GetHealthyHarlem.org: Developing a Web Platform for Health Promotion and Wellness Driven by and for the Harlem Community

Sharib A. Khan MBBS, MA, MPH[1], Jessica S. Ancker MPH, MPhil[1], Jianhua Li MS[1], David Kaufman PhD[1], Carly Hutchinson[2], Alwyn Cohall, MD[2], Rita Kukafka DrPH, MA[1,2]

[1]Department of Biomedical Informatics, Columbia University, New York, NY
[2]Department of Sociomedical Sciences, Mailman School of Public Health, Columbia University, New York, NY

Abstract

GetHealthyHarlem.org is a community website developed on an open-source platform to facilitate collaborative development of health content through participatory action research (PAR) principles. The website was developed to enable the Harlem community to create a shared health and wellness knowledgebase, to enable discourse about local and culturally relevant health information, and to foster social connections between community members and health promotion organizations. The site is gaining active use with more than 9,500 unique site visits in the six months since going live in November, 2008. In ongoing research studies, we are using the website to explore how the PAR model can be applied to the development of a community health website.

Keywords: Health promotion, collaborative technologies, participatory action research

INTRODUCTION

Disadvantaged communities such as Harlem in New York City bear a heavy burden of disease and health disparities.[i] Web 2.0 technologies afford potential new tools to engage communities in health promotion and chronic disease prevention[ii]. The new web empowers consumers to become creators, organizers, evaluators and promoters of the content. In essence, Web 2.0 is reversing the top-down, one-way direction of information flow, through the use of blogs, wikis, social networking, and other collaborative content creation methods.

Our research suggests that Internet technologies can be useful even in disadvantaged communities: a random-digit telephone survey of 646 Harlem residents found high rates of computer use (77%) and Internet use (65%)[iii]. The survey also established the community's strong desire for accurate information relevant to their health conditions, contextualized for their community, and based on the experience of local residents. Furthermore, this survey and subsequent qualitative research confirmed that community members themselves are repositories of this information.

We reported earlier on our formative efforts, beginning in 2007, to use an open-source social software (Drupal) to enable Harlem residents and community-based organizations to form on-line social networks necessary to share knowledge and experiences, and provide the motivation and support required for health behavior change[ii,iii]. According to social constructivist theory, participation and dialogue enabled by social software offers participants the opportunity to construct and organize knowledge, building social capital[iv]. Similar uses of Web 2.0 technologies have begun to diffuse in the healthcare domain over the last few years[ii,v].

In this paper we describe the construction of a Drupal-based website, www.GetHealthyHarlem.org (GHH), including requirements gathering, participatory and iterative development, and the architecture of the Drupal system. We provide preliminary statistics on site usage. We conclude by discussing challenges and on-going efforts to extend the site into a platform for other health promotion interventions.

METHODS

The research framework upon which this project operates is community-based participatory action research (PAR). The philosophy of PAR is to engage and empower the participants in the process of creating a resource[vi]. Instead of a single-source, top-down approach to design or development of an intervention (such as a website), PAR principles advocate bottom-up involvement of the people for whom the intervention is planned. All aspects of this project (including deciding to build the website at all) involve community members. The involvement of the community as partners and collaborators is critical, as it ensures that any planned intervention is grounded in the needs of the community and is created for and by the community. PAR is particularly important when working with disadvantaged populations and communities such as Harlem where there is distrust toward the research establishment.

Community involvement was ensured in several ways. First, the project is a collaboration with the Harlem Health Promotion Center (HHPC), a prevention research center funded by the Centers for Disease Control (CDC). HHPC, as described briefly above and in detail elsewhere[iii], conducted a random-

digit dial telephone survey. Second, the survey was supplemented with 6 focus groups to further understand the nuances of attitudes and perceptions of technology use[iii,vii]. Five additional focus groups were conducted to examine risk communication and to also view and discuss early mockups of the GHH site[viii]. Focus group transcripts were analyzed using grounded theory, which provides a means of discovering categories and relationships in the data[ix]. Survey data and emergent themes were translated into requirements for the portal (such as supporting the need for discussion of local resources and health traditions). Third, a Community Advisory Board (CAB) comprised of volunteers and community stakeholders meets regularly to brainstorm about the community's needs and priorities with regard to health, to provide feedback on HHPC activities, and to view and contribute to (first) mock-ups and (later) interactive versions of the GHH site. Fourth, a series of usability tests with community members provided useful design feedback.

Policy: Open participation versus expert control

The recurrent debate throughout this process was whether to give users full ability to create and post health content, or whether to control content by recruiting experts to create it or at least screen it for quality and accuracy. Members of our internal Web Portal Development Group (WPDG) and of the CAB expressed concern that the website might become a forum for inaccurate or harmful information. On the other hand, every focus group and CAB meeting included instances in which participants told the development team about new local events, resources, and other information, demonstrating that the collective knowledge of the community was greater than the knowledge of the expert team.

