How a Personalized Geowiki Can Help Bicyclists Share Information More Effectively

Reid Priedhorsky, Benjamin Jordan, Loren Terveen

Department of Computer Science and Engineering
University of Minnesota, Minneapolis, Minnesota
{reid,benjordan}@umn.edu, terveen@cs.umn.edu

Abstract

The bicycling community is focused around a real-world activity – navigating a bicycle – which requires planning within a complex and ever-changing space. While all the knowledge needed to find good routes exists, it is highly distributed. We show, using the results of surveys and interviews, that cyclists need a comprehensive, up-to-date, and personalized information resource. We introduce the *personalized geowiki*, a new type of wiki which meets these requirements, and we formalize the notion of *geowiki*. Finally, we state some general prerequisites for wiki contribution and show that they are met by cyclists.

Categories and Subject Descriptors H.5.3 [*Group and Organization Interfaces*]: Collaborative computing, computer-supported cooperative work, web-based interaction.

General Terms Human factors, design.

Keywords Wiki, geowiki, geography, personalization.

1. Introduction

1.1 Bicycling and the Personalized Geowiki

The bicycling community is focused around *doing*: navigating a bicycle in the physical world. This activity raises interesting challenges. First, it is inherently geographic and typically local (i.e., people mostly ride in the area where they live). Second, the *planning* task – *deciding where to go, and how to get there* – is hard. It is hard because cyclists must navigate a transportation network largely designed for another purpose (driving motor vehicles) and because they must do so under continually changing conditions (weather, motor traffic, construction, etc.). Third, cyclists have significant individual differences in purpose, attitude, and abilities.

Many related domains exist. One is shopping in a mall: this is geographic and local, focused around real-world doing, and requires planning in a complex, changing space. Another is finding one's way around a new city or neighborhood: this too is geographic, local, and occurs in a complex, changing space. A third is natural resource management, for example monitoring off-road vehicle activity. In other words, the personalized geowiki has many applications beyond cycling.

The complexity of bicyclists' tasks results in complex information needs. Cyclists have a strong tradition of sharing information, but their existing practices are relatively inefficient. There is no comprehensive, up-to-date information resource that helps users find routes that meet their personal preferences. This is true even today, when at least four major vendors offer geographic web search and automated motor vehicle route finding[1] and there are hundreds of Google Maps "mashups", covering everything from where to find an apartment[2] to where crimes have occurred[3] to "If the Earth Were a Sandwich"[4] [7]. Why is there no Google Bicycle?

First, the cycling community is relatively small. In the United States, cycling is two to three orders of magnitude less frequent than driving, whether measured by number of trips or number of miles traveled per person [12]. Second, cyclists require detailed and continually changing information to plan their routes. Together, these two factors make the problem too hard for hobbyists and economically unattractive for businesses.

However, *wikis change the equation*. While it's still hard to gather and maintain the information, this work is distributed across many motivated users, rather than being the responsibility of the system builder.

WikiSym'07, October 21–23, 2007, Montréal, Québec, Canada.

[1] Google, http://maps.google.com; Yahoo, http://maps.yahoo.com; MapQuest, http://www.mapquest.com; and Microsoft, http://local.live.com.

[2] http://www.housingmaps.com

[3] http://www.chicagocrime.org

[4] http://www.zefrank.com/sandwich/tool.html

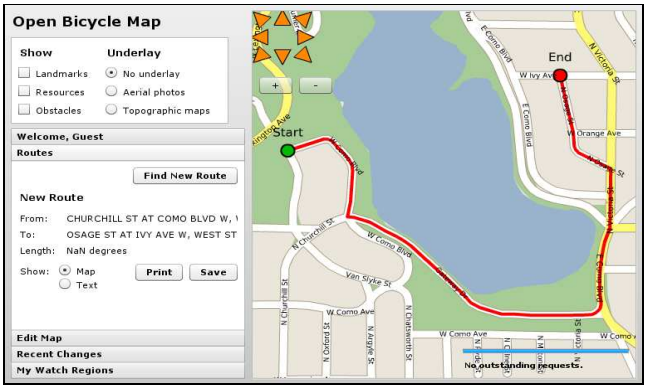

Figure 1. Screenshot of our current application, with an automatically generated route.

