

Community in Mashups: The Case of Personal Geodata

Elizabeth Goodman

UC Berkeley School of Information
102 South Hall Berkeley, CA 94720-4600
egoodman@ischool.berkeley.edu

Andrea Moed

UC Berkeley School of Information
102 South Hall Berkeley, CA 94720-4600
amoeda@ischool.berkeley.edu

ABSTRACT

A mashup is an information system that combines two or more separate data streams to create original content. An online community is a system where two or more people create original content in the process of interacting with one another. In many cases, developers seek to build mashups that are *also* communities. We argue that best practices for mashup development and the imperatives of community building online can conflict. Remixing several streams of content into an evolving interface can interfere with community-members' needs for stability, persistence, and control of their online self-presentation. Based on observations of several map mashup sites, we consider the implications of this conflict for collaboration around personal geodata. We argue that there is no experiential precedent for online place-sharing, and the new mashups of mapping and social software require more explicit agreements among system developers and individual contributors.

Author Keywords

Mashups, online community, geodata

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

Though musical remixing and sampling has a long history, mashups – the musical form – emerged as a broadly popular form of musical expression only recently. In 2001, a DJ called Freelance Hellraiser had an international bootleg hit with “A Stroke of Genie-us,” which combined vocals and music from two artists into one catchy track. A musical mashup, according to Wikipedia is: “At their best,

[mashups] strive for musical epiphanies that add up to considerably more than the sum of their parts” [10].

Online, a mashup is something a little different: “A mashup is a website or Web 2.0 application that uses content from more than one source to create a completely new service” [11]. Like their musical namesake, web-based mashups are hybrids. But unlike “A Stroke of Genie-us,” many mashups have not in the past competed with their “parents” on richness of features or depth of content.

In this paper, we will focus on “geo-mashups” – hybrid sites that draw on freely available online map functionality. Geo-mashups that expose personally relevant data on publicly-viewable maps present new questions for designers and developers concerned with sustainable, enjoyable online social interaction.

Some caveats

Our observations in this paper are preliminary. We base them on Moed’s own engagement with the process of building geo-mashups and on a series of five semi-structured 1.5-hour usability activities conducted with first-time users of OurPlaces.com¹, an online map-based community site.

The activity consisted of a thirty-minute interview focused on participant’s existing mapping practices, then a task-completion exercise using OurPlaces.com that combined a pre-set list of tasks with goals and scenarios derived from the participants’ responses to the interview questions. On the assumption that they would be the likely target audience for a geo-mashup, participants in the usability activity were chosen based on a reported interest in maps and mapping and for a range of expertise in Internet use. They included a transportation planner, a fitness coach, an avid amateur cyclist, and a community technology activist. All were daily Internet users, but their time online each day ranged from one to 10 hours. All were familiar with online mapping software (such as Google Maps), and had used online maps within the past month for purposes such as locating a

¹ Goodman conducted user research activities on behalf of OurPlaces.com during July – August 2006. OurPlaces.com is a pseudonym that is similar in tone and structure to the name of the original sponsoring company.

nearby restaurant, getting directions to a known destination, and planning a bicycle ride.

As well, we have generated a provisional list of map-based mashups and have done some comparative analyses of their interface design, market communications strategies, and functionality. None of these methods, however, have been exhaustive. For us, this workshop constitutes an opportunity to refine them with a knowledgeable and critical audience.

WHAT IS AN ONLINE COMMUNITY?

The study of online communities is a flourishing field with a rich history of literature. For the purposes of this paper, however, we will use Wellman's insight that online communities are essentially *social networks* mediated by communication technologies that unbind the traditional notions of community from spatial constraints [19].

Building and sustaining communities

Using Kollock's Design Principles for Online Communities [8] as a basis, we have identified three characteristics of successful online communities that are relevant in understanding the implications of geo-mashups for building and sustaining community:

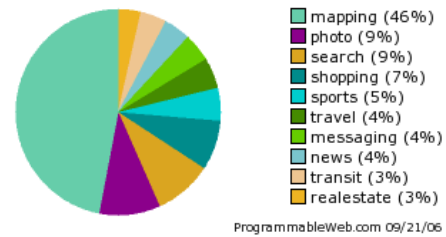
- *Control of self-representation*: User must be able to control their own social space, which includes the way they present themselves to others.
- *Persistence*: Persistent user identities allow users to establish reputations and evaluate each other's trustworthiness on the basis of past actions.
- *Stability*: Successful communities maintain well-defined boundaries. They provide continuity over time, so that user expectations of the space in which social interactions occur remain consistent.

