ADVANCING THE DESIGN OF TECHNOLOGY-MEDIATED SOCIAL PARTICIPATION SYSTEMS

Ed H. Chi, *Palo Alto Research Center*
Sean Munson, *University of Michigan*
Gerhard Fischer and Sarah Vieweg, *University of Colorado at Boulder*
Cynthia Parr, *University of Maryland*

Technology-mediated social participation systems can dramatically change the way science, government, healthcare, entertainment, and enterprise operate. Using research from several fields, we can learn how to design better TMSP systems and to better integrate lessons learned in practice back into theory.

The emerging cyberinfrastructure is providing enormous unexplored opportunities for individuals, groups, communities, and societies at large. One of the most exciting of these opportunities, with profound implications for years to come, is the use of innovative information and communication technologies that support the public in becoming cocreators of new ideas, knowledge, and products in personally meaningful activities.[1]

The increasing use of social media has led to dramatic changes in the way science, government, healthcare, entertainment, and enterprise operate. Large-scale participation in technology-mediated social participation (TMSP) systems has opened up incredible new opportunities to employ online communities for greater good. The time is ripe to capitalize on this activity and create sociotechnical ecosystems that can harness the wisdom of crowds.[2]

Research on social media has in part emerged from scientific studies aimed at understanding and developing technologies that enhance the intelligence of both individuals and groups.[3] It has also naturally evolved from efforts to improve information seeking and sense making on the Web.[4]

To realize social media's potential for positive change, we must understand how to design for more effective social participation[5] and to employ techniques from theoretical models of social capital and networks.[6,7]

Social media create technological, economic, and organizational disruptions that provide foundations to redefine learning, working, and collaboration. Social systems in turn create new support environments for social creativity and collaborative design. This transformation complements consumer cultures with cultures of participation, in which active contributors who want open environments seek to create solutions by and for themselves.[8] By providing infrastructures, seeds, and mechanisms for user-generated content, the participatory Web 2.0 transcends Web 1.0 environments, which are dominated by broadcast media developed by professional designers for passive users.

0018-9162/10/$26.00 © 2010 IEEE Published by the IEEE Computer Society

VISIONARY SCENARIO

Consider the following scenario.

It is 2014 and Venezuela is struggling to meet the new targets set for reducing biodiversity loss under the Convention on Biological Diversity. Angela is a college student passionate about rainforests. She loves watching documentaries and knows that her country is rich in biological diversity, but figures there is not much she can do about it other than write term papers and send e-mails demanding action. She is a biology major, but as she lives in Caracas, the closest she has come to biological diversity has been photographing flowers in city parks.

Angela accepts an invitation to add one of the many flower images she has posted on Flickr to the growing online Encyclopedia of Life (EOL) images pool (www.eol.org). She and her family are excited to see her photo illustrating text from Wikipedia as well as a highly technical description from botanists on the eFloras project (www.efloras.org). She submits additional images to other EOL pages, and is pleased that EOL curators always rate them highly.

In research, many attempts to design from theory have resulted in naïve and not always effective instantiations.

Angela would like to translate the technical English accompanying some of her photos on EOL into readable Spanish but is not sure how to do that. Also, an EOL curator commented that she misidentified one of her Flickr images. She fixes the error right away but has mixed feelings—are her photos unwelcome? Many of her EOL images remain unreviewed as there are no active curators for those organisms. One of Angela's professors has the expertise to review her photos but does not have time; also, he prefers to contribute only to scientific journals. On the bright side, while the professor does not allow his students to cite Wikipedia in their term papers, he does permit them to cite EOL because it has expert reviewers.

Through a discussion on an EOL page for a native plant she has seen, Angela meets several amateur and professional botanists. The map on EOL for that species shows no points in her area, but her new friends explain that it is rare and often confused with another species. They add this information to the EOL page so others will know, and use EOL to make a list of related species that might be found in a nearby national park. During the next year, Angela's friends make several field trips to take photos and slowly build rich EOL pages for these plants. They help the park naturalist make a field guide using EOL and also help to organize a "bioblitz," a one-day inventory event. Together with scientists, they document new populations of rare plants. Angela's friends mount a citizen science effort to collect data to determine if these plants are at risk of extinction due to climate change. Their findings are freely available and downloadable from EOL pages. This has been the most rewarding set of experiences Angela has had in college, and it did not happen in her classroom.

