The need for innovation in assessing web content performance: A case study

Robert B. Watson
University of Washington
rbwatson@uw.edu

Jan H. Spyridakis
University of Washington
jansp@uw.edu

Abstract** – **Best practices provide convenient, but general, guidelines for authoring web content; however, as general guidelines, they cannot possibly provide the best advice for every context. Testing is required to know if their application works in a specific context. However, testing web content is challenging, especially content that does not support the funnel-shaped interactions typically found in commerce-oriented web sites. This paper presents the case of an experiment that tested how varying two common best practices for web content—content minimalism and visual navigational affordances—affected software developers' perceptions and task performance. Applying these best practices improved some metrics as expected, but also reduced or had no measurable effect on others. The experience from this case study revealed opportunities for innovation in web-content authoring: (1) describing best practices to include the contexts in which they are most effective and the factors they might improve, (2) developing methods for modeling audience requirements and goals as they relate to the effective use of information concepts and visual design, and (3) improving the tools and methods used to measure web content and integrate such measurements into the authoring process.

Index Terms** – **Authoring for the web, Best practices, Case study, Web analytics.

INTRODUCTION

Best practices and rules-of-thumb provide efficient shortcuts and guides; however, and almost by definition, they are approximations and generalizations. As such, knowing the context in which a practice is actually a *best* practice and knowing the effects that a given practice has in a specific context are critical to achieving the desired results in real-world applications, or, more to the point, in *your* applications. Knowing how a practice will affect your content is essential to selecting the practices that work for your content, your audience, and your organization. The diverse and often conflicting goals that much web content must meet make selecting and validating content practices difficult, while the value to be gained by validating and improving content practices makes this an area ripe for innovation.

Best practices for web content generally start out with the best of intentions—to help ease the pressure that authors face when they create content. However, the contexts in which the practices have been demonstrated to improve some aspect of the web-content experience tend, over time, to lose their finer points, eventually becoming only a catch phrase or slogan [1]. Many authors of web content are familiar with some of the more popular best-practices-turned-catch phrases in web-content authoring, such as, "Less is more," "cut-cut-cut," and "people don't read, they scan [web content]." These so-called best practices have become so prevalent that an entire web site (uxmyths.com) was created to catalog them.

While their abbreviated forms make catchy seminar titles, tweets, and bullet points, their condensed forms lack the nuance and context that gave them their original value. Further, when separated from the context for which they were developed, they become too easy to generalize and apply in situations for which they were not intended and for which they may not be effective. "Cut! Cut! Cut!" for example, is the catch phrase (or perhaps, the battle cry) for simply "write what readers want to read (no more and no less)." The latter may be more pedestrian, but "Cut! Cut! Cut!" omits the difficult, yet critically important, prerequisite of knowing what the reader wants to read.

The devolution of best practices into catch phrases would quickly correct itself if suitable feedback systems existed to test the effect of a practice, best or otherwise. For web sites with a funnel-shaped interaction, many tools are available to instrument and assess tasks that are accomplished completely in a web context. (Funnel-shaped interactions are found in web sites designed to attract a large population and direct them to a specific task—such as to make a purchase—a task flow that is frequently depicted as a funnel.) However, for web content produced to satisfy tasks and reader goals that cannot be completed in a web context, comparatively few tools are available for assessing its success in meeting reader needs. Web sites with educational content, self-

help content, and reference content support readers as they perform tasks outside of the web and do not have a funnel-shaped interaction. As such, tools and metrics oriented around funnel-shaped interactions do not model or represent reader interactions with these sites very well. Applying metrics designed for funnel-shaped interactions in such cases can produce inappropriate and perhaps misleading data that could potentially encourage continuing practices that do not accomplish the content's goals for readers or organizations.

The experiment in this case study [2] provides an example in which the best practices it tested (content minimalism and use of visual design to provide navigational affordances) delivered unexpected results in a common, non-funnel-shaped interaction. The experiment measured reader performance as task-completion time and task accuracy and also measured readers' perceptions of the content's professionalism and credibility. The experiment tested the application of two best practices to API reference topics and found that some metrics improved, as expected, some unexpectedly degraded, and, in some cases, made no difference. This paper reviews the context in which these results were obtained and how the best practices applied could be described better to predict their effects more accurately.