To meet the needs of such communities, we propose a new type of wiki: a *personalized geowiki*, which is a geowiki with personalized planning features. We have designed the user interface of our personalized geowiki for cyclists and are currently implementing it; Figure 1 is a screenshot of our application at the time of this writing.

1.2 The Geowiki

The personalized geowiki is an extension of the geowiki, which is itself an extension of the wiki. *Geowiki* is already a well-established term and is generally considered to mean a geographically contextualized wiki. However, we argue that, in a *true* geowiki, the geographic context itself, as well as links between geographic context and non-geographic data, enjoys wiki editing features – i.e., the structure of the map itself can be edited, not just items on the map – and this is the notion of geowiki that we extend.

Specifically, a geowiki is a geographically contextualized wiki which implements the following features:

1. Graphical **web-map interface** with standard web-map navigation operations.
2. **On-line map editability**, i.e., if geographic data or links involving geographic data are reasonably editable at all, they can be edited in the browser. For example, the names of lakes and the locations and connectivity of streets are editable, but aerial imagery is not, because aerial images, once loaded, typically are not edited by anyone.
3. **WYSIWYG editing** of geographic data, with a reasonably complete set of editing operations. While wikitext markup works reasonably well for textual data, geographic data is too complex to edit using text markup.
4. **Robust linking** to geographic data. For example, text which describes a street must be linked *to the street itself* rather than to *a region which is co-located with the street*. Co-location links are weak, breaking down both when geodata moves and when multiple geographic objects are themselves co-located.
5. **Comprehensive data monitoring**, i.e., revision history browsing and automatic watching of geographic data, non-geographic data, and links between data. We introduce the *watch region*, a generalization of the watch list feature standard in traditional text wikis: users graphically indicate regions (or byways) they care about. When editing changes data within a watch region, the system then notifies the user who defined it.

We are not aware of any true geowikis, though numerous partial implementations exist (which we discuss further in Section 2 below). However, as we show, this complete editability is key to the success of wiki as a cycling resource.

In the remainder of the paper, we sketch related work, present the design of an empirical study of cyclists, identify their unmet needs, show how these needs can be met using a personalized geowiki, and explain why we expect cyclists to contribute to such a wiki. We also illustrate the key features of our personalized geowiki.

2. Related Work

The wiki model, where any random visitor can make unreviewed changed to a website, is a relatively new method of collective work. The Wikipedia[5] online encyclopedia is the most famous example of this model and is generally quite successful; e.g., Giles found that the accuracy of Wikipedia and *Encyclopedia Britannica* were roughly equivalent [5]. Other work has shown that users are as effective as experts in reviewing other users' work [3], and that the wiki model and traditional review-before-publication result in the same quality, but the wiki model achieves it faster [4].

Partial implementations of the geowiki concept exist. For example, PlaceOpedia[6] maps Wikipedia articles to locations, OpenGuides[7] is a wiki travel guide, WikiMapia[8] allows users to tag places and rectangles with wiki-editable information, and there are numerous other similar sites. However, these sites implement only geographically contextualized wiki information without offering any editability of the geographic context (i.e., they implement only the first of our five geowiki features).

Other sites allow sharing of WYSIWYG-edited geographic and geographically contextualized data but do not have wiki features. Gmaps Pedometer[9] allows sharing of routes, and Google My Maps[10] allows sharing of points, paths, and polygons which can be annotated with text, images, and videos. In both cases, information is geographically contextualized using co-location only, though because

[5] http://www.wikipedia.org
[6] http://www.placeopedia.com
[7] http://openguides.org
[8] http://wikimapia.org
[9] http://www.gmap-pedometer.com
[10] http://maps.google.com

routes sometimes deviate from stored geography, this is necessary in a fixed-geography system.

Open Street Map[11] is an effort to build street maps using the wiki model. This project is a fuller geowiki, implementing four of our five geowiki features: a web-map interface, on-line map editability[12], WYSIWYG editing, and some robust linking. Geographic data is robustly linked with itself, but there is no non-geographic data at all, making that robust linking issue moot. OSM's programmatic interface would permit such robust linking and could be used in the construction of geowikis, though we are not aware of any which make use of this feature. Similarly, while the programmatic interface makes available the revision history of geographic data, this is not exposed in the Web interface, and no automatic watching is available.

