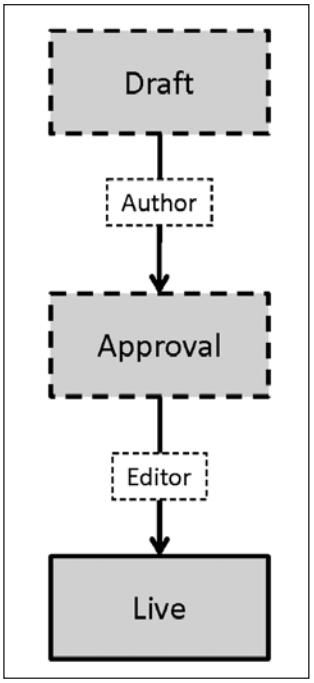

Figure 1. Very Basic Workflow

A book publishing workflow involves perhaps a dozen steps in which the manuscript passes between the author, his or her agent, and various editorial staff. A year can pass between receiving the manuscript and publishing the book. The reason for that careful, conservative process is that it is very difficult to fix a book once thousands of copies have been printed in hardcover. By contrast, consider a newspaper: a new version appears every day and contains corrections from previous editions. A newspaper workflow is hardly going to take a full year. A website is even more flexible than a newspaper because it can be fixed or improved at any time.

The kind of multistep process used for books and newspapers is effective, but not practical for websites. A website should have a workflow for editorial quality control, but it should be proportional to the format in terms of the number of steps, the length of the process, and the number of people involved.

Alternate Workflow Models

This paper focuses on a contributor/editor model in which multiple authors create material that is vetted by a central authority: the editor. Other models could be implemented with much the same tools.

For example, in a peer-review system as is used for academic journals, there is a reviewer role, and an article could have states like "published," "under review," "conditionally accepted," and so forth.

In an upvoting system like Reddit (http://reddit.com), content is published by default, any user has the ability to upvote (i.e., approve) a piece of content, and the criterion for being featured on the front page is the *number* of approvals.

In a moderation system, any user can submit content and the default behavior is for the moderator to approve anything that is not outright offensive. The moderator never edits, just chooses the state "approved" or the state "denied." Moderation is often used to manage comments.

Another model, not considered here, is to create separate "staging" and "production" websites. Content and features are piloted on the staging site before being pushed to the live site. (NYUHSL's workflow occurs all on the live site.) Still, even in a staging/production system the workflow is implicit in choosing someone who has the permission and responsibility to push the staging site to the production site.

Problems at NYUHSL

In 2007, the web services librarian position at NYUHSL had been open for nearly a year. Librarians who needed to post material to the website approached the head of library systems or the "sysadmin." Both of them could post pages, but they did not proofread. Pages that became live on the website stayed: they were never systematically checked. If a librarian or user noticed a problem with a page, it was not clear who had the correct information or was responsible for fixing it. Often, pages that were found to be out-of-date would be delinked from other pages but were left on the server and thus findable via search engines or bookmarks. Because only a few people had FTP access to the server, but authored little content, the usernames shown on the server were useless for determining who was responsible for a page. Similarly, timestamps on the server were misleading; someone might fix one link on a page without reviewing the rest of it, so the page could have a recent timestamp but be full of outdated information.

Even after a new web services librarian started in 2007, problems remained. The new librarian took over sole responsibility for posting content, which made the responsibility clearer but created a bottleneck, for example, if she went on vacation. Furthermore, in a library with five locations and about sixty full-time employees, it was hard for one person to do justice to all the libraries' activities. If a page required editing, there was no way to keep track of whose turn it was to work on the document. There also was no automatic notification when a page was published. This made it possible for content to go astray and be forgotten.

These problems added up to frustration for would-be content authors, a time drain for systems staff, and less time to create new content and sites. The most significant effect was on the quality of the website, which contained mistakes and confusing information.