GEO-MASHUPS

The appeal of geo-mashups

Though mashups need not have a geographic component, the practice of making mashups is strongly tied to use of Application Programming Interfaces (APIs) from companies such as Google or Yahoo. These geo-mashups typically mix one or more feeds of data from a large, pre-existing online collection with a map built an API. Unlike earlier online map-based interfaces, these APIs require little expert knowledge of geographic information systems (GIS). They are also (at least for the present) free of charge. As well, many popular online communities (such as Craigslist.org) offer feeds of their contents already labeled with street addresses or other geographic locators.

Given the distributed nature of web development, tracking the number of geo-mashups on the Internet as a whole is unrealistically difficult. However, developer forums are one indicator. As of September 21, 2006, mashups tagged "mapping" formed 46% of the 1001 mashups listed on one mashup community site [14].



ProgrammableWeb.com 09/21/06

Figure 1: Mashups labeled as mapping-related dominate one mashup developer community's listings. Source: programmableweb.com [14]

Mashups as browsers

Some examples of popular geo-mashups from the past three years include the Chicago Crime Map [2], Housing Maps [5], and Mappr [9]. These popular early sites act primarily as spatial visualizations of database content. These "listing" mashups can extend the efficiencies of the original datasource by broadening their audience. For example, real estate postings from a community site like Craigslist.org may be more likely answered if shared with more potential seekers on Housing Maps.

However, browser-based geo-mashups are not online communities themselves, since they do not provide the conditions for social networks to form. They are designed for browsing, not content creation². Chicago Crime Map is one such site. It maps Chicago police department's publicly available crime reports using the Google Maps API. The crimes are browsable by location, time, and type of crime, and even include some summaries pulled from another local Chicago website. However, there is no provision for users to comment on the crimes, chat about crimes, or tell stories about crimes that were not reported to the police. With no history of individual user interactions with the site, an online community cannot form.

The lack of community-forming mechanisms should not be understood as a criticism. As one user wrote, "I was living in Chicago at the time it came out, and (of course) used the listing to find out just how safe my neighborhood was. The information always had been available from the police department, but it was only in the context of a mapping application that I really was able to understand and internalize this data" [7]. At least for this person, the mapping of data to geography has social benefits.

Mashups and online communities

However, some newer websites are more complex. They are closer to social arenas than geospatial data browsers.

² Gmaps Pedometer [3] does allow users to create content (in this case, new jogging or cycling routes), but provides no mechanism for sharing those routes with other users.

They use mapping APIs and even data from other sites, but they focus on enabling social interaction between users. Platial.com, one such site, advertises itself as “the people’s atlas” – a space for public place-sharing. 43places.com, another such site, asks users to “Share stories about places in your city and around the world.” Wayfaring.com, a third site, has this mysterious tagline: “follow you, follow me.” And Plazes.com simply announces: “we know where.” This newer category of mapping sites sees maps as a means to an end -- independent online communities.

Not all of these new mapping sites are truly “mashups.” Plazes.com uses Flickr.com photo and Platial.com integrates feeds from Flickr.com, Upcoming.com, and Youtube.com. However, neither Wayfaring.com nor 43Places.com includes data from other sources. They are simply using mapping APIs to represent data contributed by their own users. Nevertheless, these new mapping sites indicate an emerging trend: online communities that explicitly allow and encourage users to repurpose geo-content on other sites. That is, they see themselves as *ingredients* for future geo-mashups. We can see this trend operationalized through:

- standardized data feeds (based on RSS or other standards)
- exposed functionality through APIs
- links to public bookmark aggregators such as del.icio.us and digg.com
- mechanisms to post community content to personal blogs

Online social spaces that had not previously included mapping functionality are also starting to provide feeds suitable for geo-mashup. The popular photo-sharing site Flickr.com introduced geo-referencing as a feature in September 2006 [1].

Building mashups

Mashups typically have followed a different development path than those of other information systems. After using successive generations of the early mapping mashup development, building our own mashups, and conducting informal observations of mashup development teams, we have distilled this account of a “typical” mapping mashup development cycle:

- 1) Developer has an idea and quickly builds a prototype
- 2) Developer waits for response, received through personal emails, mass media attention, reviews on colleague’s weblogs, and/or comments on his/her own weblog
- 3) Developer watches the progress and reception of other, similar mashups
- 4) Developer makes changes. These can include: adding or deleting data feeds, changing the user interface or choice of data presented.