The geographic information systems (GIS) community has also proposed various types of collaboration in mapmaking [2, 6, 9, 10]. However, GIS researchers have not considered the open editing that is the essence of the wiki model.

Finally, there are many websites dedicated to bicycle navigation. Bikely[13] allows users to create and share routes. TopoRoute[14] does this and also offers automated route finding, but this is of limited help because it uses motor vehicle routing data. byCycle[15] offers cyclist-oriented automated route finding but offers only the route choices "Normal" and "Safer", which does not fully meet the routing needs of cyclists, as we show below.

3. Research Design

We studied the cycling community using surveys and interviews, recruiting subjects from the Minneapolis-St. Paul metropolitan area. We sent invitations to three local cyclists' discussion lists with about 950 total members and also encouraged recipients to forward the invitation. All participants were over 18 and had spent at least 200 miles or 25 hours cycling within the local area during the past year.

73 respondents finished the survey, and most questions had about 75 responses. 68% of survey respondents were male and 32% female; most were between 18 and 64 years of age with a roughly uniform distribution. Survey questions focused on attitudes regarding map errors and existing planning and navigation methods. We also used the survey to recruit for interviews.

We completed 19 semi-structured interviews lasting 60 to 90 minutes each; 13 subjects were men and 6 women. Some questions elaborated on issues touched on by the survey; others explored topics more suitable to an interview setting, such as attitudes toward the wiki model and privacy concerns. We also presented lo-fi paper mockups of three core personalized geowiki features – the wiki geodatabase, bikeability ratings, and monitoring of edits. Finally, each subject sketched a map of a familiar route to provide non-verbal data on geographic thinking. Interviews were recorded by two note takers and subsequently coded.

4. Unmet Needs

We identify three key unmet needs: no comprehensive and up-to-date information resource, no automated route finding, and no personalized ratings of byway bikeability. *Resource* in this context means anything bicyclists use to plan routes, common examples being bicycle maps or guidebooks. We use *byway* to mean the smallest identifiable segment of a bike route, e.g. a section of bike path between intersections with roads or other paths.

4.1 No Comprehensive, Up-To-Date Resource

To plan a route, cyclists need to know how they can travel through geographic space *now* and what they will find within the space *now*. Some of this information isn't recorded at all, and the rest is distributed across numerous electronic and non-electronic resources. For example, some bike trails and bike lanes are mapped by the state, others are mapped by municipalities, and still others are mapped by park boards.

Bicycle travel is accomplished using byways, so cyclists need to know where the byways are and how they connect. Byways clearly include dedicated bike trails and many (but not all) roads. However, subjects suggested several additional classes of surfaces that are sometimes byways: alleys, sidewalks, parking lots, etc. More classes certainly exist, and it would be difficult to enumerate them even if all possible cycling routes were known. Regardless, current information resources do not include these types of byways. Furthermore, only actual byways should be included, in order to avoid overwhelming the resource and users with unimportant information. For example, most sidewalks should be excluded, but some, such as those that bridge highways, are key byways. *Cyclists know the locations and properties of byways because they themselves travel upon them.*

The locations of *landmarks*, *resources*, and *obstacles* are also important to bicyclists. 16 out of 19 subjects mentioned landmarks. These, in addition to street names, are used for orientation and navigation, and subjects cited objects like businesses, highways, and water bodies, i.e., both point and non-point landmarks. Resources are things helpful to cyclists in some way; 11 subjects cited a wide variety of resources including restrooms, water sources, and restaurants. Obstacles cause cyclists difficulty or frustration. 12 subjects mentioned a variety of obstacles, including construction and traffic lights.

A few of these items – e.g., water bodies – are found on current information resources, but most are typically not. It is also time-consuming to identify which landmarks, resources, and obstacles are important; for example, members

[11] http://www.openstreetmap.org

[12] On-line editing uses a Java applet, somewhat limiting its reach.

[13] http://www.bikely.com

[14] http://toporoute.com

[15] http://bycycle.org

of the local cycling club frequently use Dairy Queen restaurants as landmarks. However, *cyclists know which of these things are important and where they are located because they themselves use or avoid them.*

These things change, both in location and other properties. Construction begins and ends, byways close and new ones open, and businesses close or move. A useful resource must be up-to-date. However, current (typically centrally-maintained) resources go stale. 11 subjects mentioned this or its consequences as a problem. *People who notice problems can fix them or at least point them out.*

These observations – that existing information repositories are widely scattered and incomplete, and that cyclists themselves know the important information – motivate the distributed editing approach of wikis.