Methods

NYUHSL Workflow and Solutions

To resolve its web management issues, NYUHSL chose to work with the Drupal content management system (CMS). The ability to set up workflow and inventory content by date, subject, or author was a leading reason for that decision. Other reasons included usability of the backend for librarians, theming options, the scripting language the CMS uses (PHP), and Drupal's popularity with other libraries and other NYU departments.[9]

NYUHSL's Drupal environment has four main user roles:

1. Anonymous: These are visitors to the NYUHSL site who are not logged in (i.e., library users). They have no permissions to edit or manage content. They have no editorial responsibilities.
2. Library staff: This group includes all the staff content authors. Their role is to notice what content library users need and to contribute it. Staff have been encouraged to view website contributions as something casual—more akin to writing an e-mail than writing a journal article.
3. Marketing team: This five-member group checks content that will appear on the homepage. Their mandate is to make sure that the content is accurate about library services and resources and represents the library well. Its members include both librarians and staff with relevant experience.
4. Administrators: There are three site admins; they have the most permissions because they also build the site and make changes to how it works. Two of the three admins have copyediting experience from prior jobs, so they are responsible for content approvals. They copyedit for spelling, grammar, and readability. Admins also check for malformed HTML created by the WYSIWYG (what you see is what you get) interface provided for authors, and they use their knowledge of other material on the site to look out for potential conflicts or add relevant links.

Returning to the themes of correctness, choice, and presentation, it could be said that librarian authors are responsible for choice (deciding what to post), the marketing team is responsible for choice and presentation, and the administrators are responsible for all three.

An important thing to understand is that each person in a role has the same permissions, and any *one* of

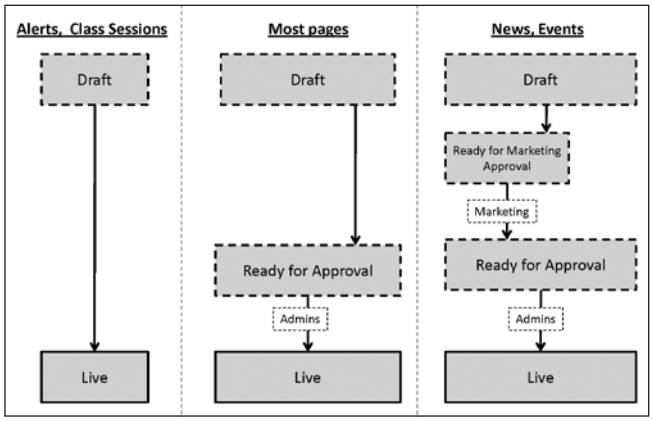

Figure 2. Approval Steps

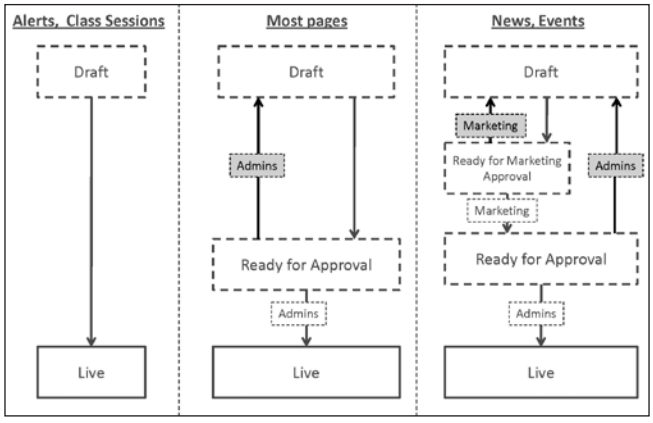

Figure 3. Returning Contents for Edits

them can perform an action. The five marketing team members do not vote on the content, nor do they all have to approve it; instead, any *one* of them, who happens to be at his workstation when they get a notification, is sufficient to perform the marketing team duty. Also, the marketing team members and administrators do not "self-approve"—no matter how good an editor someone may be, he or she is rarely good at editing her own work.

NYUHSL's workflow considers three cases:

1. Most types of content are reviewed by one of the administrators before going live.
2. Content types that appear on the homepage (i.e., at higher visibility) are reviewed by a member of the marketing team before being reviewed by an administrator.
3. Two types of content do not go through any workflow. *Alerts* are urgent messages that appear in red at the top of the homepage. Their appearance should not be delayed, so any staff author can publish one. *Class sessions* are specific dates, times, and locations that a class is being offered. These posts are assembled from prewritten text, so there is no way to introduce errors and no reason to route them through an approval step.

