DOI:10.1145/1995376.1995389

Ben Shneiderman, Jennifer Preece, and Peter Pirolli

Viewpoint

Realizing the Value of Social Media Requires Innovative Computing Research

How social media are expanding traditional research and development topics for computer and information scientists.

Social media technologies such as Facebook, Twitter, blogs, wikis, Flickr, and YouTube have garnered more than a billion users. These platforms enable more than friendly chatter and individual expression; they facilitate remarkably diverse and broad participation while accelerating the formation of effective collaborations.

Promising social media projects suggest that dramatic transformations are possible in health care, energy sustainability, environmental conservation, disaster response, and community safety.[14] Some commentators even see social media as a means for economic revitalization through business innovation, educational transformation, and civic revival.[15] However, there are deep challenges in understanding the benefits of social media and ameliorating their dangers. Computer, information, and social scientists, network analysts, system developers, community managers, and many other professionals will have important roles to play as they extend their disciplines with innovative research and development agendas.

The potential for social media impact is illustrated by international upheavals such as the Iranian elections,[9] Wikileaks information releases, and Egyptian democratic movement. In addition, a variety of U.S. and other open government efforts have been launched recently to promote transparency, participation, and collaboration. For example, *data.gov* promotes access to detailed U.S. government agency performance data and *recovery.gov* provides contracting information on the county-by-county use of stimulus money, leading to broader discussion, plus invitations to report fraud, abuse, and waste. Increased participation and collaboration that changes the relationship between government agencies and the general public is beginning with *challenge.gov*, which invites solutions to problems, *serve.gov* to expand volunteering, and wiki-based deliberative Web sites to request commentary on agency directions or regulatory plans.

There are deep challenges in understanding the benefits of social media and ameliorating their dangers.

Social media present dangers too. These include the potential for more polarized discussions as users selectively view only materials aligned with their world view and scientists retreat to narrow research topics ("balkanization") that limit the healthy interchange with related disciplines.[16] Another risk is reduced credibility of online resources as rumors and misinformation spread, unfiltered by traditional journalistic verification. Social media can distract from deep reflection as individuals respond to frequent interruptions and collaborative production methods with free distribution can undermine established reward systems, as journalists have painfully discovered.[6] Breaches of privacy and security are frequently mentioned topics and so is identity theft, online bullying, and disclosure of potentially damaging or embarrassing personal information.

Goals and Challenges for Computing Research

Realizing the full value of social media requires research agendas that include understanding the mechanisms for unleashing chain reactions of human contributions and collaborations while preventing harmful outcomes such as privacy violations, malicious attacks, and misuse by terrorists, oppressive regimes, and criminals. Evo-

lutionary patterns of activity within homogeneous or heterogeneous small, medium, and large organizations could be studied with network analysis tools to identify highly productive individuals or groups.[5,8] Understanding the dynamics of collective action, governance, and leadership in networked organizations can present grand scientific challenges that are worthy of Nobel Prize recognition, such as bestowed on Elinor Ostrom.[11] However, early successes such as Wikipedia and health discussion groups generate the impression that success in using social media is inevitable, but the reality is that failure is the norm and even successful projects have problems. For Wikipedia, only one out of every 1,000 readers registers to make contributions—and even fewer participate in durable collaborations. Higher rates of participation are needed for smaller projects to succeed.

One model of how participation evolves is the Reader-to-Leader Framework (see Figure 1), which also offers usability and sociability design guidelines.[13] This framework describes how some of the large numbers of readers mature into contributors who offer user-generated content such as videos, photos, reviews, and ratings. A smaller segment becomes intensely involved in collaborative groups who discuss substantive changes and expansions of content. Finally, a small group of leaders emerge to set policies, deal with attacks, resolve disputes, and mentor newcomers. A major research effort could validate and refine such frameworks, providing deeper insights into the nature of human motivation in different contexts.

Not every computing scientist will be interested in studying social media, but computing science social media research can have a profound impact on every discipline.

