Optimizing crowdsourcing websites for volunteer participation / A case study: What's on the Menu? New York Public Library

Donelle McKinley

School of Information Management

Victoria University of Wellington

donelle.mckinley@vuw.ac.nz

@donellemckinley

www.digitalglam.org

Table of Contents

1. Slide: Title

2. Slide: Project list

As many of you are aware, a growing number of academic and collecting institutions are outsourcing tasks to the general public through an open call, otherwise known as crowdsourcing. This approach is being used to create and enhance online resources more cost-effectively, engage the wider community, and enable research.

Online volunteers are assisting with a wide range of tasks such as tagging, identification, proofreading, transcription, text encoding, translation, and contextualisation.

Today we'll be looking at What's on the Menu? (WOTM), a project focused on collection preservation, access and research. New York Public Library's restaurant menu collection holds about 45,000 menus dating from the 1840s to the present ("What's on the Menu?," n.d.).

With insufficient staff and financial resources to transcribe the digitized menus NYPL looked to crowdsourcing (Taranto, 2011). The aim of the project is to create a searchable database that can be browsed by dish, beverage, price, restaurant name, and date.

A custom-designed website was launched in April 2011, and a redesign in June 2012. To date, over one million dishes have been transcribed from more than 15,000 menus. This progress, and the project's commitment to enabling the use, reuse and 'remix' of

collection data makes the project an interesting example for other collecting institutions and the Digital Humanities.

So if that's crowdsourcing in a nutshell, at least on our terms, what's website optimization?

Website optimization involves clearly defining the objectives of the website and aligning them with the objectives of the online visitor. It aims to optimize the page elements and processes that impact on the perceptions, actions and experience of online visitors.

3. Slide: Working bee

Optimizing a website for volunteer participation is perhaps best understood by envisaging the offline equivalent. The project website is where volunteers congregate; ideally, it's welcoming, with clear signage and appealing workspaces. It's obvious who's in charge, why the volunteers are here, what they need to do and how it's done.

Volunteers know where to find help at any stage, and their valuable free time is not wasted being confused, frustrated, or bored. There are opportunities to mix with and learn from other volunteers. They sense the buzz of activity and see progress being made. Perhaps they've only dropped by for a short time or they stay several hours; either way, their contributions are acknowledged, and they leave in a positive state, looking forward to their next visit.

Usability author Steve Krug (2006, p. 162) talks about the "reservoir of goodwill" that visitors bring to a website, and explains that each problem they encounter lowers its level. To maximize the potential of that reservoir, we need to provide a rewarding user experience by actively optimizing our site for participation.

4. Slide: Trove

Rose Holley (2009, 2010) has shared tips for crowdsourcing based on her experience as manager of the Trove historical newspaper project at the National Library of Australia. She emphasizes the importance of clearly stating the goal of the project, identifying 'the crowd', understanding volunteer motivations, and providing relevant incentives. The online environment must be intuitive, reliable, quick and easy to use, and include a chart of progress. Holley suggests volunteers should be given the choice to identify themselves on the site and receive acknowledgement, and be provided with tools to contribute to a dynamic and supportive team environment. Inviting feedback from volunteers from the early stages of project development helps to ensure the crowdsourcing platform meets users needs.

Holley's work has been hugely influential for a number of crowdsourcing projects, including WOTM, as NDF keynote Michael Lascarides (2011) mentioned this time last year.

5. Slide: Heuristics

Marketing Experiments Labs (or MECLABS) is an independent research lab for optimization experimentation. Their methodology includes a small set of heuristics, or core principles, for optimizing websites and online communication that reflects Holley's

observations. These are the five main elements that impact on conversion, which is the process of achieving the primary objective of a specific page or website.

Clarity of the value proposition (why the visitor should take action), user motivation, incentives, friction and anxiety. In the context of website optimization, friction refers to the aggravation, fatigue or confusion of the online visitor, and anxiety to concern about aspects of the website or online process (McGlaughlin, 2010).