The experiment's findings identified opportunities for innovation in the methods and instruments used to describe best practices, content goals, and content measurement. Making it easier to collect meaningful data about web content that does not have a funnel-shaped interaction can help improve how authoring practices are described and selected. Ultimately, improving authoring practices and evaluation methods will improve content experiences for readers.

BACKGROUND AND LITERATURE REVIEW

The best practices applied and tested in the experiment described later in the case study are found in practitioner and academic references and do not seem limited to a particular genre. These best practices are reviewed here.

Reducing the length of web content is a recurring best practice. In the 1980s, Carroll presented work on the advantages of minimal content in task instructions [3], proposed Minimal Instruction in the 1990s [4,5], and critically reviewed it in 1998 [6]. Jakob Nielsen has published articles since the 1990s that describe how little people read on the web—encouraging shorter, more succinct content [7-9]. Redish [10] encourages shorter content over longer content; however, how short "short" is depends on the context—that is the readers and their goals [1,10,11]. The experiment described in this case study based its definition of "short" on prior research into the content of API documentation [12-15] and operationalized the variable as information-concept elements [2].

Redish [10], Nielsen [8], and Krug [16], to name a few, also recommend that web pages provide navigational affordances such as visually distinct headings that support skimming. Stating that, "online, readers skim and scan" and that they do not read, they suggest using visually distinctive headings to help facilitate such interaction.

An attractive visual appearance is also encouraged as a best practice for web content [10,17]. Best practices abound for effective use of typography, color, layout, and other visual style elements to improve readability [10,18] and to instill confidence in the content [19]. The variations of visual styles and affordances used in the experiment described here were derived from variations observed in API reference documentation [14] and visual design guidelines presented in popular best practices as seen applied in open-source API reference documentation [2].

The context of web content in the eyes of readers can be studied and described as personas [10,20,21], which can provide insight into use cases and reader goals that inform content and document design.

While the preceding best practices mention audience context and readers' goals, their context and goals do not always map audience characteristics directly to content elements (either in terms of information or context). Many encourage empirical testing of web-content performance, ideally in a closed feedback loop that applies "best practices," observes the result, makes adjustments, and then takes new measurements [10,21]. Unfortunately, observing and measuring the result can be problematic. If readers' goals for reading the web content lie outside of the web experience, measuring readers' success in achieving their goals with such content can be difficult [22].

The funnel-shaped interactions found in e-commerce and similar, action-oriented (and often, revenue generating) web content are compelling for several reasons. The funnel accurately describes common, and ideally profitable, use cases. Their convergence towards a single goal, such as a purchase, makes the use case easy to understand and measure. Their popularity, focused goals, and profitability have motivated the development of many measurement and analysis tools.

There exist, however, many sites with interactions that are not funnel-shaped. Brandt et al. [23] described a software development scenario that is similar to those of the experiment. The scenario and tasks converge towards a goal—completing a programming task—but that goal exists outside of the web experience. While the web-interactions they observed helped their participants advance towards completing software development tasks, the web-interactions and related software development actions converged to a point outside of the web. The web interaction in the scenario that Brandt et al. observed is an intermittent one in which readers dip in and out of the web content. The web content plays an important, yet

only supporting, role in accomplishing readers' goals. In contrast, with funnel-shaped interactions, the goals and the supporting steps to achieve those goals generally exist entirely within the context of the web site. Funnel-interaction metrics, such as *bounce rate* [24], that indicate a content failure in a funnel-shaped interaction (bounce rate describes how often readers visited a page and left after only a short time) indicate the expected, if not desired, interaction in the case of a reference topic, such as Brandt et al. observed.

For authors of content designed for a funnel-shaped use case, best practices are plentiful as are the tools to measure and optimize their effectiveness in a specific context. Authors and stakeholders of content that does not support a funnel-shaped use case, however, are at a comparative disadvantage. Assumptions and scenarios on which best practices for funnel-shaped interactions are based, might not apply to scenarios that do not have a funnel-shaped interaction. The experiment described in this case study is such an example.