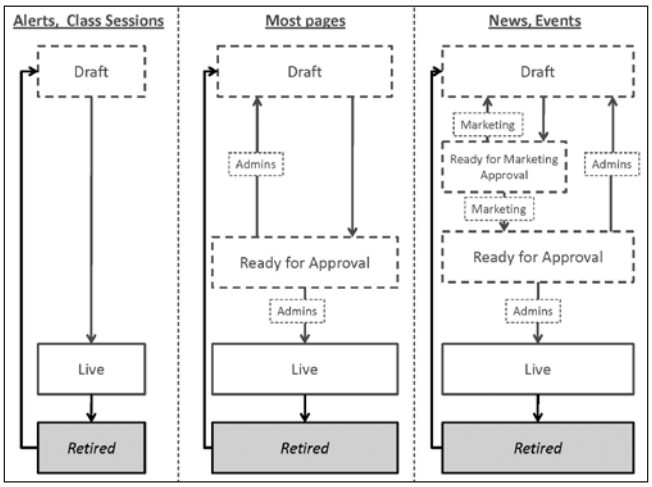

Figure 4. Retirement

Figure 2 illustrates the main steps of the three cases. The names of the states are shown with arrows indicating which role can make each transition. Unlabeled arrows mean that any staff member can perform that step.

Figure 3 shows how, at each approval step, content can be sent back to the author (with comments) for revision. Although this happens rarely, it is important to have a way to communicate with the author in a way that is traceable by the workflow.

Figure 4 illustrates the concept of retirement. NYUHSL needed a way to hide content from library users and search engines, but it is dangerous to allow library staff to delete content. Also, old content is sometimes useful to refer to or can even be republished if the need arises. Any library staff user can retire content if they recognize it as no longer relevant or appropriate. Additionally, library staff can resurrect retired content by resetting it to the draft state. That is, they cannot directly publish retired content (because they do not have permission to publish), but they can put it back on the path to being published by saving it as a draft, editing, and resubmitting for approval.

Figure 5 shows that library staff do not really need to understand the details of workflow. For any new content, they only have two options: keep the content in the draft state or move it on to whatever next step is available. All

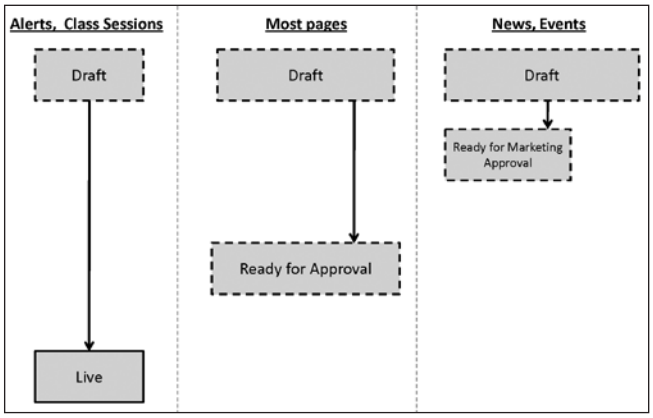

Figure 5. Workflow Choices for Library Staff Users

of the other options are hidden because staff do not have permission to perform them.

The status of content in the workflow can be checked by clicking on the workflow tab of each page, but it also is tracked by notification e-mails. When the content enters a state requiring an approval, each person in that approving role gets an e-mail letting them know something needs their attention. The e-mail includes a link directly to the editing page. For example, if a librarian writes a blog post and changes its state from "draft" to "ready for marketing approval," he or she gets a confirmation e-mail that the post is in the marketing approval queue. The marketing team members each get an e-mail asking them to approve the post; only one needs to do so. Once someone has performed that approval, the marketing team members receive an e-mail letting them know that no further action is required. Now the content is in the "ready for approval" state and the author gets another e-mail notification. The administrators get a notification with a link to edit the post. Once an administrator gives the post final approval, the author gets an e-mail indicating that the post is now live.

The NYUHSL website workflow system also includes reminders. Each piece of content in the system has an author (authorship can be reassigned, so it is not necessarily the person who originally created the page). The author receives an e-mail every four months reminding him or her to check the content, revise it if necessary, and re-save it so that it gets a new timestamp. If the author does not do so, he or she will continue to get reminders until the task is complete. Also, the site administrators can refer to a list of content that is out of date and can follow up in person if needed. Note that reminders only apply to static content types like pages and FAQs, not to blog posts or event announcements, which are not expected to have permanent relevance.