After several rounds of discussion, we chose to be more inclusive and enable greater participation, given the precedent set by user-generated sites such as Wikipedia (www.wikipedia.com). Others have found that wiki communities tend to be self-correcting and are not substantially more error-prone than expert-written encyclopedia[ii,x]. However, to balance the risks, members of HHPC and the WPDG also decided to monitor for inappropriate content. We decided that inaccurate postings would provide experts the opportunity to dispel misconceptions by posting rebuttals or explanations.

Iterative development

Throughout development of the GHH website we have incorporated the feedback and advice from our internal research team, the CAB, and our internal Web WPDG. We employed a participatory and iterative process of translating the high-level requirements into actual implementation details such as, but not limited to determining the underlying data models, the wire-frames, information flow, system features, graphic design, navigational elements and access permissions. Our development followed iterative cycles of making incremental additions to the prototype of the website and obtaining feedback from the research team and the WPDG. For instance, the advisory board and the WPDG contributed feedback on the website name, navigation options, content requirements and even the color choices for the site. For usability, we used expert-driven heuristic evaluation and usability testing with representative users[xi].

Technical implementation: Drupal architecture

Drupal (www.drupal.org) is just one example of social software that includes many of the characteristic Web 2.0 technologies and provides a sophisticated, scalable and high performance platform for building a community-driven information and networking website. Several innovative architectural features of Drupal, which align with the requirements of the site as determined by our PAR approach, informed the selection of this platform. We chose this content management system (CMS) because it allows easy creation of different types of content, has fine-grained user permissions, allows for "plug and play" added functionality, provides mechanisms for easy customization (not just through configuration parameters but programmatically), and has active development support from the technical community.

The structural framework of the Drupal system is centered on the "nodes" (pieces of content) and "users". Each node has properties (such as an identifier, title, description, etc) and Drupal provides the ability to define different node (content) types. Users can perform functions (e.g., creating content) based on their permission levels. Drupal supports creation of user groups and allows administrators to assign granular (down to a specific node level) read/edit/write permission for each group. Nodes can be given associated functions such as the ability for people to rate them, comment on them or forward the URL of the post via email by installing and activating modules. Another innovative feature about Drupal is its support for associating "tags" or keywords with content posts using user-driven taxonomies (also referred to as folksonomies). These taxonomies can contain terms and parent-child relationships between terms, offering the ability to categorize the content into as many different slices as possible and to restrict access based on the content's taxonomy classification.

Drupal separates the page layout and presentation (look and feel) into a separate theme system. This system renders each page as composed of different regions (such as a header, footer and a sidebar); any number of blocks can be placed within each region, at

positions controlled through an administration panel. This architectural separation provides the flexibility to customize a Drupal installation to suit specific needs by activating/de-activating modules through an administration panel or to change the default behavior of the system by overriding or extending the core functions and system variables.

The core Drupal system contains functions needed to generate a page, including node creation, access control, and security. Add-on modules, created by the community of developers extend the functionality of the system. Several modules exist that provide collaborative features such as the ability to create different content types, perform social networking through user profiles, and collaboratively filter content through tags, ratings, and reviews. Other modules supports dissemination and aggregation of content through RSS feeds, or integration with other sites. Drupal is written in the popular programming language PHP and works with several relational databases such as MySQL.

Organization of www.GetHealthyHarlem.org

The core setup of any Drupal website involves building content types, defining user roles and permissions, adding functionality through modules, determining navigation structures and page layout elements. Each of these tasks was informed by our requirements as determined in the preliminary research studies described above. We created content types for users to post a) articles (factual descriptions of health topics), b) blogs (personal experiences or journals), c) web links,   d) event listings (Figure 1) and e) forum discussions. The focus group data highlighted strong desires in the community to know more about local resources. Thus, we also created several content types to enable users to submit information about resources such as places to eat or exercise, environmental organizations, and healthcare facilities. The resulting directories of health-related services and resources are self-updating.

We created a taxonomy for classifying content under six topic categories: Fitness, Food, Environment, Spirit and Mind, Health Information, and Around Harlem. Menu tabs (Figure 1) reflect these categories: clicking one displays a page with all the content in that category, sorted in reverse chronology and by content type. The user can also sort by number of views, comments and rating of each content type to exploit collaborative filtering. The focus group data suggested that Harlem residents might suspect the motives of a website that required registration for all users because of the collection of personal information. Thus, the website is open for viewing by all site visitors. However, previous

experience with social networking websites demonstrates the counterbalancing need for accountability among users. For example, web communities need to block people who post offensive content. We balanced openness and accountability by requiring registration before posting content. To reduce collection of personal information, registration entails only specifying a username and a working email address. Verified users can post, comment on other posts, or rate posts.
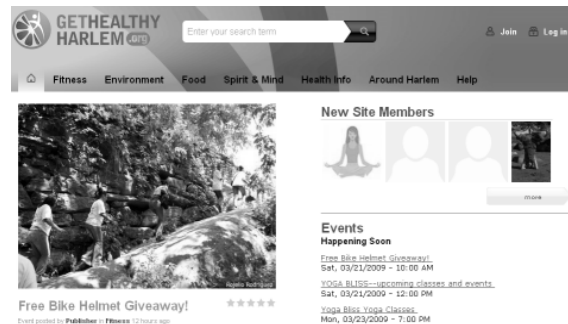

Figure 1. Screenshot of the GHH website.