CASE STUDY REVIEW

This case study reviews an experiment that tested the effects of two common best practices and produced unexpected results [2]. The best practices tested concerned the topic (document) length, as measured by the number of unique information concepts, and the number of visual-design elements used. The experiment tested software developers' task-completion performance and their perceptions of the topics' professionalism and credibility as they performed information-seeking tasks with application programming interface (API) reference topics.

The experiment used sample topics taken from open-source API reference documentation and varied their content and appearance based on the best practices of content minimalism and visual navigational affordances. Figure 1 is an example of a topic variation that applies neither best practice and Figure 2 is an example of a topic variation that applies both best practices by adding visual design elements and navigation affordances and reducing the content to a minimum. Each topic in the experiment had four variations to test each combination of the two best practices studied. The experiment measured how the application of these best practices influenced software developers' task performance and perceptions.

The experiment was conducted as an online, remote, unmoderated, user-experience survey. Participants were invited through a snowball sampling method and completed a demographic questionnaire. Participants were guided through a practice task before performing up to four measured tasks, each of which consisted of: (1) a scenario describing a situation and the information to seek and (2) a topic in which to find the information requested of the scenario. Participants indicated whether they felt that the topic contained the information and indicated the location in the topic that most influenced their decision. After each measured task, participants responded to a perception questionnaire and after participants completed all tasks, they had the option to register for a chance to win an Amazon gift certificate.

FIGURE 1. EXAMPLE OF TOPIC VARIATION WITH FEW VISUAL-DESIGN ELEMENTS AND MANY INFORMATION CONCEPTS

FIGURE 2. EXAMPLE OF TOPIC VARIATION WITH MANY VISUAL-DESIGN ELEMENTS AND FEW INFORMATION CONCEPTS

We expected the experimental conditions that applied the best practices to show significantly faster task-completion time and higher ratings of readers' perception of the topics' professionalism and credibility. Rather than striking and significant differences, the results we observed were mixed. For example, while the participants who used versions of the topics with fewer information concepts (the best practice) completed the information-seeking tasks more quickly than those who used topics with more information concepts, the shorter topics were considered to be less credible and less professional than the longer topics. Participants who used topics with more visual design elements and navigation affordances (the best practice) completed the tasks as quickly as participants given topics with fewer visual design elements did. At the same time, participants who used topics with more visual design elements and navigation affordances reported them to be more credible and professional than those who used topics without fewer design elements.

On average, participants in the experiment spent an average of 44 seconds reading the document presented for each information-seeking task—about the time required to read each document completely two times.

<center>DISCUSSION</center>

Observations from the experiment in this case study reveal how the effects of different best practices can interact and how, in the experiment's context at least, can provide unexpected results. This section reviews the experiment's findings in the context of the best practices that were applied and some of the questions the findings raised.

Do Readers Skim, Read, or do Both?

The assertion that readers do not read web content, rather they simply skim it for snippets of information, does not completely describe the results observed in this experiment. The average time that participants spent evaluating the topics in this experiment was long enough to read the topics completely almost two times. Unfortunately, no eye-tracking data were collected during the experiment to know whether they were skimming, reading, or, most likely, some combination of both as McGovern describes [25]. Reading, skimming, or both, the time spent on each topic seems quite long to consist of only skimming.

The experiment timed the duration required for participants to evaluate the topic and complete their information-seeking task and after completing the timed, information-seeking task, participants indicated the spot in the document that most influenced their decision. That the information-seeking tasks were completed more quickly in the topics with fewer information unique information concepts suggests that the best practice to

reduce the amount of content in a topic encourages faster access; however, it does not provide any information about how they read the topics.

In the absence of eye-tracking data, which is difficult to collect in an unmoderated, remote survey, some information about the reading patterns could be inferred by comparing the time required to find the information in the topic and the location of the information as indicated by the participant. A high positive correlation between distance from the top of the topic and the time spent in the topic might suggest that participants were following a linear path to the information while a low or negative correlation could suggest that the participant was following a more complex pattern to locate the information.

How Minimal is too Minimal?