TMSP SYSTEM DESIGN CHALLENGES

This scenario highlights many of the obstacles that TMSP designers face. The Grand Challenge, then, is to learn how to design better TMSP systems using theories and models from several fields and to determine how to better integrate lessons learned in practice back into theory. At the high level, this includes multiple subchallenges:

- How can we best generate, refine, and evaluate TMSP system design ideas, and who should be involved in this process? How do we choose among conflicting design ideas?
- In TMSP systems, how do we motivate contributions? What incentives are inherently present and should be provided or designed into the system? Controversy surrounding the quality of information in TMSP systems such as Wikipedia raises the question as to whether the so-called wisdom of crowds actually results in mediocre products and ideas.[9] How can we detect and prevent participation that attempts to cheat or game the system, without deterring other participation?
- How can we match tasks to people with appropriate skills and interests? How should we tailor a TMSP system to people with different skills, interests, and needs? How can the system effectively aggregate and integrate information from experts and novices, and from people with different training, perspectives, and vocabulary?
- How do we integrate information provided by people with data gathered from sensors or other behavior traces? What data integration technology do we choose when partners vary in ability, platform, and schemas?

These are no longer entirely open questions; the efforts of researchers and practitioners have substantially advanced both theory and design practice. Technologies that facilitate data access, aggregation, manipulation, and presentation have also broadened the ranks of designers, builders, and administrators of tools and systems to support social participation.

Much more work remains to support TMSP system designers. Many attempts to design from theory have resulted in naïve and not always effective instantiations. On the other hand, theory does not adequately explain or predict many highly effective designs. Our goal is to help

close this gap. Toward that end, we propose studies that will let researchers and practitioners

- explore particular design principles for TMSP systems, focusing specifically on three classes of design issues: the understanding of what domains and contexts might benefit from mass social participation; the exploration of components, tools, and techniques for social participation; and the mapping of factors that affect users and communities.
- understand the overall architecture and model of social participation; determine how the model interacts with social computing techniques such as recommendation engines, visual analytics, and social voting systems; and explore the interaction between users and communities with human and machine intelligence.
- develop best practices for designing in different contexts and to meet various types of challenges, and learn how to exploit the long tail[10] to improve social creativity.

Infrastructure investments that span research and practice communities will be necessary. In particular, we propose a depository for datasets from existing communities. This *data depository* will be responsible for facilitating data contributions, archiving the data, and granting appropriate levels of access to researchers with prereviewed research plans requiring the deposited data. We also propose a research center that experiments on live platforms of TMSP systems—a *living laboratory* that focuses on impacts in real domain areas.[11]

Toward these ends, we seek to engage in efforts that are multidisciplinary—bringing together researchers, practitioners, and citizens from different disciplines; multisector—supporting collaborations among academia, industry, and start-up companies as well as among domains such as open government, social healthcare, and education; and international—understanding and fostering developments that transcend national and societal boundaries.

DESIGNING FOR SOCIAL PARTICIPATION

When we consider the potential benefits of TMSP systems, large-scale participation is a necessity. Designers must synthesize different perspectives, exploit ideas from multiple disciplines, manage large amounts of information relevant to a collaborative task, and understand the decisions that determine a system's long-term evolution.

The principal concern for TMSP system designers is ensuring that participants both give to and get something back from the system. This may take the form of debating ideas, contributing to a domain's knowledge, or sharing experiences. A good example of this is Wikipedia, in which individuals with distinct motivations and perspectives are contributing to a single, encyclopedic knowledge base. With Wikipedia as an example, TMSP systems should facilitate collective action by encouraging users to learn from one another and to form consensus.

Four concepts are at the heart of effective design.

- *Usability* refers to the ability of all users to contribute, regardless of accessibility requirements, computing experience, and costs associated with working with TMSP systems.
- *Sociability* refers to users' skill in networking, making connections, and interacting well with others. Designers can facilitate social interactions among system users through detailed research and careful consideration of offline sociability.