This may sound like a large volume of e-mail, but it does not appear to bother library staff. The subject line of every e-mail generated by the system is prefaced with "[HSL site]" for easy filtering. Also, every e-mail is signed with "Love, The NYUHSL website." This started as a joke during testing but was retained because staff liked it so much. One described it as giving the site a "warm, fuzzy feeling."

Drupal Modules

NYUHSL developers used a number of different Drupal modules to achieve the desired workflow functionality. A simple system could be achieved using fewer modules; the book *Using Drupal* offers a good walkthrough of Workflow, Actions, and Trigger.[10] Of course, it also would be possible to implement these ideas in another CMS or in a home-grown system. This list does not describe how to configure each module because the features are constantly evolving; more information is available on the Drupal website.[11]

The Drupal modules used include:

- Workflow
- Actions
- Trigger
- Token
- Module Grants
- Wysiwyg, IMCE, IMCE Wysiwyg API Bridge
- Node Expire
- Taxonomy Role
- LDAP integration
- Rules

Results

Participation

Figure 6 shows the number of page revisions per person from July 14, 2009, to November 4, 2010. Since many pages are static and were created only once, but need to be updated regularly, a page *creation* and a page *update* count equally in this accounting, which was drawn from the node_revisions table in Drupal. It gives a general sense of content-related activity.

A reasonable number of staff have logged in, including all of the librarians and a number of staff in key positions (such as branch managers). The black bars represent the administrators of the website. It is clear that the workflow system, while broadening participation, has hardly diffused primary responsibility of managing the website. The web services librarian and web manager have by far the most page edits, as they both write new content and edit content written by all other users.

Figure 7 shows the distribution of content updates once the web team members have been removed. It is clear that a small number of heroic contributors are responsible for the bulk of new content and updates, with other users logging on sporadically to address specific needs or problems.

How Editorial Workflow Addresses NYUHSL's Problems

Different aspects of the NYUHSL editorial workflow address different website problems that existed before the move to a CMS. Together, the workflow features create a clearly defined track that marches contributed content along a path to publication while always making the history and status of that content clear.

- Keeping track of who wrote what when: This information is collected by the core Drupal software and visible on administrative pages. (Drupal also can be customized to display or sort this information in more convenient ways.)
- Preventing mistakes and inconsistencies: This requires a human editor, but Drupal can be used to formalize that role, assign it to specific people, and ensure nothing gets published without being reviewed by an editor.
- Bottlenecks: NYUHSL eliminated bottlenecks that stranded content waiting for one person to post it by creating roles with multiple members, any *one* of whom can advance content to the next state. There is no step in the system that can be performed by only one person.
- Knowledge: The issue of having too much going on in the library for one person to report on was addressed by making it easier for more people to contribute. Drupal encourages this through its usability (especially a WYSIWYG editor), and workflow makes it safe by controlling how the contributions are posted.
- "Lost" content: When staff contribute content, they get e-mail notifications about its status and also can check the status by clicking on the workflow tab. This eliminates the discouraging mystery of having content get lost on the way to being published.
- Identifying "problem" content: The Node Expire module has been modified to send e-mail reminders about stale content; as a result, this "problem"