The notion of what constitutes minimal content is very task, audience, and context specific. Since Carroll's initial publications on the topic, it has been the subject of much interpretation and misapplication [6]. A recurring theme in the discussion is the importance of matching the content to the task, audience, and context, which seems difficult to do in practice without effective methods to model and evaluate these parameters. While participants using the shorter topic variations in the experiment (fewer information concepts) completed information-seeking tasks more quickly than those using the longer topic variations (more information concepts), they perceived the topics with fewer concepts to be less credible and professional. In this case, depending on which property was most valuable to the author or the author's organization, less information might or might not be better than more.

These observations illustrate several problems with applying best practices out of their intended context. The first problem concerns how to know when your context and key performance metric match those that a particular best practice will improve—in this case, what criteria were used to identify the task, audience, and context of a document and how those properties translate to the application of a specific best practice. The next challenge that web-content developers face is to identify and operationalize the measures that will tell them whether their content is having the desired effect. Many of the web tools available for commercial content can be applied to measure reader tasks that can be completed in a web context; however, measuring tasks for which the web content plays only a supporting role is more difficult.

Is Reading Behavior Motivated by Task or Genre?

One possible explanation for the unexpected results seen in the experiment is that other factors had a greater or overwhelming influence on the experiment's measured values than the best practices did. One aspect of the experiment reviewed in this study that might not be

apparent to those who are unfamiliar with API reference documentation is that such content tends to follow a familiar format—a format that has changed very little in over 30 years [2,14,15]. The document variations used in the experiment also followed this pattern. It is possible, although not tested in the experiment, that the participants' information-seeking and reading patterns were influenced by their expectations of the document genre and that these expectations had a greater influence, or initial influence, on their reading pattern than the experiment's variations of navigation element design and the number of information concepts. If this were the case, then the most influential best practice for design of documents in this genre could be simply "follow the existing or established pattern." If the pattern or readers' expectation of a pattern is strong, other document variations might have little effect on readers' perception or performance.

While the genre of API reference documentation and API documentation, in general, has been fairly consistent over the past several decades, it is also a frequent object of dissatisfaction among its target audience of software developers [12,26]. In recent years, however, API reference topic design has started to show signs of evolving. API documentation generators for Representational State Transfer (REST) APIs that have been increasing in popularity and population are experimenting with new layouts to address how those APIs are developed and used (examples include swagger.io, box.com, readme.io). While these new formats push the frontiers of the API documentation genre and they have been generally well received, it is still too early to tell whether the genre will evolve to accommodate new technologies or return to the format that has endured the past several decades.

What is the Goal of the Content?

The experiment illustrated the importance of knowing the goal or goals of the content—specifically, measuring the values that reflect how well the content is achieving those goals. Applying best practices in the experiment had varying effects on the measured variables. The topic variations in the experiment with fewer information concepts showed faster reading-task performance, while, at the same time, participants' rated the topics' professionalism and credibility lower. While there might be a shortage of meaningful metrics for measuring a given content goal, web content provides no shortage of aspects that can be measured. Knowing which measurement reflects how well the content is achieving its goal is essential not only to applying the practices that will improve the content, but also to not being distracted by measurements that do not measure what is important.

CONCLUSION

The experiment described in this case study illustrates an example of how applying so-called best practices to web content produced results that were not entirely well informed by the best practices. Part of the experiment's surprise resulted from the best practices not providing sufficient context—or having lost their context over time—so as to predict their effect in a specific application in advance. If a best practice is not always a best practice, the context in which it is should be clear. If the context of a best practice cannot be clearly specified, the testing of its application should be a priority. If a practice's context cannot be specified nor its effects tested, then can it really be a *best* practice? In the experiment, the nature of the documents tested made it difficult to state reader goals in a way that could be measured solely by web-based analytics because they supported a task in which readers' goals existed mostly outside of the web experience.

The difficulty in characterizing the contexts and effects of a given best practice is further complicated by difficulty in measuring the performance of web content that supports an interaction that is not funnel-shaped. In content that is not designed for a funnel-shaped interaction, it is difficult to know which practices are really the best for the effects their application will have. A shortage of accurate models and measurement methods in this scenario provides many opportunities for innovation.