The principal concern for TMSP system designers is ensuring that participants both give to and get something back from the system.

- *Social capital* refers to the positions people occupy in social networks and their ability to utilize those positions to achieve goals. Designers can enable people to comfortably align themselves within social networks in both leadership and follower positions.
- *Collective intelligence* refers to intelligent behavior among groups of people in defining and reaching goals. Designers can create mechanisms such as voting systems, folksonomies, and other opinion aggregators to ensure the emergence of collective intelligence over time.

Designers must strive to reduce the burden on users to acquire the technical knowledge and social skills necessary to participate as well as motivate and reward active contributors. They must also understand the conditions under which humans will take the time and effort to contribute.

CATEGORIES OF DESIGN PROBLEMS

Broadly speaking, TMSP system designers must solve problems that fall into three major categories: knowledgeware, toolware, and peopleware.

Knowledgeware

Knowledgeware refers to the understanding of domains and contexts large-scale social participation could greatly impact. TMSP systems are appearing across various domains, and designers must understand how their choices interact with domains and the different classes of TMSP systems. Consider a simple design claim: design feature x will result in behavior or outcome y. On its own

this claim seems straightforward, but it is also incomplete. Feature *x* might produce outcome *y* in the health domain but an entirely different outcome in the open government domain, or it might hold true for systems in which people must participate but not for opt-in or opt-out systems. In real system design, context and domain matter.

Basic questions designers seek to answer include these: In which domains would massive-scale social participation dramatically disrupt current solutions? For example, how can we radically change textbooks by making them a complete social experience? How do we design TMSP systems for cold start, adoption, growth, and maintenance in multiple domains? Other contexts that designers must consider include community life stage (a choice that works in the start-up phase might not be appropriate for the same community as it matures) and individual differences (a feature that motivates one user might demotivate another user).

The real-world successes and failures of design ideas have driven and will continue to drive theoretical refinements.

Some candidate domains to explore initially include healthcare, education, open government, and enterprise innovation. Other domains worth considering include energy and sustainability, citizen science, and emergency response and preparedness.

Toolware

Toolware refers to the system components that enable effective social participation. The continued advancement of a broad range of techniques and technologies will be necessary to enhance and efficiently apply many human capabilities:

- *The ability to recognize, extract, and identify patterns.* Techniques and technologies include machine learning methods such as clustering and classifiers, social network analysis, expertise finding and matchmaking, social search, and recommendation and voting systems.
- *The ability to augment and aggregate human intelligence by providing powerful analysis tools.* Technologies and techniques include augmentation of human curation or organization of content, workflow and process control and management, decision support systems, prediction markets, and tools for machine and human summarization and distillation in knowledge production.
- *The ability to communicate complex ideas with many people.* This includes the ability to share without creating clutter or overload, collaborative information visualization, and group and personal dashboards
- *Matching tasks with people who are both motivated and appropriately skilled.* How can designers make a system inherently appealing or desirable to use, or how can they design appropriate extrinsic rewards? What should designers do when different types of incentives compete? Microtask markets such as Amazon Mechanical Turk, question-and-answer sites, and peer production systems fit into this space.

Of course, many low-level infrastructure issues, including cloud computing, underlie these techniques and technologies.

Peopleware

Peopleware refers to understanding how people act in social cognitive systems, in which users function as both individual and social agents. Designers need better models of users and communities, and more knowledge of how to build TMSP systems from theory that facilitate the emergence of collective intelligence. Key design questions include the following:

- Beyond opinion aggregation and crowdsourcing, how do social cognition systems work with human intelligence? Here, human intelligence broadly refers to new techniques utilizing human computations such as Amazon Mechanical Turk and games-with-a-purpose techniques.
- How do we design systems that make social interactions between users seamless? Social interactions here broadly include communication behaviors such as commenting, replying, annotation, and distribution and broadcasting of information.
- How do we design for conflict and coordination in TMSP systems? What are the governing policies and design principles, and how are they implemented? How do we represent multiple points of view in a TMSP system? How do we encourage debate and discussion between potentially balkanized groups?
- How do a system's governance, policies, and norms emerge? What behaviors should be designed in code, mandated in policy, or encouraged by norms?