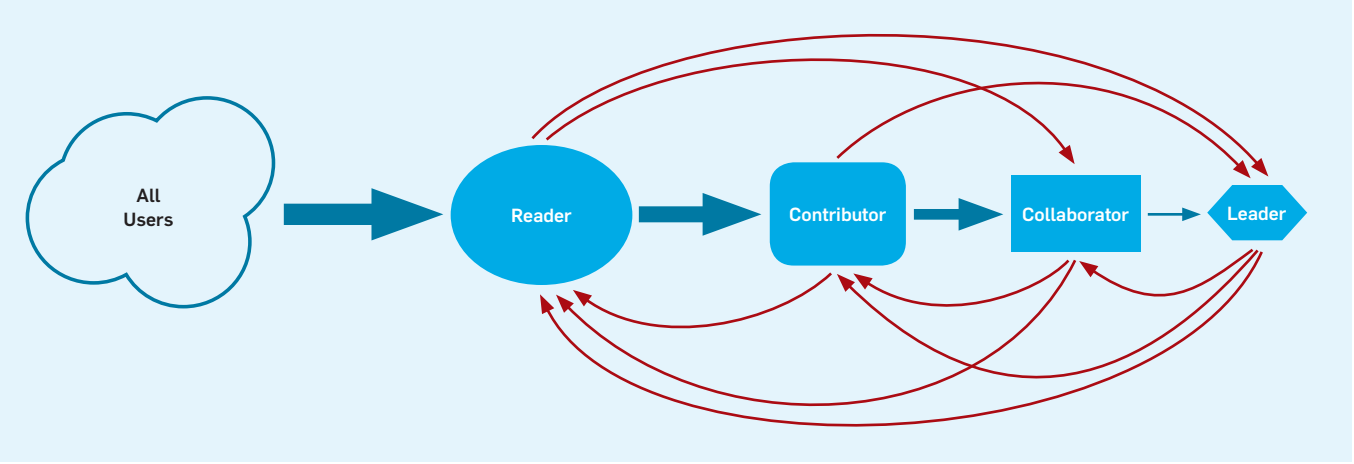
The emerging science of online motivation draws on sociological studies and political science theories, as well as on statistical methods, agent-based simulations, linguistic sentiment analysis, and network analysis/visualization.[4] For example, studying trust, in its many forms, would lead to improved designs that facilitate collaboration so that participants can rapidly resolve their differences and act effectively when needed, as some environmental groups did following the Gulf Oil spill.[3]

Another research topic is the growing availability of *big social data*,[8] which presents significant challenges to algorithm designers and mathematicians possibly requiring innovative chip designs to accelerate the necessary computations. Just as graphical processing units (GPUs) have enabled rapid 3D exploration, *social processing units* (*SPUs*) may be needed to enable scalable social network analysis for computations such as eigenvector centrality, community clustering, and comprehensible layouts. While Moore's Law has signaled the steady progress of hardware technologies in petaflops and gigahertz, new laws could describe the growth of massive projects by measuring peta-contribs and giga-collabs.

New scientific measures are also needed for trust, empathy, responsibility, and privacy, and new mathematical operators could characterize the relationship among relevant usability and sociability measures. The rich contextual and volatile temporal dependencies among these measures mean traditional reductionist models need to be enriched with inter-variable sensitivity analysis and informed by qualitative studies. The motivations for early Wikipedia users may be very different from the community safety organizers who must develop trust and ensure privacy over many years. Similarly, those engaged in collective intelligence projects may respond to very different motivations from those who conduct collective action initiatives. Weak ties are sufficient for early stages and for spreading ideas, but strong ties also become vital for the deep commitments necessary to produce substantial change.

Multidisciplinary network science is rapidly emerging with models of network growth/decay, strategies for comparing thousands of apparently similar networks, and algorithms for detecting unusual bursts of activity.[1,2,4] These methods, strategies, and algorithms

Figure 1. The Reader-to-Leader Framework suggests the evolutionary path for participants in social media communities. Some users may move smoothly through the four phases, while others may take different paths as indicated by the arrows in the figure.