The easier it becomes to collect and analyze information about how content is performing, and then to apply that knowledge towards improving the content, the more likely such an iterative process can be adopted to improve the content. As the modeling factors and content goals become more known and accepted, they can be incorporated into the authoring tools used to create and update content. The real opportunity is to make the process of collecting, analyzing, and incorporating data about readers' goals a natural part of the content authoring process. Ideally, collecting and applying this information would become an authoring practice that is as natural as formatting a paragraph.

OPPORTUNITIES FOR FUTURE WORK

This case study presented an experiment that tested two common best practices for designing web content and found that the best practices did not influence performance as expected and, in some cases, negatively affected perception. While this study presented only one case, it highlights some of the aspects of best practices that warrant scrutiny before a practice is applied to a specific case. Understanding the goals of the content, the goals of the audience when they use the content, and goals for which a best practice actually improves the content are important considerations. Unfortunately, as vital as that information is, it is often difficult to obtain and apply. Finally, the experiment demonstrated the value

of testing web content. In this case, the experiment measured the effects of variations in the information concept count and visual design elements in a specific context that provided feedback that could be applied to the content authoring practices. The experiment had detailed instrumentation to measure the effects that the best practices had on the readers' performance and perceptions. Innovations in the aspects that the experiment studied could make it easier to perform such testing in the future and help identify and refine the practices applied to web content.

Content Goals versus Practice Goals

One could argue that the best practices applied in the experiment are not intended to reduce task-performance time or improve readers' perceptions of API reference topics and, as such, the experiment's results should come as no surprise. Such an argument would be valid, however, if the practices were described such that an informed decision could be made in advance. Ideally, the list of applicable or exceptional use cases and the improvements one could expect would have been stated with the best practices. For example, a best practice could provide that information if it were stated as, "in online informational content, reducing the word count 50% reduces the reader's task performance time by 20%." Such a claim, of course, would be risky to make in the absence of a more comprehensive context. Yet, the much less specific "best practice" of "Cut! Cut! Cut!" is repeated frequently.

Following some best practices to their logical extremes is clearly absurd. Taking "Cut! Cut! Cut!" to one extreme suggests that the best results would be found with no content whatsoever. Taken to the other extreme, it would suggest that continuing to add content would ultimately make the result valueless. In reality, and as many best practices state in their original form, most best practices describe a *sweet spot* or an optimal application. For web content, knowing where the optimal application of a practice is requires both accurate audience modeling (to get as close as possible to the optimum on the first try) and testing in context (to learn how to adjust your model and the resulting practice). Unfortunately, the methods for determining the optimum application of a practice are difficult and have become lost in the sloganization of the practices.

The experiment described in this study identified an example of where applying best practices had a mixed influence on the aspects measured compared to what was anticipated or expected. The shortcoming is not so much with the best practices, but with the ability to map best practices to specific contexts and then to test the applications in context. At the same time, the experiment provides a ripe context for identifying opportunities for innovation in audience modeling, content modeling to meet business needs, instrumentation and authoring processes, and sharing information across genres

Opportunities in Audience Modeling

There are many opportunities to improve how audiences are modeled, how those models relate to the content that helps audiences achieve their goals, and the measurements taken to evaluate the success of the model. Task analysis of readers is not novel; however, it is not particularly easy to perform or translate to specific practices. The time pressures that practitioners face require that audience modeling and analysis methods be as quick and easy to apply as possible. The process of analyzing an audience in a way that more directly informs the content requirements needs to be easier to accomplish, especially for content that does not support a funnel-shaped interaction.

Audience modeling must clearly describe how the content supports readers and organizational goals. Without a clear relationship between the audience model and the content goals, it will be impossible to identify the content that the audience requires, how well the content meets those requirements, and what should be improved, if it does not.

Existing audience modeling methods, such as segmentation and personas, can be time-consuming and expensive, which makes them difficult to apply when time and budget are in short supply. Innovations are needed in ways to make collecting, applying, refining, and sharing this information less expensive and less time-consuming.