4.2 No Automated Route Finding

Cyclists told us that they want automated route finding, i.e., "find me a good route from point A to point B": 4 of 19 subjects mentioned this desire specifically, 5 described a problem that could be solved with such a tool, and 5 expressed dissatisfaction using motor vehicle route-finders for cycling (11 total). No such tool is available.

Tools for automated motor vehicle route finding are very successful, but our interviews reveal that they are unsuitable for cyclists because they do not know about all byways and they do not take into account the complexity of cyclists' routing needs. While a handful of bicycle-specific routing tools exist, most notably byCycle, they suffer from the same basic problems of incomplete coverage and simplistic routes.

Automated bicycle route finding can use the same basic approach as motor vehicle routing: the network of byways is represented as a weighted graph, and a minimum-weight path is calculated. However, while motor vehicle routing uses simple factors like distance and travel time to calculate edge weights, effective bicycle weights are based on many more additional factors. Subjects cited factors both objective and subjective, including the locations of hills, presence and quality of pavement, motor vehicle traffic levels, motorist attitude, and numerous others.

4.3 No Personalized Bikeability

Furthermore, people's ratings of and preferences for any given byway are a matter of personal taste: cyclists do not agree on which quality factors should be considered and what their relative importance should be. This led 8 of 19 subjects to question the utility of existing generic bikeability ratings (in the cases where they are available), expressing either a general concern that their own notion of what made for a good byway might differ or else that they had actually encountered ratings they disagreed with. Additionally, existing resources offer bikeability ratings for only a subset of byways. What cyclists really wanted was a way to get personalized ratings for *any* byway.

Figure 2. Our current automatic route finding interface. Edge weight is calculated as a weighted sum of the (normalized) priority factors.

5. The Personalized Geowiki

A personalized geowiki can meet these three unmet needs. In the context of cycling, a personalized geowiki has four core features, the first two forming the geowiki and the latter two adding the necessary personalization:

1. **Wiki map.** A user-maintained geographic information system (GIS) storing the byway network. Editing byways is done with a simple interface analogous to standard vector drawing programs. High-resolution aerial photos help orient users[16]. Also, precise positioning of byways is not as important as connectivity, which is easier to edit.

   This component is closely related to Open Street Map. While OSM builds street maps from scratch in order to avoid the great cost of base maps in Europe, in our American locale, good base maps of both streets and cycling infrastructure are available for free and high-resolution color aerial photography for a nominal fee. We therefore use these as a starting point for our wiki map.

2. **Wiki geodatabase.** A user-maintained database of geographic objects important to cyclists: the locations and details of landmarks, resources, and obstacles. These are contextualized using the wiki map and robustly linked to geographic objects in the map and to one another.

   Monitoring of both the wiki map and wiki geodatabase is achieved using watch regions. When editing occurs inside the geographic boundary of a watch region, whoever defined the region is notified, regardless of whether or not the edit was geographic.

3. **Route finding.** Automatic generation of routes through the byway network based on personal preference and immediate needs. Figure 2 shows our current route finding interface design.

[16] A few subjects noted that they already use aerial photos to plan.

Willing to spend ...	If other users saw corrections ...	
	in six months	immediately
1 minute or more	67%	96%
5 minutes or more	44%	73%
10 minutes or more	21%	43%

Table 1. Willingness to correct map errors.

As noted above, there is strong personal variation on what makes a good byway. Our approach is to isolate a few objective metrics generally considered useful – for example, distance and hilliness – and address the rest with a purely subjective "bikeability" metric, as discussed below.

4. **Personalized bikeability ratings** of *every* byway. Clearly, any individual cyclist would be able to give ratings for only a small subset of the byways in the system. However ratings for the remaining byways can be *estimated* using a collaborative filtering recommender system [8, 11]. After a cyclist rates familiar byways, the system finds other cyclists with similar preferences on those byways and uses these users' ratings to estimate ratings of unfamiliar byways, the result being a personalized bikeability map of the entire byway network for each individual cyclist.