The real-world successes and failures of design ideas have driven and will continue to drive theoretical refinements. Researchers will need listening posts and communication channels that enable learning from these real-world experiences and support the production of new design ideas in an iterative fashion.

For peopleware, understanding users' and groups' goals, and how to support and augment their ability to think and reason, is vital. This requires building better

social participation models and understanding the architecture of participation. By exploring why and how people contribute, we will be able to create TMSP systems that efficiently diffuse information, provide accurate data, are broadly engaging, respond quickly, enable emergent practices, and result in higher productivity.

Integrating domains, techniques, and people

Designers must understand how to integrate knowledgeware, toolware, and peopleware in TMSP systems, and how these aspects interact with each other. The complex interrelationships among domains, techniques, and people make designing for social participation different from designing for other interactive systems. For example, in social healthcare systems such as CureTogether (www.curetogether.com), designers must understand the techniques required to support a patient community as it comes together to solve medical problems. What characteristics of the medical domain dictate the interaction techniques we can employ? Do privacy issues trump openness design principles? Do we have to use different sharing techniques to preserve anonymity? Designers must know what kinds of parameters they can change in a TMSP system that would result in different behaviors.

TMSP INFRASTRUCTURE

To facilitate TMSP system design, we propose infrastructural investments in two research centers, possibly distributed across organizations, that would be part of a broader collaboratory network: a data depository responsible for accepting datasets, cataloging and archiving them, and making them available to qualified researchers; and a living laboratory, where researchers and practitioners can come together to study new and existing systems and communities. Both centers would serve as brokers across communities.

Scientific studies have established that innovations and rewards are more likely to occur in certain niches in a social network. Those entities that span more tightly formed communities are more likely to see opportunities earlier and to reap the benefits of arbitrage. We envision the living laboratory as being at the intersection of academic research (for example, in microeconomics, social networks, network theory, and cognitive modeling), industrial research, and start-up companies.

Data depository

Many important contributions to TMSP studies have come from the careful analysis of datasets from large-scale existing systems. Some datasets are released publicly; researchers assemble others by scraping websites or APIs, and in some cases they negotiate access to company datasets on a case-by-case basis. Researchers can use such data to learn about behavior in real systems with millions of users, and the resulting findings have high external validity.

Companies are not uniformly eager to release this data, and even those that see potential benefits cannot always do so. Netflix recently cancelled the successor to its Netflix Prize over fears that it could not adequately anonymize the data that competitors would use. Furthermore, data that seems anonymous today may be identifiable when paired with other future data or resources and methods for deanonymizing data.

Our proposed data depository would accept datasets from industry and other sources, and grant access to researchers with qualified study plans and appropriate human-subject approvals. Analysis of more sensitive datasets might be further restructured or only permitted on the depository's computing infrastructure. The depository would also organize courses on large-scale data analysis by experts in the field for professors, students, and practitioners.

Designers must know what kinds of parameters they can change in a TMSP system that would result in different behaviors.

Precedents for such an organization, such as the Inter-University Consortium for Political and Social Research (www.icpsr.umich.edu/icpsrweb/ICPSR), which has been archiving and making social science data available for more than 40 years, offer some guidance about how the data depository might be organized and what its budget might be. The depository will be a single site, with start-up funding initially provided by the federal government. Over time, the site might either continue to be federally funded—the Library of Congress's recent acquisition of Twitter data suggests that this is a priority—or be supported by dues from member institutions.

Living laboratory

It is important that TMSP researchers study data on real users in live systems and not work only with "dead data," in which the system is no longer evolving.[8] To fully explore the research issues around social participation, we propose a living laboratory that focuses on designing, building, and growing TMSP systems. This collaborative research center would develop a platform for technology prototypes, live experiments, and scientific research in a wide variety of domains. It would also be a place to exhibit research wares, attract user participation, and voice thought leadership. In addition to delivering Social Web technologies—applications, features, tools, and so on—the living lab would operate in a social fashion.