Opportunities in Content Modeling to Meet Business Needs

As with audience modeling, opportunities exist to improve how to model content goals that suit the needs of other stakeholders such that the organization's goals are met. Modeling business goals and reader goals in terms of content requirements can be challenging, even when readers' goals and organizational goals are confined to the web experience. When either or both of those goals are supported by, but exist outside of, the web experience, modeling and measuring them become much more challenging. The opportunity in this area is to develop goal models that support reader goals that are outside of, the web experience.

Opportunities exist in professional and academic fields to explore this modeling from a broader perspective. Supporting a broader view of readers' and organizational goals will enable those goals to be expressed and measured in ways that are not artificially constrained to a web-only experience. After the goals of the audience, content, and organization can be modeled in a consistent and coherent form, instrumenting and measuring how the content performs with regard to those goals will become more straightforward.

Opportunities in Instrumentation and Authoring Processes

Innovations in goal modeling, instrumentation and content authoring processes can occur in parallel—iterating off each other. Effective test and measurement methods that are easy to apply and not limited to the web-experience are necessary to support more general content models. Improved instrumentation is also necessary to enable the use of informal and expedient modeling methods that can then be refined in iterations over time.

Innovations in instrumentation and modeling, however, must be accompanied by innovations in the authoring process. An iterative authoring process that is supported by planning, authoring, publishing, and measurement tools is necessary to apply the information gleaned from the improved instrumentation.

Opportunities in Sharing Information

Creating innovations to improve sharing information across web-content genres and applications will help accelerate the iterative cycle of innovation in all genres. A good place to start is identifying what we already know about modeling and measurement from other genres. For example, take our understanding of the methods and measurements used by commercial web content designed for funnel-shaped interactions and identify how they can (and cannot) be applied to other genres. The innovations will likely come in the form of finding ways to share that information in a way that enables the information to be applied to other types of content.

Another opportunity exists in sharing ways to develop authoring practices that support an iterative feedback cycle. Many genres of web content exist that are written to support goals and actions outside of a web context, so there are many opportunities to share and learn. While it might seem most practical to develop these models separately and specifically for each genre, with innovations in these areas, there is an opportunity to develop them in way that facilitates learning and sharing from other genres such that they all improve more quickly.

The Best Best Practice

If this case study has identified a universal best practice, it is to test, measure, and iterate the content. While this is not a new discovery, this study identifies some of the challenges writers face trying to apply it. For this practice to be applied more frequently requires innovation in the areas described here—and, most likely, in many more areas that are not described here. Developing methods to share and learn from the different genres will take advantage of what we already know and accelerate our learning of what we do not know.

REFERENCES

[1] J. M. Carroll and H. van der Meij, "Ten misconceptions about minimalism," in *The Nurnberg Funnel: Designing Minimalist Instruction for Practical Computer Skill*, MIT Press Cambridge, MA, 1990, pp. 55–90.

[2] R. B. Watson, *The Effect of Visual Design and Information Content on Readers' Assessments of API Reference Topics*, Doctoral dissertation, University of Washington, Seattle, WA, 2015.

[3] J. M. Carroll, P. L. Smith-Kerker, J. R. Ford, and S. A. Mazur-Rimetz, "The minimal manual," *Human–Computer Interaction*, vol. 3, no. 2, pp. 123–153, Jun. 1987.

[4] J. M. Carroll, *The Nurnberg Funnel: Designing Minimalist Instruction for Practical Computer Skill*. MIT press Cambridge, MA, 1990.

[5] J. M. Carroll, "Techniques for minimalist documentation and user interface design," in *Quality of Technical Documentation*, Rodopi, 1994, pp. 67–75.

[6] J. M. Carroll and S. for T. Communication, *Minimalism Beyond the Nurnberg Funnel*. MIT Press, 1998.

[7] J. Nielsen, "Be succinct! (Writing for the web)," *Nielsen Norman Group*, 15-Mar-1997. [Online]. Available: https://www.nngroup.com/articles/be-succinct-writing-for-the-web/. [Accessed: 17-Apr-2016].

[8] J. Nielsen, "How users read on the web," *Nielsen Norman Group*, 01-Oct-1997. [Online]. Available: https://www.nngroup.com/articles/how-users-read-on-the-web/. [Accessed: 17-Apr-2016].