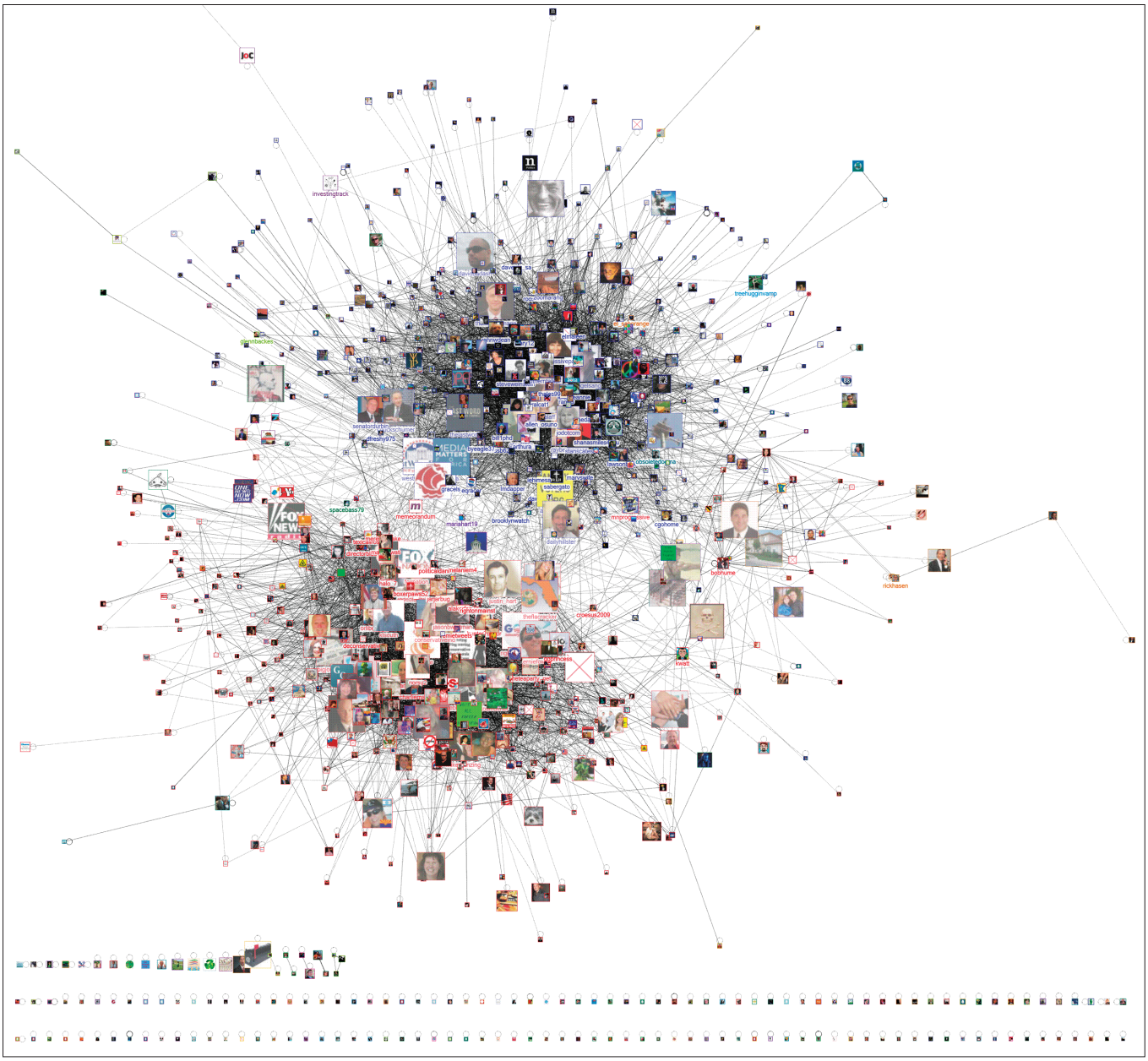


Figure 2. Connections among Twitter users who recently mentioned GOP when queried on July 6, 2011, with vertices scaled by numbers of followers. The clusters are created by the patterns of connections (follows, replies, and mentions) among the authors in the graph. The clusters were based on Clauset-Newman-Moore algorithmic analysis in which the red cluster is composed of largely GOP supporters, while the blue cluster contains largely critics and opponents of the GOP as indicated by the content of the tweets from each cluster. Name labels are attached to the top 10% of vertices as measured by betweenness centrality. Other colored or shaped nodes are not strongly affiliated with either major cluster. Users on the bottom are not connected with any of the other Twitter users.