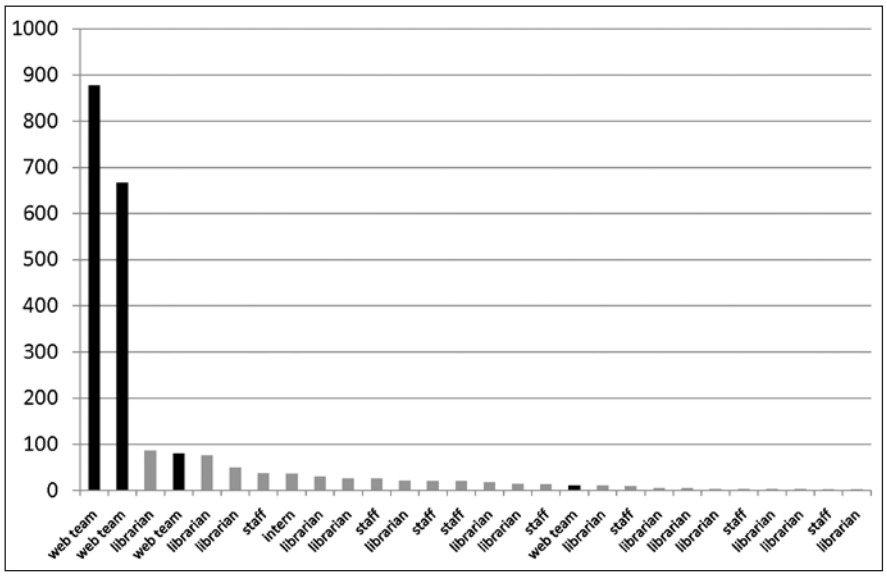

Figure 6. Number of Revisions by User
Each user is indicated by their employee type rather than by name.

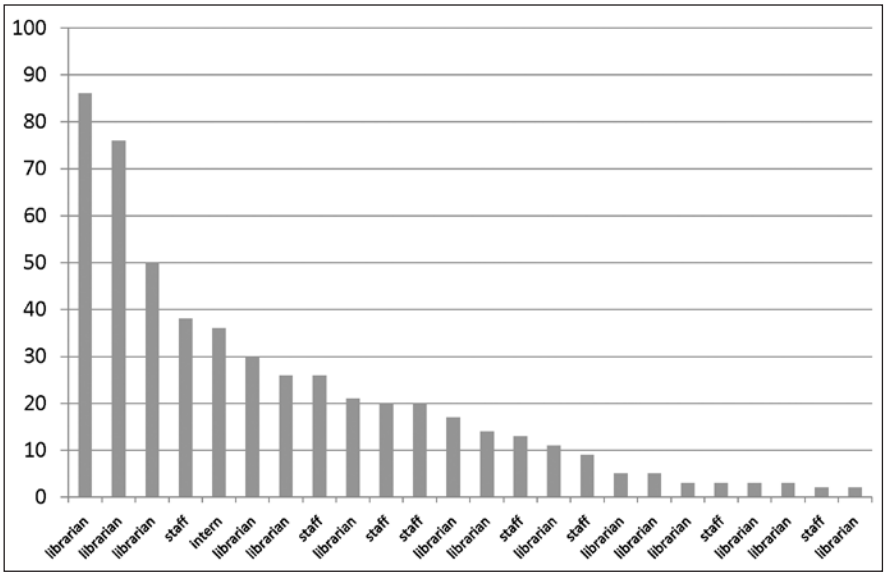

Figure 7. Number of Revisions by User, Minus Web Team
Each user is indicated by their employee type rather than by name

content is usually addressed by library staff without the administrators/editors doing anything at all. The administrators also can access a page that lists all the content that has been marked as "expired" so they know with whom to follow up.

- Outdated content: Some content may be outdated and undesirable to show the public or be indexed by search engines, but be useful to librarians. It also is not safe to allow staff to delete content, as they may do so by accident. These issues are addressed by the notion of "retiring" content, which hides content by unpublishing it but does not delete it from the system.

Future Work

The workflow system sets up an environment that achieves NYUHSL's goals, structurally speaking, but social (nontechnology) considerations prevent it from living up to its full potential. Not all of the librarians contribute regularly. This is partly because they are busy, and writing web content is not one of their job requirements. Another reason is that some staff are more comfortable using the system than others, a phenomenon that reinforces itself as the expert users spend more time creating content and become even more expert. A third cause is that not all librarians may perceive that they have something useful to say. Reluctant contributors have no external motivation to increase their involvement.

It would be helpful to formalize the role of librarians as content contributors. There is presently no librarian at NYUHSL whose job description includes *writing* content for the website; even the web services librarian is charged only with "coordinating, designing, and maintaining" sites. Ideally, every librarian job description would include working with users and would mention writing website content as an important forum for that. That said, it is not clear what metric could be used to judge the contributions fairly.

It also is important to continue to emphasize the value of content contributions so that librarians are motivated and feel recognized. Even librarians whose specialties are not outreach-oriented (e.g., systems librarians) have expert knowledge that could be shared in, say, a short article on how to set up RSS feeds.