Rather than the linear “waterfall” process of traditional HCI-oriented development or the rapid-fire teamwork of agile methods [18], in mashup development the “prototype” is what is released, and there are no final specifications. We can see this phenomenon in the longtime “beta” standing of some seemingly established websites (i.e., Plazes.com jokingly labels itself “still beta” on its homepage).

WHEN WORLDS COLLIDE

Developers report substantially altering the content and form of mashups for many reasons, including user requests and changes to access policies and protocols at data sources. The reported reason that concerns us most is the way in which remixing data streams can transgress social boundaries. For example, one industry observer reported that a site was altered because “correlating crime data which became sensitive when mixing it with another source” [15]. “Sensitive,” in this case of web services and geo-mashups, does not mean “illegal.” It suggests instead that remixing, visualizing, and repurposing publicly available data with mapping APIs can violate previously unsuspected social norms and understandings.

The recent furor within Facebook.com’s user community is a helpful example of what “sensitive” can mean within an established online community. According to the *New York Times*, online protests exploded after Facebook.com introduced a new feature that streamed recent activity within users’ social network to users’ homepages on the site. The *New York Times* quotes one upset user: “It’s like someone peeking in on my conversations. People now know exactly when you became friends with somebody...Before it took some extra effort” [17]. The information was already available within the social network; users apparently disliked the aggregation of the information into an automatically-updated feed.

Through our work with geo-mashups that aim to create new online communities, we have identified some typical features of mashup developments that violate design principles for successful communities, potentially creating “sensitive” situations.

A lack of precedent

All participants in the usability exercise brought with them assumptions about how some functions of an online community “should” work based on their engagement with other sites (notably, Flickr.com and del.icio.us). They repeatedly compared the access permissions, bookmarking functions, and social networking mechanisms OurPlaces.com to those of existing sites that they used. They could – and did – predict the probable results of choices they made about presentation of self in terms of user names, profile data, et cetera. However, when users had to make choices about revealing personally meaningful geodata, they often were often unable to forecast the likely results of their choices. Though more investigation is needed, we believe that this inability to forecast probable

results of their choices indicates the absence of a cognitive model for geodata-based online interaction.

Control over self-representation

Keeping some aspects of our characters or everyday lives private in social situations is a key aspect of self-representation [4]. Current geo-mashups extend past models of online disclosure of offline activities. From the creation of the Internet until the emergence of location-based network services, self-disclosure has primarily occurred through textual description and photographic documentation of one's own past activities or those of others. While personal geographic data has been available online – from Internet registry whois searches to web-based White Pages – the expense and difficulty of establishing geospatial interfaces has limited the spatial visualization of that data to specialists. Now, free API access has expanded that capability.

As we saw in the Facebook and Chicago Crime Maps examples, changes in the way online systems aggregate and represent information can have profound social effects. In the case of geo-mashups that produce or rely on content created by end users, that information can be intensely personal, even intimate.

When it comes to geographic or location-based data, definitions of what should be private are both personal and contextual [5]. Participants in the usability activity had similar – though varying – definitions of “private” geospatial information. To one, the address of her apartment was “private.” To another, cycling routes he might use as a fitness coach were “private.” A third felt that the locations of potential apartments he was considering renting were also “nobody’s business.” And another speculated that if he had a child, the location of the child’s bus stop would also be “private.”

The addition of photographs further complicated the definition of “private.” One participant who had been comfortable with a public listing of her home address became agitated after linking a photograph of her home with that listing, then decided that the listing itself should be private. While we should not generalize from one example, it does suggest that multimedia geo-data can complicate previously straightforward self-disclosures online. In general, however, participants were interested in sharing some of this “private” information within friendship and business networks, but feared public release.

Recall our previous discussion of the operationalization of remixing and repurposing designed into the interfaces of many geo-mashups. While we did not intend a thorough study of users’ attitudes towards geographic privacy, there are some implications for the remixing of geotagged information into contexts its creators did not expect. As Flickr creator Butterfield acknowledges in his announcement of Flickr geo-tagging [1], even if users consider a photo to be public, they may wish to keep the

location private. Our observations imply that the converse may also be true: a photo of a woman’s bedroom may be posted publicly to Flickr. But associating it with a specific geographic location may not be acceptable. And sharing any personal geo-data outside the context of its creation may also be unacceptable to users.