VISUALIZATION BY MARC A. SMITH USING NODEXL; HTTP://WWW.CODEPLEX.COM/NODEXL.

will benefit from coupling with natural language processing and discourse analysis to identify nexuses of positive collaborations as well as threatening activity from hate groups, terrorists, and criminals (see Figure 2).

Still more ambitious research goals are to identify key influencers, successful discussion generators, and reliable answer providers in discussion groups with millions of participants while curbing the damage caused by scammers, spammers, and troublemakers of many kinds who seek to undermine the efficacy of social media platforms (see http://www.wikitrust.net).

While many of these topics will be new to computer and information scientists, the social media will dramatically expand their traditional research and development topics such as large-scale heterogeneous distributed systems design, exploratory search tasks across enormous multimedia databases, and visual analytic tools with statistical components that produce valuable insights even from voluminous and noisy data. Other traditional challenges that will become even more central include context-aware systems that work on mobile, laptop, Web, and cloud-based platforms, and policy-aware systems that allow successful operation in different cultures, languages, and political systems.

Broad Scholarly Payoffs

Not every computing scientist will be interested in studying social media, but computing science social media research can have a profound impact on every discipline. Social media are

already restructuring the ways in which scholars form collaborations and communicate their results.[10] What used to be called the *invisible college* of personal scholarly communications is now a vast and highly visible, searchable, and influential infrastructure. These new scholarly social networks, the *visible commons*, ignite hot topics, accelerate data sharing, and enable rapid refinements to theories in ways that were never before possible. For example, in August 2010, when a researcher claimed to have proven one of the most profound, challenging, and elusive problems in all of mathematics and computer science (P=NP?), blogs (such as http://rjlipton.wordpress.com), wikis, and other forms of online communication conveyed active discussion about the proof—and ultimately enabled a form of real-time "peer review" that called into question the researcher's approach.

Scientists also have begun to use social media to conduct new forms of scientific research. NASA's use of clickworkers to measure Martian craters (http://beamartian.jpl.nasa.gov) or the Encyclopedia of Life's (http://eol.org) integration of professional scientists with trained citizen scientists and nature enthusiasts are examples of even more potent methods. Scientists can now engage with thousands of peers as in the GeneWiki (http://genwiki.eva.mpg.de), with serious amateurs as in star surveys (http://galaxyzoo.org), or with numerous paid workers through services such as Mechanical Turk (http://mturk.com). Such large-scale collaborations could produce conflict over credit for breakthroughs unless new strategies for supporting trust are created.[10,15] Other ethical dilemmas come from the appropriateness of existing Institutional Review Board oversight processes or fairness of using low-paid Web-based labor in place of traditional research assistants or experimental participants.

Call to Action

These topics provoked lively discussions at two National Science Foundation (NSF)-funded workshops held in the past year. The final report[12] covers descriptive, explanatory, prescriptive, and predictive theories; opportunities in health care/wellness and e-government; ethical issues for researchers; design strategies for practitioners; motivational challenges for community managers; research infrastructure proposals; and innovative educational reforms (http://www.tmsp.umd.edu). Some steps in expanding research have already begun with the NSF's Social Computational Systems program (http://www.nsf.gov/pubs/2010/nsf10600/nsf10600.htm) and the National Institutes of Health's two programs on Social Network Analysis and Health (http://obssr.od.nih.gov/funding_opportunities/foas/faqs.aspx).

The next step will be paradigm-shifting methods for conducting scholarly research in the computing sciences and in every discipline.