Workflow is part of a group of concerns being called "content strategy." This concept, which has grown in popularity since 2008, includes editorial quality alongside issues like branding/messaging, search engine optimization, and information architecture. A content strategist would be concerned with why content is meaningful in addition to how it is managed. In her brief, useful book on the topic, Kristina Halvorson places web content in the context of other communication methods, like e-mail marketing, press releases, and social media.[12] In her view, it is not enough to consider a website on its own; it has to be part of a complete strategy for communicating with an organization's audience. Libraries embarking on a website redesign would benefit from contemplating this larger array of strategic issues in addition to the nitty-gritty of creating a process to ensure quality.

Conclusions

NYUHSL differs from other libraries in its size, status as an academic medical library, level of IT staffing, and other ways. Some aspects of NYUHSL's experience implementing editorial workflow will, however, likely be applicable to other libraries.

It does not necessarily make sense to assign editorial responsibility to IT staff; instead, there may be someone on staff who has editorial or journalistic experience and could serve as the content approver. Many universities offer short copyediting courses, and a prospective website editor could attend such a course.

Implementing a workflow system, especially in Drupal, requires a lot of detailed configuration. Developers should make sure the workflow concept is clearly mapped out in terms of states, roles, and transitions before attempting to build anything. Workflow can seem complicated to users too, so developers should endeavor to hide as much as possible from nonadministrators.

Small mistakes in Drupal settings and permissions can cause confusing failures in the workflow system. For example, a user may find him or herself unable to advance a blog post from "draft" to "ready for approval," or a state change from "ready for approval" to "live," and may not actually cause the content to be published. It would save time in the long run to thoroughly test all the possibilities with volunteers who play each role before the site is in active use.

Finally, when the workflow is in place, the website's managers may find themselves doing less writing and fewer content updates. They have a new role, though: to curate the site and support staff who use the new tools.

The concept of editing is not yet consistently applied to websites unless the site represents an organization that already relies on editors (like a newspaper)—but it is gaining recognition as a best practice. If the website is the most readily available public face of an institution, it should receive editorial attention just as a brochure or fundraising letter would. Workflow is one way that libraries can promote a higher level of quality and perceived competence and reliability through their website presence.

Acknowledgments

Thank you to Jamie Graham, Karen Hanson, Dorothy Moore, and Vikram Yelanadu.

References

1. Charles Martell, "The Absent User: Physical Use of Academic Library Collections and Services Continues to Decline 1995–2006," *Journal of Academic Librarianship* 34 (2008): 400–407.

2. Jeanie M. Welch, "Who Says We're Not Busy? Library Web Page Usage as a Measure of Public Service Activity," *Reference Services Review* 33 (2005): 371–79.

3. B. J. Fogg and Hsiang Tseng, "The Elements of Computer Credibility" (paper presented at CHI '99, Pittsburgh, Pennsylvania, May 15–20, 1999): 82.

4. B. J. Fogg et al., "What Makes Web Sites Credible? A Report on a Large Quantitative Study" (paper presented at SIGCHI '01, Seattle, Washington, Mar. 31–Apr. 4, 2001): 67–68.

5. Andrea Everard and Dennis F. Galletta, "How Presentation Flaws Affect Perceived Site Quality, Trust, and Intention to Purchase from an Online Store," *Journal of Management Information Systems* 22 (2005–6): 79.

6. Pamela Briggs et al., "Trust in Online Advice," *Social Science Computer Review* 20 (2002): 330.

7. Patrick J. Lynch and Sarah Horton "Online Style," *Web Style Guide*, 3rd ed., http://webstyleguide.com/wsg3/9-editorial-style/3-online-style.html (accessed Dec. 1, 2010).

8. Janice (Ginny) Redish, *Letting Go of the Words: Writing Web Content that Works* (San Francisco: Morgan Kaufman, 2007).

9. Emily G. Morton-Owens, Karen L. Hanson, and Ian Walls, "Implementing Open-Source Software for Three Core Library Functions: A Stage-by-Stage Comparison," *Journal of Electronic Resources in Medical Libraries* 8 (2011): 1–14.

10. Angela Byron et al., *Using Drupal* (Sebastopol, Calif.: O'Reilly, 2008).

11. All Drupal modules can be found via http://drupal.org/project/modules.

12. Kristina Halvorson, *Content Strategy for the Web* (Berkeley, Calif.: New Riders, 2010).