Persistence

In evolving and improving mashups, a developer must be free to add, subtract and reorganize component data sources. Communities, however, rely on persistent, reliable access to information over time. In the weblog world, permalinks serve this function: they guarantee a persistent, stable referent for conversation. This in turn helps support reputation systems based on patterns of user activity over time. When developers mix-and-match sources within an online community, they may either

- remove access to a previously stable historical referent upon which the community relies
- remove access to their information for other communities that depend upon it

Stability

In his research on Usenet newsgroups, Marc Smith has noted that “high levels of cross-posting” of messages to multiple forums “are indications that a newsgroup lacks a clear boundary, possibly indicating the lack of focused content and a stable population” [16, p. 208]. We speculate that frequently reconfiguring the data feeds on a mashup site can similarly defocus the identity of any community in formation there. As Kollock’s principles remind us, a community requires stable, visible boundaries.

In the mashup ecosystem, remixability is critical. Data streams can be imported and exported freely. In communities, however, participants generally tailor their contributions to the unique context of the people and affordances there at the time. Often, they do not consider how the information they share might be re-presented somewhere else or sometime in the future, and how its meaning might change.

IMPLICATIONS

We expect to see continued diversification among geo-mashups, both in their selection of data sources and in their strategies for integrating data viewing and community functions. Sensitive situations that arise around personal geodata are likely to affect all participants in the geo-mashup ecosystem, including the major data and API publishers, mashup creators and individual content contributors. If these situations are to be resolved in favor of sustaining community building, we believe the following design challenges will need to be met:

- Build design vocabularies to convey the nature of authorship and collaboration in mashups.

- Help novice users and contributors understand the configuration of a mashup and where their contributions fit in.
- Allow contributors to grant or withhold permission for their data to be included in outbound feeds.
- Promote the expectation that mashups can and should be fluid environments that may be reconfigured at any time.

Mashups emerge from the desire to appropriate and reconfigure interfaces and data sources. In responding to these design challenges, we hope that their ease of creation and access is maintained so that the original project that inspired so many people continues: giving individuals and communities the option to “mash up” their content with the applications, data and services that best fit their needs and interests.

REFERENCES

1. Butterfield, S. Great Shot – where’d you take that? http://blog.flickr.com/flickrblog/2006/08/great_shot_where.html. Posted August 28, 2006. Accessed September 12, 2006.
2. Chicago Crime Map. <http://www.chicagocrime.org/>. Accessed September 12, 2006.
3. Gmaps Pedometer. <http://www.gmap-pedometer.com/>. Accessed September 12, 2006.
4. Goffman, E. *The Presentation of Self in Everyday Life*. Doubleday Anchor, New York, 1959.
5. Housing Maps. <http://www.housingmaps.com/>. Accessed September 12, 2006.
6. Lederer, S., Mankoff, J, Dey, A.K. Towards a Deconstruction of the Privacy Space. Presented at the UBICOMP 2003 Workshop on UbiComp Communities: Privacy as Boundary Negotiation, October 12, 2003.
7. Lerner, R. 2006. At the forge: Creating mashups. *Linux J.* 2006, 147 (Jul. 2006), 10.
8. Kollock, P. 1996. *Design Principles for Online Communities*. Harvard Conference on the Internet and Society.
9. Mappr. <http://www.mappr.com> Accessed September 12, 2006.
10. Mashup (music). Wikipedia. http://en.wikipedia.org/wiki/Mashup_%28web_application_hybrid%29. Accessed September 12, 2006.
11. Mashup (web application hybrid). Wikipedia. http://en.wikipedia.org/wiki/Mashup_%28web_application_hybrid%29. Accessed September 12, 2006.
12. Platial. <http://www.platial.com>. Accessed September 12, 2006.
13. Plazes. <http://www.plazes.com>. Accessed September 12, 2006.
14. ProgrammableWeb. <http://www.programmableweb.com>. Accessed September 12, 2006.
15. Rubio, D. Mashups and Web Services. http://searchwebservices.techtarget.com/tip/0,289483,sid26_gci1216674,00.html?asrc=SS_CLA_299037&psrc=CLT_26. Posted September 19, 2006. Accessed September 21, 2006.
16. Smith, Marc. Invisible Crowds in Cyberspace: Mapping the Social Structure of the Usenet. In Smith and Kollock, eds., *Communities in Cyberspace*. Routledge, 1998.
17. St. John, W. When Information Becomes T.M.I. *The New York Times*. September 10, 2006.
18. Tomayko, J. and Hazzan, O. *Human Aspects of Software Engineering*. Charles River Media, Boston, MA, USA 2004.
19. Wellman, B. and Hampton, K. Living Networked On and Offline. *Contemporary Sociology*. Vol. 28, No. 6 (Nov., 1999), pp. 648-654