[9] J. Nielsen, "How little do users read?" *Nielsen Norman Group*, 06-May-2008. [Online]. Available: https://www.nngroup.com/articles/how-little-do-users-read/. [Accessed: 17-Apr-2016].

[10] J. Redish, *Letting Go of the Words, Second Edition: Writing Web Content that Works*, 2nd edition. Amsterdam; Boston: Morgan Kaufmann, 2012.

[11] S. Rosenbaum, "Follow-up on training in minimalism: How are technical communicators using minimalism?" in *Minimalism Beyond the Nurnberg Funnel*, MIT Press, pp. 119–147, 1998.

[12] M. P. Robillard, "What makes apis hard to learn? Answers from developers," *Software, IEEE*, vol. 26, no. 6, pp. 27–34, 2009.

[13] M. P. Robillard and R. DeLine, "A field study of API learning obstacles," *Empirical Software Engineering*, vol. 16, no. 6, pp. 703–732, 2011.

[14] R. B. Watson, "Development and application of a heuristic to assess trends in API documentation," in *Proceedings of the 30th ACM international conference on Design of communication*, Seattle, WA, pp. 295–302, 2012.

[15] R. Watson, M. Stamnes, J. Jeannot-Schroeder, and J. H. Spyridakis, "API documentation and software community values: A survey of open-source API documentation," in *Proceedings of the 31st ACM International Conference on Design of Communication*, pp. 165–174. 2013.

[16] S. Krug, *Don't Make Me Think, Revisited: A Common Sense Approach to Web Usability*, 3rd edition. Berkeley, Calif.: New Riders, 2014.

[17] J. Tidwell, *Designing Interfaces*. Sebastopol, CA: O'Reilly, 2011.

[18] W. Lidwell, K. Holden, and J. Butler, *Universal Principles of Design: 125 Ways to Enhance Usability, Influence Perception, Increase Appeal, Make Better Design Decisions, and Teach through Design*. Beverly, Mass.: Rockport, 2010.

[19] D. Robins and J. Holmes, "Aesthetics and credibility in web site design," *Information Processing & Management*, vol. 44, no. 1, pp. 386–399, 2008.

[20] A. Cooper, *The Inmates Are Running the Asylum: Why High Tech Products Drive Us Crazy and How to Restore the Sanity*, 1st ed. Sams - Pearson Education, 2004.

[21] G. McGovern, *Killer Web Content: Make the Sale, Deliver the Service, Build the Brand*, 1st edition. London: A&C Black Trade, 2006.

[22] R. B. Watson and J. Spyridakis, "Using readers' and organizations' goals to guide assessment of success in information websites," in *Human-Computer Interaction: Design and Evaluation*, Springer, pp. 283–294, 2015.

[23] J. Brandt, P. J. Guo, J. Lewenstein, M. Dontcheva, and S. R. Klemmer, "Two studies of opportunistic programming: interleaving web foraging, learning, and writing code," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1589–1598, 2009.

[24] "Bounce rate - Analytics help," *Google Analytics Help*. [Online]. Available: https://support.google.com/analytics/answer/1009409?hl=en. [Accessed: 23-Apr-2016].

[25] G. McGovern, "Block reading: how we read on the web | Gerry McGovern," *New Thinking*, 05-Oct-2008. [Online]. Available: http://www.gerrymcgovern.com/new-thinking/block-reading-how-we-read-web. [Accessed: 17-Apr-2016].

[26] C. Parnin, "API documentation: Why it sucks," *ninlabs research*, 04-Mar-2013. [Accessed: 04-May-2013].

ABOUT THE AUTHORS

Robert B. Watson is an Assistant Professor in Technical Communication at Mercer University. He graduated from the University of Washington with a Ph.D. in Human Centered Design & Engineering and his research focuses on how writers create, evaluate, measure, and manage technical documentation. Before earning his Ph.D., his professional career spanned over 30 years during which he worked as a senior programmer/writer and as a senior software developer.

Jan H. Spyridakis is a Professor in the Department of Human Centered Design & Engineering at the University of Washington. Her research focuses on designing and investigating tools for remotely testing users, understanding best practices for the design of online information, and designing online tools and information to support leadership and advancement of women in science and engineering. She is a Fellow of the Society for Technical Communication