   This requires a critical mass of ratings, but while this is collected, estimation can rely on simple aggregate popularity or generic ratings based on objective metrics.

Relying on cyclists themselves to enter and maintain data seems plausible and attractive. They know what information is important to cyclists and why it is important, because they use it (or wish they had it) on every ride. Therefore, a comprehensive, up-to-date personalized geowiki is possible – but only if cyclists are willing to contribute.

6. Will Cyclists Contribute?

A wiki is successful only if users do the necessary work: creating, editing, and monitoring the data. We discuss four prerequisites for getting work done: *propensity to share*, *trust*, *propensity to monitor*, and *privacy*.

Our subjects expressed a strong *propensity to share* what they knew with other cyclists. 83% of survey respondents reported asking other cyclists for route planning help. They were also willing to spend substantial effort correcting map errors, especially if their work was available to others immediately; see Table 1. Furthermore, some cyclists already spend considerable effort on helping one another. In 2006, the local recreational cycling club's 100 volunteer ride leaders led over 1,400 rides [1]. The duties of a ride leader are to obtain or create a route, scout the route regularly, maintain and distribute maps and turn lists, and lead rides along the route – many hours of effort on behalf of other cyclists.

In our interviews, we asked subjects if they would share routes with the general public; 17 of 19 subjects said yes. 14 mentioned textual information they would share. When asked if they would rate the bikeability of byways to help other cyclists, 13 subjects said yes. When asked if they believed that they *knew how to* correct map errors they had encountered themselves, 14 said yes, and when asked why they would fix map errors, 7 gave helping others as a reason.

We also observe cyclists using existing collaborative technology, however cumbersome, to share information. For example, the following routing help request recently appeared on a local cycling web forum:

> How do I get to Khan's in Roseville, from the St. Paul campus [of the University of Minnesota], without being killed in traffic?
>
> Khan's Mongolian Barbeque
> 2720 Snelling Ave N
> Roseville, MN 55113

It generated 8 responses, including (a) a detailed Gmaps Pedometer route posted within about 7 minutes, (b) another Gmaps Pedometer route recommended for use after dark, (c) an endorsement of the second route, and (d) a warning that a specific byway contained many potholes. Another thread began with a warning that a particlar bridge was closed, and a third, titled "streets to avoid", had 24 posts.

These results reflect the tradition of information sharing within the bicycling community. We expect the geowiki initially to benefit from this tradition and subsequently to allow for even more effective sharing.

6 interview subjects hesitated to *trust* the wiki model, noting its vulnerability to vandalism and misinformation. These concerns are valid. However, as we have noted, successful and trusted wikis have emerged. When they do, wikis work because of the motivation of users to correct mistakes and vandalism, aided by mechanisms to help notice and fix the damage. Subjects were enthusiastic about using watch regions to monitor areas they knew and cared about: 16 identified specific geographic regions they would want to watch, typically near their homes or on routes they rode frequently.

We asked subjects about any *privacy* concerns they had with the wiki information sharing model. 10 subjects mentioned various concerns. Most could be mitigated with standard techniques (e.g., using pseudonyms instead of real names) or though selective information withholding (e.g., by not sharing a few revealing routes such as those used for commuting). 7 subjects told us that they wanted to keep the locations of their homes private; however, most were willing to reveal nearby locations, e.g. "If some creepy guy wants to come to the corner of my block, more power to him".

Traditional wikis expose editing history, labeling each action with the pseudonym of the user who performed it; if geowikis do this as well, it might be possible to infer the locations of users' homes, areas they're likely to be found, or other potentially sensitive places. What is and is not acceptable to users, specific information that can be inferred, and techniques to mitigate risks should be explored.