Researchers from many disciplines can build on the ideas generated at these workshops and summarized here by working with funding agencies to restructure existing programs so that social media research becomes more widely supported. Evaluations of civic social media projects could make them more reliably successful by developing validated design guidelines, effective community management strategies, advanced visual analytic and statistical tools, and broader theories. Academics can spread this new knowledge by introducing segments on social media into existing courses, adding new courses, and planning degree programs for professionals and researchers.

Adventurous researchers are already using social media to improve or speed their research, but the next step will be paradigm-shifting methods for conducting scholarly research in the computing sciences and in every discipline. Faster paths to curing cancer, tracking climate change, mapping species distribution, and much more seem within reach. However, there is also a risk that social media researchers will soon confront ethical challenges as serious as those that the nuclear physicists faced in the 1950s. This time the concerns will be about inequities in Internet access, violations of privacy, vulnerability to attacks, as well as technical failures and social chaos during crises. We believe the computing sciences community can rise to these challenges and find effective solutions. C

References

1. Barabasi, A.-L., *Bursts: The Hidden Pattern Behind Everything We Do*. Dutton, NY, 2010.
2. Easley, D. and Kleinberg, J. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*, Cambridge University Press, NY, 2010.
3. Golbeck, J. Weaving a web of trust. *Science, 321*, 5896 (2008), 1640–1641.
4. Hansen, D., Shneiderman, B., and Smith, M.A. *Analyzing Social Media Networks with NodeXL: Insights from a Connected World*, Morgan Kaufmann Publishers, San Francisco, CA, 2011.
5. Hendler, J. et al. Web Science: An interdisciplinary approach to understanding the Web. *Commun. ACM 51*, 7 (July 2008), 62–69.
6. Lanier, J. *You Are Not a Gadget: A Manifesto*, Knopf Publishers, NY, 2010.
7. Latour, B. and Woolgar, S. *Laboratory Life: The Construction of Scientific Facts*. Princeton University Press, Princeton, NJ, 1986.
8. Lazer, D., et al. Computational social science. *Science 323* (Feb. 6, 2009), 721–723.
9. Lichtenstein, J. Digital diplomacy. *New York Times Magazine* (July 18, 2010), 24–29.
10. Olson, G.M., Zimmerman, A., and Bos, N., Eds., *Scientific Collaboration on the Internet*, MIT Press, Cambridge, MA, 2008.
11. Ostrom, E. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press, NY, 1990.
12. Pirolli, P., Preece, J., and Shneiderman, B., Eds., Technology-mediated social participation (cover feature with seven articles). *IEEE Computer 43*, 11 (Nov. 2010), 20–67.
13. Preece, J. and Shneiderman, B. The Reader-to-Leader Framework: Motivating technology-mediated social participation. *AIS Transactions on Human-Computer Interaction 1*, 1 (Mar. 2009), 13–32; http://aisel.aisnet.org/thci/vol1/iss1/5/
14. Shirky, C. *Cognitive Surplus: Creativity and Generosity in a Connected Age*. Penguin Press, NY, 2010.
15. Tapscott, D. and Williams, A.D. *MacroWikinomics: Rebooting Business and the World*. Portfolio, NY, 2010.
16. Van Alstyne, M. and Brynjolfsson, E. Global village or cyber-balkans? Modeling and measuring the integration of electronic communities. *Management Science 51*, 6 (June 2005), 851–868.

Ben Shneiderman (ben@cs.umd.edu) is a professor in the Department of Computer Science, the founding director of the Human-Computer Interaction Laboratory, and a member of the Institute for Advanced Computer Studies at the University of Maryland at College Park.

Jennifer Preece (preece@umd.edu) is a professor and dean of the Information School at the University of Maryland at College Park.

Peter Pirolli (pirolli@parc.com) is a research fellow in the Augmented Social Cognition Area at the Palo Alto Research Center (PARC).

We appreciate National Science Foundation support (IIS-0956571) to conduct the two workshops and all the participants in those workshops. We appreciate the comments we received from the reviewers and James Hendler.