Based on real-world experimentation with partners that include the New York Times, Microsoft and others, MECLABS has observed that the probability of conversion is subject to the relationship between these five elements; clearly defined reasons for the visitor to take action need to be aligned with visitor motivations, incentives need to reflect these motivations, and sources of friction and anxiety should be minimized (McGlaughlin, 2010).

To meet the needs of the online visitor undergoing these often simultaneously occurring thought processes, and fulfil the objectives of the project website, every web page element needs to serve a clearly defined purpose.

6. **Slide: WOTM homepage (full)**

So how can we apply this set of heuristics to websites designed for volunteer participation?

Let's start with the value proposition, which represents the mutually beneficial exchange between the institution and the volunteer. This is our invitation to participate. It should concisely and clearly state who is driving the project, the nature and purpose of the project, and the action visitors are being asked to take (Sherman, 2011, p. 52). Krug (2006, p. 31) explains that an effective invitation relies on a "clear visual hierarchy" presenting page elements in order of priority, and a clear call to action.

In this example, the value proposition is "What's on the Menu? Help the New York Public Library improve a unique collection". It's supported by a brief explanation, and a clear call to start transcribing. These page elements have been prioritized at the top of the homepage, so within seconds of arrival visitors know who's driving the project, what the project is, why they should get involved and how they can get started.

The second element impacting on conversion is motivation. Based on volunteer surveys conducted by Holley and other project teams, common motivations for participation include the size of the challenge, the necessity for volunteer contribution, collaboration with prestigious institutions, contribution to research, education, mental stimulation, being part of a community, personal research interests, and enhancing a resource from which they will benefit (Holley, 2010; Simon, 2010, p. 195; Smith, 2011).

WOTM reflects these likely volunteer motivations in various ways. Homepage visitors are told, "it's a big job, and we need your help". Reading on, they'll discover that they will be helping to digitize "one of the largest menu collections in the world". By joining "tens of thousands of patrons" transcribing menus, they will be part of a community, which they can interact with via Twitter, Facebook, and the project blog. Visitors with a

particular interest in food and culinary history can aid the research of fellow "historians, chefs, and food enthusiasts", some of whom have published testimonials at the bottom of the homepage.

The third element impacting on conversion is incentive, which should reflect an understanding of volunteer motivations. Crowdsourcing websites commonly include regular updates indicating project progress, evidence of how new data is being used, and leader boards that acknowledge top contributors. Project teams may acknowledge volunteer contributions in published research, provide open access to the online resources being created or enhanced, and promote a sense of community by enabling volunteers to interact with each other and the project team (*RunCoCo*, n.d.).

7. Slide: WOTM home (above fold)

What's on the Menu updates transcriptions on the home page in real time. It also strives to maintain volunteer momentum with the banner across the top of the homepage, which indicates that batches of freshly digitized menus are ready for transcription.

8. Slide: WOTM home (below fold)

Appealing images are used throughout the site to illustrate the unique collection, and encourage visitors to explore. The homepage also demonstrates how the digitized collection is being used, by showcasing themed sets and exhibitions.

9. Slide: Task interface (Ticks)

The available time volunteers will have to contribute on any given day will differ considerably; some may have several hours to spare, others only minutes. The WOTM

dish-by-dish transcription approach means even volunteers with only a few minutes to spare can contribute. Updating the number of dishes transcribed for each menu in real time demonstrates to volunteers that every line of transcription makes a difference, and brings the institution closer to achieving its goal.

10. Slide: Menus (timeline)

Volunteers can choose from a range of menus based on the historical period that interests them, restaurant name, or menus at varying stages of completion.

11. Slide: Dishes (Data visualization)

Interactive visualizations, faceted browsing, and an open API[i], show volunteers some of the ways that the data they are contributing can be accessed and reused.

Friction, the fourth element impacting on conversion, refers to the aggravation, fatigue or confusion of the online visitor. Sources of friction can include confusing navigation, readability issues, uncertainty over what something means or does, difficulty locating information on a 'busy' page, and the number and complexity of processes required to complete the desired action. Sources of friction can sometimes be difficult for project teams to identify when they're familiar with the project and the website. In *The Art of Community* Jono Bacon (2009, p. 124) encourages us to ask "Is this step really required? How easy is this step to understand? Could it be simplified?".