The living laboratory concept, in which the understanding of social cognitive systems derives from a real, natural, and unconstrained environment, constitutes a research platform on which a live community engaged in real tasks and collective behavior might emerge out of the use of the system's features. This suggests a great need for academia to connect with industry, and vice versa. It also calls for building bridges to large open source initiatives such as Wikipedia and government efforts such as Data.gov.

The living laboratory's main strategic goals include the following:

We envision the living lab as a multidisciplinary, multisector, and internationally collaborative sociotechnical environment.

- *Center of excellence.* The living lab must promote innovation in academia, business, and the public sector. While utilizing traditional methods such as publications and conference presentations, it must also obtain recognition from the Social Web research community via Web 2.0 mechanisms such as blogs, tagging, and wikis. It should establish itself as a center for ideas about TMSP design through its websites.
- *Portfolio of high-value experiments.* The living lab must create a self-sustaining model through a portfolio of experiments with clearly understood solutions to important problems, milestone metrics to evaluate continued investment or abandonment, and desired end goals. This will enable the center to post prototype applications on the Web, get user feedback for rapid iterative improvement, and attract interest from potential sponsors and partners. One mantra of the living lab may be to "fail fast and fail often" to learn what does and does not work.
- *Deep science.* Unlike start-ups and other shorter-term experiments, the living lab will benefit from the ability to do deep science. This might include developing techniques for extracting structures from social data; developing models of interference and coordination costs; understanding expertise, social capital, and social network models; gaining insight into reducing user friction; and evaluating different design methods across time and various communities.

One consequence of the living laboratory model is a focus on impact and not just demos. Whatever gets developed must be usable—systems must benefit individuals as well as improve collective intelligence. This means that evaluations need to be applied often and iteratively throughout the process. A substantial focus of the lab's research must be on engineering the user interfaces and back ends to be reasonably robust and scalable.

We envision the living lab as a multidisciplinary, multisector, and internationally collaborative sociotechnical environment that will create a strong synergy among academia, industry, distributed scientific communities, and the public. It will develop frameworks in specific contexts that are applicable to numerous domains for which social creativity and collaborative design are important.

The living laboratory will support the education of a new generation of scholars motivated to engage in social participation. Members of the living lab community will coordinate with other researchers to understand new theories and models of social media and participation that might give rise to unique analysis methods and algorithms.

Engagement

To support collaborations across TMSP research communities, the living laboratory should be collocated with one or more institutions recognized as leaders in the TMSP field, preferably in regions with a high concentration of private and public sector companies and organizations working in this space and possibly in areas that are experiencing economic hardships and transitions.

Researchers and practitioners will come together in the lab to study existing technical systems as well as develop new ones from the ground up. These new systems will include large, ongoing research platforms—such as the successful MovieLens recommender project (http://movielens.umn.edu)—as well as systems that are built with both research and commercial goals but, unlike traditional start-ups, are designed from the beginning to facilitate research. A fellows program will support practitioners who wish to rotate from industry to the lab to learn how to better design and analyze their own communities.

The living lab must be located in proximity to good partners in industry and academia. There may thus be one or more physical research centers. Different funding models are possible: the centers might be entirely federally funded, they could be funded by a mixture of academia and industry, or there could be full partnerships between different stakeholders.

Researchers can also learn much from real-world communities by spending time embedded in organizations not collocated with the living laboratory that are designing TMSP systems. These roving researchers could include graduate students and academics on sabbatical or other types of leave. A portion of their stipends could be funded under the National Initiative for Technology-Mediated Social Participation, with the requirement that all work resulting from the fellowship be published or otherwise made publicly available.

This is the "decade of social media." We now have a taste of how mass-scale participation in technology-mediated social participation systems can dramatically change the way science, government, healthcare, entertainment, and enterprise operate. We have an unprecedented opportunity to enhance the science, models, techniques, and evaluations of TMSP systems. Our overall recommendation is threefold.

First, we propose exploring unique design principles relating to three different areas: the understanding of what domains and contexts might benefit from mass social participation (knowledgeware); the exploration of components, tools, and techniques for social participation (toolware); and the mapping of factors that affect users, groups, and communities (peopleware).

