

## Proposal Outline

### General research question:

#### **How does social language use affect information sharing in user-contributed content repositories?**

### Specific research questions

#### **1: How does common ground (from shared past experience) affect language choices in a collaborative naming task?**

- language use by friends v. strangers in a collaborative naming task (taboo)
- same-gender pairs of participants who were either friends or strangers played taboo (in different media conditions)
- results will tell us whether common ground affects language choices. this type of common ground has typically not been explored in a lab setting; it is really more of a subject variable than something "manipulated". but it is closer to the kind of common ground we're talking about in an online repository used by a group, than community membership or physical & temporal copresence
- data is collected, quantitative & content analysis is underway. can also transcribe and analyze face-to-face condition

#### **2: How do groups organize and share information via an online repository of user-contributed content? What problems do individual members encounter during information seeking within the repository?**

*A: describe how people are using CTools project sites, and the situations in which information seeking occurs (ctools interviews & log data)*

- results will provide context, allowing us to describe the users, the sites, and the environment in which it occurs, for a specific real-world system in use by thousands of people
- data is collected, transcribed, and cleaned. analysis will begin when del.icio.us paper for GROUP is finished

*B: information seeking in CTools project sites*

- data collection will take place in early fall 2007 at the latest; possibly over the summer if suitable project sites are found whose members are around and active over the summer
- users will be asked to do specified-target information seeking tasks in their ctools project site. targets will be selected after initial interviews with an 'informant' from the site. subjects will 'think aloud' during the task. targets will be selected according to dimensions such as frequency and recency of use. issues of counterbalancing and "contamination" still have to be worked out. "type" of site to be used in this study is also yet to be worked out; three possible ways to look at it are based on: content analysis of information in site, usage log data, stated purpose.
- audio and screen actions will be recorded for analysis. it should also be possible to gain access to ctools event logs for the time periods in which the search tasks took place.
- results will allow us to quantify and describe the problem in greater detail, in the context of the particular system of interest.

**3: How do the source of common ground and the intended audience affect file and folder name choices, and the effectiveness of those names for future information sharing?**

*A: the effect of 'community membership' common ground and intended audience on the creation of filenames*

- data collection will take place in summer 2007. development of a system for conducting the experiment and collecting the data is currently underway. pilot will take place ASAP.
- participants will be recruited in ann arbor; and will either have 'community' membership common ground or not. precisely how this will be operationalized has yet to be determined, but it will probably consist of a common hobby interest (like amateur gourmet cooking for example) so that content can be selected appropriate to that interest.
- independent variables: common ground (y or n), instructions (none, for yourself, for others)
- method and procedure, phase 1: subjects will view a series of 'content items' and will give them a filename and then organize them into folders that they create themselves, while thinking aloud
- quantitative measures: time, similarity/distance measure of file structure, interuser agreement on filenames (other stuff?)
- qualitative measures: content analysis of file and folder names, and verbal protocols
- method and procedure, phase 2: after a few weeks have passed, subjects will try to find the content items using some of the possible 'categories' of structures that were created; for example, their own structure, someone else's structure they created for themself, etc., while thinking aloud
- quantitative measures: time, number of "clicks", wrong paths taken, etc.
- qualitative measures: analysis of verbal protocols

*B: the effect of 'shared past experience' common ground and intended audience on the creation of filenames*

- data collection will take place toward the end of fall semester 2007
- participants will be either be members of project groups in a course like SI 501 that has a semester-long group project, or be SI students that took the course the previous year, or be non-SI students. this adds "shared past experience" common ground as a level of the 'common ground' independent variable, for students who are working together in the same project group. students who took the course in different years can be expected to have community membership common ground. materials (content items) will be 501 or user-centered-design related
- independent variables: common ground (shared past experience v. community membership v. none), instructions (none, for yourself, for your group)
- same basic method and procedure and measures as above