12. Slide: Transcribe drop-down instructions

From the WOTM home page, it's possible for visitors to start transcribing a menu in just three clicks. They click the help transcribe button, select a menu, and then click a

dish to transcribe. Concise instructions are delivered in a step-by-step manner on a 'need to know' basis, and link to more detailed guidelines on the Help page.

Requiring online visitors to register before they participate enables us to identify and acknowledge top contributors, but visitors may be hesitant to provide personal information, and some may want to simply try it out without going through the process of sign up. WOTM volunteers aren't currently required to register. Minimizing the steps between arrival and actual transcription allows new visitors to quickly determine whether or not the task is going to be manageable, enjoyable and rewarding in a way that reflects their personal motivations.

On the other hand, there is evidence to suggest that overall contribution may suffer without a system in place to track and acknowledge participation. Reporting on the Trove project Holley (2009) observed, that "Initially 50% of the corrections were done by anonymous users and 50% by registered users but by the end of 6 months 80% of corrections were being done by registered users". A survey of volunteers contributing to the Transcribe Bentham project indicated that while competition and recognition were not significant motivational factors for most, they were for the top contributors. During the pilot period of Transcribe Bentham, the seven most active volunteers represented 70% of actual contributions (Causer, Wallace, & Tonra, 2012, p. 126). For these volunteers registration is in their interests, and the project leaderboard, which acknowledges top contributors with a points system, is an incentive.

WOTM addresses this issue in the frequently asked questions (FAQ) section of the About page. In answer to 'Can I create an account?' visitors are informed that 'During

this first experimental phase, we're trying to keep things as open as possible, but we intend before long to add a user account system to start more visibly tracking contributions from the community. We're grateful for the time/effort you devote to this endeavor, and hope to be able to recognize some of our top contributors down the road.' In this way the site minimizes friction for new visitors to the site, aligns visitor motivation with future incentive, and maximizes the potential for contribution during the first phase of the project.

13. Slide: Transcribe detail (arrow)

Let's move on to the fifth element impacting on conversion: anxiety. MECLABS Director Flint McGlaughlin (2010) explains that anxiety is closely related to perception. He urges not to undermine this aspect of the user experience, and to over-correct potential sources of anxiety where possible.

In the context of non-profit crowdsourcing, visitor concerns may include the credibility of the project, and whether it's being actively managed and supported. Is it worth investing their time? Who are the people behind the project? Is there someone to answer questions? New volunteers might also be concerned about their ability to use the task interface, and enter data correctly.

WOTM uses real-time progress updates, instantly available data and social media to demonstrate the project is being actively managed and supported. The 'About' page includes profiles of the project team, a contact email is included on several web pages, and a personal, informal tone is used throughout the site.

Transcribing the menus one dish at a time minimizes the potential for error and confusion. For better accuracy an arrow is placed alongside the dish to be transcribed, and users have the ability to magnify the highlighted text. Concise instructions are presented in simple language, and the large fields provided for transcription are limited to dish and price. No superfluous page elements distract from the task at hand, and volunteers can indicate if they've had trouble reading the text after typing their 'best guess'.

14. Slide: Google Analytics

While it may seem like a lot of work to design and evaluate each page element with these heuristics in mind, Jesse Garrett (2003, pp. 161–167) argues that "the right approach is one in which no aspect of the user's experience is left to chance". He explains that every decision should be made consciously and deliberately, and be grounded in our understanding of the underlying issues at play. Because these issues aren't always immediately apparent to us, we need to test and track user behaviour.

One approach is to have a 'soft' launch, whereby select groups of volunteers are invited to test the live website before an open call goes out to the public. During the six-month pilot period, the WOTM team sought to build up public participation, and improve the user interface (National Endowment for the Humanities, n.d.). The project blog continues to be used to field volunteer questions, and inform users of changes to the website.

