Terminology and Evaluation: Two Challenges for Awareness Systems Research

Position paper for CHI 2005 Workshop on Awareness systems: Known Results, Theory, Concepts and Future Challenges

A.J. Bernheim Brush
Microsoft Research Community Technologies Group
One Microsoft Way
Redmond, WA 98052
ajbrush@microsoft.com

ABSTRACT
Technology offers many ways to help us stay aware of a wide variety of things we are interested in. Based on my research experience I believe that there are two challenges for awareness systems research that we should address in the CHI 2005 workshop: terminology and evaluation. Clarifying the wide variety of terms related to awareness systems will help the research community communicate more easily about their work and provide a framework for this rich design space. By exploring the best practices for evaluating awareness systems, we can work together to identify successful methods of evaluation and determine ways to compare systems that attempt to provide similar types of awareness.

INTRODUCTION
Awareness is a popular, (perhaps overused), term in human computer interaction literature. A search on "awareness" at the ACM digital library hints at how many different adjectives we apply to the term. The titles and keywords from just the top 20 hits include eleven different terms related to awareness: activity awareness, context-awareness, contextual awareness, group awareness, groupspace awareness, information awareness, peripheral awareness, presence awareness, situational awareness, social awareness, and task awareness. It is time to ask ourselves, what is an awareness system? Or more appropriately, what are the different types of awareness systems?

One thing is clear; there is interest in providing awareness about a range of things from people to information, making this an exciting and dynamic space for further research. Based on my personal experience in two research projects, I believe the CHI 2005 workshop on awareness systems should address two important challenges: *terminology* and *evaluation*.

While perhaps seeming merely semantic, the lack of established terminology in how we as a research community talk about different types of awareness makes it more difficult to clearly communicate the type of awareness that a particular system strives to provide. The workshop also offers a valuable opportunity to tackle the question of *evaluation.* What are the best ways to show that a particular system improves awareness of some object or information? As a research community, we need to explore evaluation methods that facilitate comparisons between different awareness systems. So, for example, two different systems providing awareness of family members could be compared and contrasted.

I now briefly describe my previous research projects which motivated my interest in addressing the challenges of terminology and evaluation. I then describe the aspects of these challenges I feel would be valuable to discuss at the workshop.

MOTIVATING RESEARCH PROJECTS
In two different projects I have focused on providing awareness to support collaboration in a work setting. In the first project, we studied two methods for providing awareness of annotations made on software specification documents to the team members responsible for reviewing them [1]. In a field study, I compared email notifications to displaying information using the Sideshow peripheral awareness system [4]. The email notifications alerted users to new annotations, provided the content of the annotations, and included links directly to the annotations in the context of the document. The annotation 'ticket' on the Sideshow sidebar constantly showed the number of total annotations and those 'new' that day for a particular document. By mousing over the ticket, users could get details and also links directly to the annotations in the document.

We found that the notifications using email and peripheral awareness both increased awareness of annotations on documents and participants used the notifications in many different ways including active monitoring and more casual tracking. More relevant for this workshop, the research highlighted for me the importance of exploring different methods of providing awareness to the same information (e.g. email vs. peripherally) and also the challenges in evaluating awareness, particularly when using logging data.

For example, if a product manager glances at the awareness information on her desktop sidebar and learns that no new annotations have been added to the specification document for which she is responsible, then she knows she does not need to visit the document. Through the awareness mechanism, she has received valuable awareness information that results in no action that can be logged and analyzed. This leads to a reliance on survey data from participants about their "sense" of awareness. While many research projects use surveys, it can feel somewhat unsatisfying as there are known problems with self-reported data.

In another project, I evaluated the role of 'today' messages, short emails sent daily by members of a project team to each other, in providing task awareness among group members [2, 3]. For this project, one main goal was contrasting the use of a very lightweight method of providing awareness (email) with other more heavyweight awareness systems. However, the lack of clear terminology made communicating this more difficult than we initially anticipated.

This project also highlighted for me the challenges in evaluation. During our field study of the use of today messages in six small workgroups we surveyed and interviewed our participants. I believe we would have benefited from having a standard survey instrument or instruments so that our results could be more easily compared with evaluations of other awareness systems.