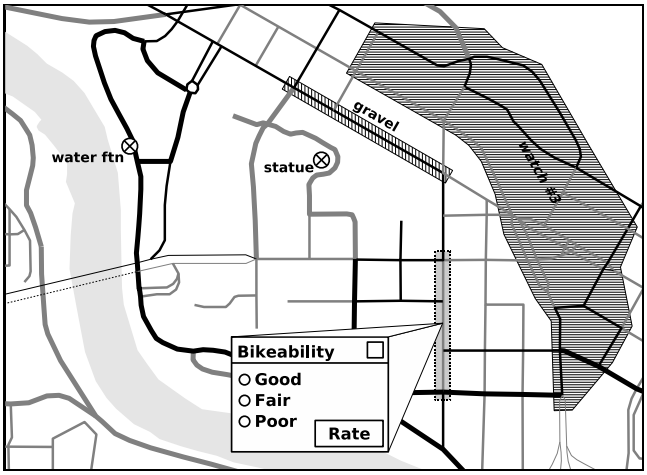

Figure 3. Mockup of the personalized geowiki we envision. A landmark (*statue*), resource (*water ftn*), and obstacle (*gravel*) are shown, as well as a watch region; clicking on any of these would bring up further details. Bolder and black indicates byways with increasing user-rated bikeability, while bolder and gray indicates increasing *system-estimated* bikeability. The highlighted byway is in the process of being rated.

7. Status and Next Steps

The results of our study, including subjects' positive response to our interface design, support the potential of the personalized geowiki approach. Our planned design is illustrated in Figure 3. It features a zoomable, pannable map with different layers of information that can be viewed or hidden, and (in most cases) edited.

To realize the potential of a personalized geowiki, we must complete our implementation and deploy it for use by cyclists. This is challenging, as existing web mapping APIs like Google Maps[17] provide insufficient functionality, offering only manipulation of data *co-located with* an unchangeable, vendor-controlled transportation network. Our study shows that we must implement not only deep access to the connectivity and properties of byways, but all of this must be editable as well.

Therefore, we are implementing our design using lower-level APIs. The client is written using the Adobe Flex[18] framework and runs in the browser on the widely deployed Flash Player 9 virtual machine[19]. It communicates over HTTP with custom mapping and route-finding servers written in Python[20]. Data is stored using the industry-standard PostGIS spatial database manager[21].

[17] http://www.google.com/apis/maps/

[18] http://www.flex.org

[19] http://www.adobe.com/products/flashplayer/

[20] http://www.python.org

[21] http://postgis.refractions.net

In summary, we show that cyclists need a comprehensive resource that contains all information relevant to planning a route and enables personalized route finding through the network of byways. While cyclists have a strong tradition of sharing information, their existing methods are limited; a wiki builds on this tradition of sharing and promises much greater coverage and reach, allowing individual cyclists to efficiently share with many others. We look forward to deploying our personalized geowiki and meeting this promise.

Acknowledgments

We gratefully acknowledge the time and effort of our subjects as well as the support and feedback of members of our research group, particularly Pam Ludford. This work has been supported by the National Science Foundation, grants IIS 03-07459 and IIS 05-34692.

References

[1] Angle, K. Annual recognition dinner 2007. *Twin Cities Bicycling Club News* (2007), 7. http://www.pdwebworks.com/tcbc/news/.

[2] Balram, S. et al. A collaborative GIS method for integrating local and technical knowledge in establishing biodiversity conservation priorities. *Biodiversity and Conservation, 13* (2004), 1195–1208.

[3] Cosley, D. et al. How oversight improves member- maintained communities. In *Proc. CHI*. 2005, 11–20.

[4] Cosley, D. et al. Using intelligent task routing and contribution review to help communities build artifacts of lasting value. In *Proc. CHI*. 2006, 1037–1046.

[5] Giles, J. Internet encyclopedias go head to head. *Nature, 438* (2005), 900–901.

[6] Gillavry, E. Collaborative mapping: By the people, for the people. *Society of Cartographers Bulletin, 37*, 2 (2004), 43–45.

[7] Google maps mania. http://googlemapsmania.blogspot.com.

[8] Konstan, J. et al. GroupLens: Applying collaborative filtering to Usenet news. *CACM, 40*, 3 (1997), 77–87.

[9] Li, S. Results of CSCW supported collaborative GIS data production: An internet-based solution. In *Proc. Hypertext and Hypermedia*. 2002.

[10] Rybaczuk, K. GIS as an aid to environmental management and community participation in the Negril Watershed, Jamaica. *Computers, Environment and Urban Systems, 25*, 2 (2001), 141–165.

[11] Sarwar, B. et al. Item-based collaborative filtering recommendation algorithms. In *Proc. WWW*. 2001, 285–295.

[12] U.S. Dept. of Transportation. National household travel survey. http://nhts.ornl.gov/2001/ (2001).