

Specialist Meeting on Volunteered Geographic Information Position Paper

Reid Priedhorsky

Department of Computer Science and Engineering, University of Minnesota
reid@umn.edu

At the University of Minnesota, we are exploring VGI in the context of an online bike map: a *geowiki* for bicyclists [5]. We are building a system to enable cyclists to collaboratively build a database of geographic information relevant to them, including a fully editable map. This VGI will then be used as input for personalized route finding.

In order to build a successful repository of volunteered information (VI), geographic or otherwise, four prerequisites must be met:

1. **Utility.** The information must be useful.
2. **Motivation.** Volunteers must be willing and motivated to share.
3. **Correctness.** The information must be largely correct.
4. **Usability.** Volunteers must be able to use the repository-building system.

The need for Prerequisite 1, Utility, is obvious, but this critical consideration must not be neglected when designing a VI system. In other words, the trendiness of VI should not interfere with a rigorous evaluation of alternative information gathering methods.

A useful model for considering Prerequisite 2 is the *collective effort model* [1]: people will do work when they believe that their efforts will result in outcomes that they value. In a collaborative context such as a VI system, this means that people will do work *only* when they both value the group outcome and believe that their efforts will meaningfully further that group outcome.

In our work with cyclists, we found that the group outcome – a comprehensive, up-to-date navigation and route planning resource designed expressly for cycling – is highly valued, and that cyclists believe that they collectively have the knowledge necessary to build such a system, that no other group does, and that as individuals each have unique information which no one else can contribute. We also believe that cyclists would be motivated to contribute because they told us in interviews that they would, and because we observe cyclists using existing technology, however cumbersome, to share geographic information.

Privacy concerns do not seem to be a meaningful obstacle to our VGI system. Cyclists noted limited geography-driven concerns for privacy: mostly, that someone might use their artifact trail within the system to infer the location of their home. It remains to be seen what can be actually inferred, to what accuracy, and what degree of geo-anonymity is acceptable.

Prerequisite 3, Correctness, can be defeated in two ways: intentionally and unintentionally. Intentional incorrectness, i.e. vandalism, is a perennial problem for well-known VI systems; however, it seems to be manageable. For example, Wikipedia is aggressively targeted by vandals due to its high visibility – about 5% of its edits are clearly damaged (assessing intent, required to label damage vandalism, is very difficult). However, damage is repaired quickly [4]: half of incidents are essentially never seen, and long-duration,

highly viewed damage is very rare. The probability of encountering a damaged article is currently about 0.7%; this probability grew exponentially until mid-2006, when widespread autonomous anti-vandalism software was introduced. It remains to be seen whether the growth has been permanently halted. Regardless, it seems that VI systems which are less visible are also much less likely to be the targets of vandals.

Unintentional incorrectness occurs when users enter incorrect information *and* the errors remain uncorrected. In other words, users must make no errors (unlikely), or they must check and repair each others' work. In the context of Wikipedia, the former is clearly untrue, but the system overall produces largely correct content [2].

When geography is introduced, correctness becomes a significant concern. (GIS professionals we encounter are consistently horrified when we suggest that average users should be allowed to directly edit our geodata.) However, in our interviews, most cyclists expressed enthusiasm for monitoring geodata in areas with which they were familiar and also thought that they would be able to fix map errors that they had identified. Other important considerations are that it's easier to point out mistakes than fix them, broadening the group of people who can effectively help, and sometimes correctness requirements can be relaxed (for example, in our system, how streets and trails connect with one another is more important than the preciseness of their geometry).

Finally, Prerequisite 4, Usability, is also seemingly obvious, but this consideration is frequently neglected in technology systems. Bad usability is a potent contribution killer, and even systems as simple as a door are frequently misdesigned [3], e.g. clearly inviting pushing when the proper action is pulling. Furthermore, the presence of geography adds an additional layer of difficulty in designing for good usability.

This issue is particularly important in the case of collaborative systems, because an overwhelming majority of users will contribute only casually if at all (a "power law" distribution of contributions). For example, nearly half of the value of Wikipedia was contributed by only 0.1% of those users who edited it at least once [4] (and many more never edit at all). Because the long tail of infrequent contributors is so long, lowering barriers to contribution even a little bit may result in much increased contribution.

In light of these prerequisites, it is clear that the *users* of any VI system – volunteers who will (hopefully) contribute useful, correct information – must be carefully considered from the earliest phases of design. *A VI system is not a magic bullet* for gathering information cheaply.

References

- [1] Gerard Beenen et al. Using social psychology to motivate contributions to online communities. In *Proc. ACM Conference on Computer Supported Cooperative Work (CSCW)*, pages 212–221, 2004.
- [2] Jim Giles. Internet encyclopedias go head to head. *Nature*, 438:900–901, 2005.
- [3] Donald Norman. *The Design of Everyday Things*. Basic Books, New York, 2002.
- [4] Reid Priedhorsky et al. Creating, destroying, and restoring value in Wikipedia. In *Proc. International ACM SIGGROUP Conference on Supporting Group Work (GROUP)*, 2007. To appear.
- [5] Reid Priedhorsky, Benjamin Jordan, and Loren Terveen. How a personalized geowiki can help bicyclists share information more effectively. In *Proc. International Symposium on Wikis (WikiSym)*, 2007. To appear.