Website analytics, also plays an important role in the post-launch process. This can provide information such as the number of unique visitors, average time spent on the

site, and average visitors per month. Considered alongside data like the online visitors who register as volunteers, actual contributors, abandoned and completed tasks, and the rate of volunteer participation, we can begin to evaluate the effectiveness of the website and the success of the project.

After 4.5 months, the WOTM website had received 58,000 unique visitors, and over 10,000 menus had been fully transcribed (Lascarides, 2011). One month later, the team reported that on average, visitors to the website spent seven minutes per visit (Vershbow, Owens, & Brumfield, 2012). After 10 months, the website had received more than 3 million page views. While Project Director Ben Vershbow (2012) considers this is a rough proxy for a "highly engaging" website, his colleague Barbara Taranto emphasises that this level of public engagement is also heavily reliant on marketing (2011). But we need to leave *that* story for another time…

15. Slide: Shirky

"What matters most now is our imaginations. The opportunity before us, individually and collectively, is enormous; what we do with it will be determined largely by how well we are able to imagine and reward public creativity, participation, and sharing."
(Shirky, 2010, p. 212)

---

[i] Application programming interface

References

Bacon, J. (2009). *The art of community : building the new age of participation*. Sebastopol, CA: O'Reilly.

Causer, T., Wallace, V., & Tonra, J. (2012). Transcription maximized; expense minimized? Crowdsourcing and editing The Collected Works of Jeremy Bentham. *Literary and Linguistic Computing*, *27*(2), 119–137. doi:10.1093/llc/fqs004

Garrett, J. (2003). *The Elements of User Experience: User-centered design for the web*. USA: Aiga/New Riders.

Holley, R. (2009). *Many Hands Make Light Work: Public Collaborative OCR Text Correction in Australian Historic Newspapers*. Australia: National Library of Australia. Retrieved from http://www.nla.gov.au/openpublish/index.php/nlasp/article/view/1406/1688

Holley, R. (2010). Crowdsourcing: How and Why Should Libraries Do It? *D-Lib Magazine*, *16*(3/4). doi:10.1045/march2010-holley

Krug, S. (2006). *Don't Make Me Think: A common sense approach to web usability* (2nd ed.). USA: New Riders.

Lascarides, M. (2011). Aim at the Chopping Block. Presented at the National Digital Forum, Wellington, New Zealand. Retrieved from http://www.r2.co.nz/20111129/michael-l.htm

McGlaughlin, F. (2010). *MECLABS Landing Page Optimization Online Course*.

National Endowment for the Humanities. (n.d.). *"What's on the Menu? Crowdsourcing Culinary History at The New York Public Library - Narrative Section of a Successful Application."*

RunCoCo - Beyond Collections: Crowdsourcing for public engagement. (n.d.).University of Oxford. (2011, May 26). Retrieved from http://podcasts.ox.ac.uk/series/runcoco-beyond-collections-crowdsourcing-public-engagement

Sherman, A. (2011). *The Complete Idiot's Guide to Crowdsourcing*. New York: Penguin.

Shirky, C. (2010). *Cognitive Surplus: Creativity and Generosity in a Connected Age*. New York: Penguin Press.

Simon, N. (2010). *The Participatory Museum*. California: Museum 2.0.

Smith, A. (2011, May 26). *A Cautionary Tale*. Presented at the RunCoCo - Beyond Collections: Crowdsourcing for public engagement, University of Oxford. Retrieved from http://podcasts.ox.ac.uk/series/runcoco-beyond-collections-crowdsourcing-public-engagement

Taranto, B. (2011, December 12). *Crowd Sourcing Metadata*. Presented at the Coalition for Networked Information (CNI) Fall 2011 Membership Meeting, Arlington, Virginia. Retrieved from http://vimeo.com/38196574

Vershbow, B., Owens, T., & Brumfield, B. (2012). IMLS Crowdsourcing panel. Presented at the 2012 WebWise Conference on Libraries and Museums in the Digital Age. Retrieved from http://www.tvworldwide.com/events/webwise/120229/default.cfm?tab=2

What's on the Menu? (n.d.). Retrieved May 25, 2012, from http://menus.nypl.org/