In addition, focus group data and the CAB opinions suggested a need to distinguish between expert content (provided, for example, by doctors or health promotion specialists at HHPC) and user-generated content (UGC). We created the role of the 'publisher' for our internal research team, which has greater control than other authenticated users. Any content submitted by the Publisher has a different visual theme to show that it is expert content. We also created an administrator role for overall monitoring of site activity, tracking inappropriate posts, spam and other housekeeping functions.

We installed several contributed Drupal modules to extend the functionality of the base system. These included modules to enables users to comment on the posts, rate them (using a 5-star rating scale) or flag them as inappropriate. A rich-text editor module supports uploading images or embedding Flash videos from third-party sites. Modules that enable users to build a basic profile, contact other members though an internal messaging system, and subscribe to receive notifications for updates or comments posted on different content types were also added. To foster greater engagement and knowledge discovery, the site also has a "related posts" block in which posts that are relevant to a given post are displayed to the user. An RSS feed aggregator that provides relevant health news stories, has also been added. We performed programmatic customizations besides the basic configuration, writing our own module to process some of the page layouts, to enable secondary navigation links and other such tasks. The WPDG

preferences and usability test findings showed that the site should reflect the cultural nuances of the Harlem community; hence we developed our own theme to change the base colors of the site, to add a block to display images of Harlem and to visually align the site to the sensibilities of the community.

Usability evaluation

After the initial site construction, we conducted expert usability inspection and 3 rounds of usability evaluation. The participants were a convenience sample recruited through word of mouth. The first round involved 2 proficient computer users. Each performed 8 tasks of varying levels of complexity. A sample basic task was to find information about yoga, whereas a complex task was to post a blog. Although both participants could complete the tasks, they had considerable difficulty especially in navigation. In particular, they could not use the primary categories (e.g., Fitness) to locate information. The navigation problem was compounded by inadequate labeling and feedback to the user. The subjects thus relied on the search function. We revised the layout of the site and made the information architecture more transparent. We have also enhanced our site search by using Google's custom site search on GHH, thus leveraging Google's search technology for our site search.

In the second round, 3 users experienced fewer difficulties and completed the tasks in less time. They also expressed greater satisfaction with the system. Most problems were comparatively minor (e.g., the article titles were not sufficiently visible), but some navigation problems persisted. The results prompted another round of design changes as well as a new visual theme resulting in a significant aesthetic improvement and greater transparency. We are currently engaged in cognitive research focused on understanding the processes of how users create their own content (e.g., articles and blogs). We are also beginning to investigate how communities form and communicate about health.

Site usage

We provide some preliminary statistics on the usage of the site in six months since launching in November 2008. Table 1 shows the total number of posts submitted on the site for each content type. Particularly popular topics include complementary and alternate medicine, diabetes, yoga, and information about local events. An article series named 'Harlem Word', created by our content team to profile local doctors, professionals, or community residents who promote health and wellness, has received good attention. There has also been encouraging discussion through comments on several posts. For instance, a post asking users to suggest way

to shop for healthy foods in grocery stores elicited a rich discussion through comments from other users, demonstrating the embedded knowledge within the community.

Content Type	Number of Posts
Articles	119
Blogs	51
Links	143
Events	247
Health Facilities	27
Places to Exercise	18
Comments	300

Table 1. Number of posts by content type

The site traffic, as measured by unique visitors per month, is also increasing (Figure 2), with search engines being the mainstay of incoming traffic. Although the site is fairly new, it shows up prominently in Google searches for several keywords such as 'Harlem fitness', names of local doctors, or events. The site has also received inbound links from other Harlem related sites, increasing the Google page rank (a marker of the site's prominence) to 5.
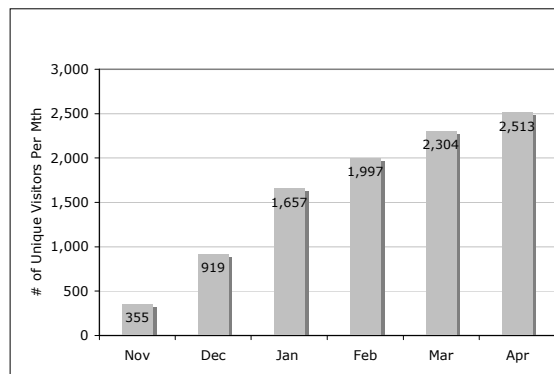

Figure 2. Number of Unique Visitors Per Month

DISCUSSION

We have described the motivation and the process of eliciting need, requirements and the development of a community-driven resource for health promotion and wellness. The site is starting to gain traction and adoption by local fitness and wellness professionals. In the process of creating this site we came across several challenges and learned some key lessons.