These two research experiences led me to believe that the awareness research community as whole could greatly benefit from discussions around terminology and evaluation. In the next two brief sections I expand on how I envision the workshop might address these challenges.

TERMINOLOGY

As I have described, in my view as a research community we do not have a clear notion of the types of awareness that systems may be attempting to provide. If possible, perhaps an existing framework could be borrowed from another discipline. Or a particular framework from human computer interaction literature could be more widely publicized. For example, in Dix et al. [5, pg. 700] the authors breakdown types of awareness related to "who is there," "what has happened," and "how did it happen."

I believe in the workshop it would beneficial to discuss how different awareness terminology relates to each other and how to begin a process of standardizing the existing terminology. I think we also need to be careful to emphasize the differences between "what" a system provides awareness of and "how" it does it. For example, the term "peripheral awareness" describes a mechanism for awareness in contrast to "task awareness" which describes the type of information for which the system is trying to provide awareness. Clear terminology and a framework of different types of awareness systems would improve our ability to communicate with one another about our research, help us better understand where research opportunities exist, and allow us to more easily compare and contrast our research.

EVALUATION

Evaluating awareness systems is critical in understanding whether or not an innovative new system provides awareness or how it might need to be improved. However, in my personal experience and in reading other research on awareness systems I have been struck by the challenges in carefully evaluating an awareness system. Many projects rely on self-reported survey data or the "lots of people used it" metric.

I would like to work as a group to brainstorm, catalogue and discuss methods of evaluating awareness systems to determine a set of best practices. We must, of course, recognize that with many different types of awareness systems different methods of evaluation will be appropriate. By working together we can leverage the collective experience of workshop participants to identify successful approaches and outline possible new ones.

I believe we must also address the question of how to structure evaluations of awareness systems so that they facilitate comparison between different systems that attempt to provide awareness of the same information. The workshop is an ideal place to discuss and debate methods for doing this, perhaps by developing standard survey questions.

CONCLUDING REMARKS

Technology offers great promise to help people stay aware of many things that interest them, from other people to a wide range of information. However, as a research community I believe there are two key challenges for awareness systems research we must address: terminology and evaluation. By coming together and clarifying the wide variety of terminology used around awareness systems we can more clearly outline the expansive and varied design space. A discussion of methods for evaluation and ways to enable comparisons of awareness systems will help us ensure that we are building systems that truly benefit our users.

In my current research focusing on digital family calendars and coordination among family members, I expect the question of how and when to provide awareness of others will be an important consideration. I would appreciate a clear way to describe the type of awareness the system attempts to provide, the method it uses to do so, and having evaluation methods that allow comparison with other techniques for providing a similar type of awareness.

BIOGRAPHY

A.J. Bernheim Brush is a researcher in the Microsoft Research Community Technologies Group. She has many different research interests related to computer supported

collaborative work. Currently she is studying family calendaring focusing on how families coordinate and collaborate to accomplish the tasks of daily life.

ACKNOWLEDGMENTS

I thank my collaborators in the research described in this paper, David Bargeron, Alan Borning, JJ Cadiz, Jonathan Grudin, and Anoop Gupta, for their assistance, guidance and thoughtful discussions. I also thank my colleagues in the MSR CTG group for their support and encouragement.

REFERENCES

1. Brush, A., Bargeron D., Grudin, J., and Gupta, A., Notification for Shared Annotation of Digital Documents. *Proc. CHI 2002*, ACM Press (2002), 89-96.
2. Brush, A.J. and Borning, Alan, 'Today' Messages: Lightweight Support for Small Group Awareness via Email, *Proc. HICSS-38*, CD-ROM, 10 pages. 2005
3. Brush, A.J. and Borning, Alan 'Today' Messages: Lightweight Group Awareness via Email, Proc. CHI 2003 Extended Abstracts, ACM Press (2003) 920-921.
4. Cadiz, JJ, Venolia, G., Jancke, G,, and Gupta, A., All ways aware: Designing and deploying an information awareness interface, Proc CSCW 2002 ACM Press (2002), 314 -323.
5. Dix, A., Finlay, J., Abowd, G., and Beale, R. Human-Computer Interaction, Third Edition Pearson Education, 2004