Second, we recommend developing a suite of methods for designing TMSP systems. We must better understand the overall architecture and model of social participation—including existing users and communities and their goals and aspirations, as well as how they interact with social computing techniques such as recommendation engines, visual analytics, and social voting systems. We must use human and machine intelligence techniques to further explore the interaction between users and communities. Close collaboration between researchers and practitioners during the development and evolution of TMSP systems will make this possible.

Finally, we propose creating both a depository for datasets from real-world sociotechnical systems and a living laboratory to focus research efforts on live platforms of TMSP systems across many contexts. We seek to engage with industry and start-up companies as well as domains such as open government, healthcare, and education.

Acknowledgments

We thank the members of the TMSP workshops who participated in the discussions that shaped this article: John T. Riedl, David Gutelius, Jeff Heer, Hal Eden, Hang Ung, Anupriya Ankolekar, Mark Ackerman, Erwin Gianchandani, Derek Hansen, Wayne Lutters, and Jenny Preece.

References

1. B. Shneiderman, "Web Science: A Provocative Invitation to Computer Science," *Comm. ACM*, June 2007, pp. 25-27.
2. J. Surowiecki, *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*, Doubleday, 2004.
3. V. Bush, "As We May Think," *The Atlantic Monthly*, July 1945, pp. 101-108.
4. E.H. Chi, "The Social Web: Research and Opportunities," *Computer*, Sept. 2008, pp. 88-91.
5. A. Girgensohn and A. Lee. "Making Web Sites Be Places for Social Interaction," *Proc. 2002 ACM Conf. Computer Supported Cooperative Work* (CSCW 02), ACM Press, 2002, pp. 136-145.
6. M.S. Granovetter, "The Strength of Weak Ties," *Am. J. Sociology*, May 1973, pp. 1360-1380.
7. R.S. Burt, *Structural Holes: The Social Structure of Competition*, Harvard Univ. Press, 1995.
8. L. Lessig, *Free Culture: The Nature and Future of Creativity*, Penguin, 2005.
9. J. Lanier, *You Are Not a Gadget: A Manifesto*, Knopf, 2010.
10. C. Anderson, *The Long Tail: Why the Future of Business Is Selling Less of More*, Hyperion, 2006.
11. E.H. Chi, "A Position Paper on 'Living Laboratories': Rethinking Ecological Designs and Experimentation in Human-Computer Interaction," *Human-Computer Interaction: New Trends, Part I*, LNCS 5610, Springer, 2009, pp. 597-605.

Ed H. Chi *is a principal scientist at the Palo Alto Research Center, where he manages the Augmented Social Cognition group. His research interests include user interfaces and human-computer interaction, Social Web/Web 2.0, social computing, Internet traffic and Web ecology analysis, and information visualization. Chi received a PhD in computer and information science from the University of Minnesota. He is a senior member of IEEE and a member of the IEEE Computer Society and the ACM. Contact him at chi@acm.org.*

Sean Munson *is a PhD candidate in the School of Information at the University of Michigan. His research interests include social computing and persuasive technology in the domains of wellness and political diversity. Munson received a BS in engineering from Olin College. Contact him at samunson@umich.edu.*

Gerhard Fischer *is director of the Center for Lifelong Learning and Design, a professor in the Department of Computer Science, and a fellow of the Institute of Cognitive Science, all at the University of Colorado at Boulder. His research interests include lifelong learning, metadesign, social creativity, distributed intelligence, human-computer interaction, and design-for-all (assistive) technologies. Fischer received a PhD in computer science from the University of Hamburg, Germany. He is a member of the CHI Academy and a fellow of the ACM. Contact him at gerhard@colorado.edu.*

Sarah Vieweg *is a PhD student studying technology, media, and society at the University of Colorado at Boulder. Her research interests include human-computer interaction, computer-supported cooperative work, social computing, and natural-language processing. Vieweg received an MA in linguistics from the University of Colorado at Boulder. Contact her at sarah.vieweg@colorado.edu.*

Cynthia Parr *is an adjunct faculty member at the University of Maryland Institute for Advanced Computer Studies as well as director of the Encyclopedia of Life's Species Pages Group. She uses Web technology to help instructors at many institutions teach about animal ecology and evolution. Parr received a PhD in biology from the University of Michigan. Contact her at parrc@si.edu.*