We realized that building an open, participatory, community-driven health resource inverts control from the experts to the users. This is a challenging proposition for the research community, and on several occasions there were considerable disagreements with our own internal team and the CAB about the tradeoff between openness and participation, and the validity of the UGC. Thus far, no obviously inaccurate or harmful health content has

been posted (other than spam links, which we have blocked using automated filters), so the long-term implications of this decision remain open for continued observation and research.

Our choice of Drupal has been fairly productive, efficient and cost-effective. However, open-source solutions pose problems as well. Developers need the time to learn Drupal, and obtaining support to deal with bugs can be time-consuming and at times, difficult. Drupal has an active user community, knowledgebase, and well-documented references, which have helped tremendously. Finally, it has been a challenge to ensure usability. For example, we have been unable to modify Drupal's text editor to improve its usability, and are instead developing tutorials.

Growing the user and content base is a challenge. The most effective marketing is through increasing search engine traffic, by following search engine optimization practices such as descriptive URLs, submitting a sitemap, and monitoring search referral traffic for keywords. Some of our highest traffic generators are keyword searches for local organization names and doctors in Harlem and we are building up content to reflect this search trend. It will also be a challenge to ensure that users return to the site after the first few visits.

We envision the GHH website becoming a platform for other health promotion interventions. One such extension will involve a health behavior assessment tool that provides tailored information based on the use of tags in the content's description[xii]. Improvements such as these would help transform the site from a repository of passive information to a site enabling action. We aim to make GHH a way to link primary care with personalized self-management resources in the community. We also plan effectiveness trials to quantify the impact on health promotion and self-management.

Conclusion

We have used PAR and Web 2.0 technologies to build a community-driven website for health promotion and wellness. Our early usage statistics demonstrate a positive trend. We are extending the base system to function as a platform for more targeted and personalized health interventions. We are seeking to ground health promotion and self-management into actionable information and resource linkages in the community. The goal of the system is to enable patients who are motivated to manage their health, to connect with others and access the resources available in the community, thus extending the continuum of health information systems from electronic health records, to personal health records, to a collaborative community health knowledgebase.

References

[i] McCord C. Freeman HP. Excess mortality in harlem. New Eng J Med.1990;322(3):173-7

[ii] Ancker JS, Carpenter K, Greene P, Hoffman R, Kukafka R, Marlow L, et al. Peer-to-peer communication, cancer prevention, and the internet. Journal of Health Communication in press 2009.

[iii] Kukafka R, Khan SA, Hutchinson C, McFarlane DJ, Li J, Ancker JS, et al. Digital Partnerships for Health: Steps to develop a community-specific health portal aimed at promoting health and well-being. AMIA Ann Symposium 2007:428-432.

[iv] Vygotsky LK. Mind Society: the Development of Higher Mental Processes. Cambridge: MA. Harvard University Press, 1978

[v] Weiss JB, Lorenzi N. Synthesizing Community Wisdom: A Model for Sharing Cancer-Related Resources through Social Networking and Collaborative Partnerships. AMIA Annu Symp Proc. 2008:793-7.

[vi] The Role of Community-Based Participatory Research. AHRQ Publication No. 03–0037, June 2003. Agency for Healthcare Research and Quality; Rockville, MD.

[vii] Khan SA, McFarlane DJ, Li J, Ancker JS, Hutchinson C, Cohall A, et al. Healthy Harlem: Empowering health consumers through social networking, tailoring and Web 2.0 technologies [poster]. AMIA Ann Symp Proc 2007:1007.

[viii] Ancker JS, Kukafka R. A combined qualitative method for testing an interactive risk communication tool. AMIA Annu Symp Proc. 2007:16-20.

[ix] Strauss A, Corbin J. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. 2nd ed. Thousand Oaks, CA: Sage; 1998.

[x] Esquivel A, Meric-Bernstam F, Bernstam EV. Accuracy and self correction of information received from an internet breast cancer list: content analysis. BMJ 2006;332:939-942.

[xi] Nielsen J. Usability Engineering. Cambridge (MA): Academic Press; 1993.

[xii] Khan SA, Kukafka R, Cohall A. A Tag based Recommendation Engine to Suggest Information Resources in an Online Community for Health Promotion. AMIA Annu Symp Proc. 2008:1002