Getting Connected: Social Science in the Age of Networks

A proposal submitted to Cornell's Institute for the Social Sciences

Michael W. Macy, Project Leader
Professor, of Sociology

Team Members
- **David Easley, Department of Economics**
- **Geri Gay, Department of Communication**
- **Dan Huttenlocher, Department of Computer Science**
- **Jon Kleinberg, Department of Computer Science**

Table of Contents:

I. GOALS AND MOTIVATION

The proposed theme project would bring together leading figures in the study of social and information networks, in a series of workshops, seminars, and lectures designed to foster multi-disciplinary exchanges that identify new research questions and approaches and that broaden and deepen substantive and methodological knowledge. The proposed activities focus on specific problem areas, with the goal of generating follow-on research projects and proposals.

Section I motivates our proposal in two steps. We begin by sketching the backdrop of growing collaborations between social and information scientists at Cornell and elsewhere, and their intellectual and institutional significance (**I.A**). We then focus more narrowly on a specific research area where these collaborations have exceptional promise, especially at Cornell – social and information networks (**I.B**). Section II then itemizes the nuts and bolts of our proposal, including the proposed activities (**II.A,B**), personnel (**II.C**), and expected outcomes (**II.D,E**).

I.A. The Big Picture

The impetus for this proposal is a growing appreciation that social and information scientists have much to learn from one another, and that these synergies have the potential to attract broad interest both within academia and beyond. We believe that the time is ripe for initiatives that bridge the social and information sciences. First, national interest in computational social science is increasing, both at peer institutions and among funding agencies. For example, both the National Science Foundation (NSF) and the Advanced Research and Development Activity (ARDA) have recently announced funding opportunities in computational social science, particularly in social and information networks. Second, Cornell is in a position to become widely recognized as a leading institution in this emerging area. We have broad expertise, several well-known individual researchers, and a history of collaboration among those involved in this proposal.

This proposal grows out of a six-year collaboration between the Departments of Sociology, Economics, Communication, Theoretical and Applied Mechanics, and Computer Science, dating back to the beginning of the program on "dynamic and complex networks" which was a core component of the IGERT program in complex systems directed by Steve Strogatz. The common denominator of that collaboration has been the "new science of networks," which Strogatz helped to pioneer, along with his student Duncan Watts, now a member of the Sociology Department at Columbia University.

In the years since the IGERT was initially funded, research at Cornell and elsewhere has created a community of scholars with similar interests but whose research lies in different disciplines. While the complex systems IGERT program has been enormously successful in many ways, it also points out some of the difficulties of doing multi-disciplinary work in social and information networks. The IGERT funded graduate students worked with one or two faculty, and as a result most of the research projects were still quite focused within the disciplines. In short, despite our separate strengths in the study of social and information networks, the whole is a good deal less than the sum of its parts, largely because our research efforts are in separate departments and colleges without systematic coordination and collaboration.

This proposal is motivated by the need to realize the full potential of the intellectual talent that is scattered around various units on campus, by supplementing on-going discipline-specific research projects with a series of intensive topical workshops, seminars, and lectures that bring together leaders at Cornell and elsewhere with similar interests but with different disciplinary backgrounds and methodologies. If successful, we believe that this project could establish Cornell as a flagship in the emerging field of computational social science, leading to increased visibility and new grants for multi-disciplinary and multi-institutional collaborations (which are required by several of the programs at NSF, such as "Human Social Dynamics" and "Mathematical Social and Behavioral Sciences").

The main focus of the proposed activity is the workshop series. Each week-long workshop will have approximately 25 presenters from within and outside Cornell. Each workshop will have a specific theme, and will be organized much like the successful week-long workshops of the NSF funded Mathematical Sciences Research Institute (www.msri.org).

The tangible product of these activities will be workshop materials – presentations, notes, and video – made available on the Web. However, the more important products will be intellectual and institutional:

- new research collaborations that grow out of the workshop series,
- research proposals to support those collaborations,
- the emergence of new norms and expectations that legitimate and sustain efforts to integrate social and information science, and
- the visibility of Cornell as a leading center for computational social science, centered on the study of social and information networks.

I.B. The Immediate Focus

In the past decade, online information networks have fundamentally transformed the ways that people obtain news and current events, communicate with one another, seek advice, shop, travel, and are entertained. One of the primary distinguishing aspects of these networks is that they are produced collectively by the actions and interactions of large numbers of individual actors. For instance, today's blogs, discussion boards, online auctions, and even the links between personal Web pages all arise out of the interactions of multiple individuals whose decisions are largely unconstrained by institutional rules and uncoordinated by procedures for centralized governance.

The rapid rise in the popularity and scope of online information networks has attracted considerable attention from scientists in a broad range of disciplines, including sociology, economics, communication, computer science, physics and applied mathematics. For example, techniques from the mathematical and computational

sciences have begun to reveal general properties of the structure of the Web. Moreover, online networks provide a highly accessible source of data for studying interactions within social networks, something that has traditionally been difficult to observe.

At the same time, techniques from the social sciences have been applied to online information networks. Methodologically, social scientists have developed highly sophisticated tools for network analysis that information scientists often lack. Substantively, social scientists have developed, refined, and tested theoretical models that address fundamental problems in information networks, such as the problems of trust and peer-based social control. The creation of trust among strangers affects peoples' daily lives as they engage in online activities such as bidding in auctions, seeking medical advice, and making purchases based on recommendations. An on-going series of international conferences introduced computer scientists to research by economists and sociologists that showed how trust can be conferred via purely online mechanisms such as seller ratings on eBay. (These bi-annual conferences, which began four years ago at Harvard and MIT, are resuming this spring in Amsterdam, and will hopefully move to Cornell in the second year of our theme project, as part of our proposed workshop series.)

Despite these shared interests, however, social scientists investigating these questions are often not well equipped to understand the technical options for conveying trust online, while computer and information scientists are not adequately trained in research on social networks, signaling, or informal social control. For instance, computer scientists tend to analyze social networks as online systems, missing decades of advanced work in social network analysis by sociologists. Meanwhile, social scientists are often ill equipped to formalize the graph-theoretic and computational intuitions that motivate their work. Even in areas with considerable cross-disciplinary collaboration, such as theoretical work in economics and computer science on mechanism design, there is relatively little connection to the empirical social sciences.

Some of the most influential work in social and information networks has occurred at the boundary of related disciplines. For instance, economists have

traditionally developed formal game-theoretic models that assume relatively homogeneous fully connected populations of identical actors. Recent advances at the boundaries of economics, sociology, and computer science incorporate complex networks into game theory by embedding games inside relational structures that constrain opportunities for interaction. Meanwhile other studies use game theory to model a dynamic network structure as a Nash equilibrium, in which no one has an incentive to search for a new partner. Yet the broader questions of how to unite empirical studies of social phenomena, network games, and computationally tractable models under a single conceptual framework remain largely unexamined.

Growing awareness of the importance of these cross-disciplinary collaborations is reflected in a recent grant to Cornell from the National Science Foundation's program on "Human Social Dynamics." The grant proposal recasts a compelling problem that has preoccupied generations of social thinkers, from Thomas Hobbes and David Hume to Mancur Olson and James Coleman. The classical problem of cooperation among egoists has been framed around the tension between individual and collective rationality. The proposal shifts attention to the configuration of social ties through which participation in collective action can propagate as a cascade. This approach bridges the disciplinary gap between the study of dynamical systems and game-theoretic models of collective action. The latter have generally relied on static conceptions of equilibrium among homogenous, fully connected populations of utility maximizers with complete and perfect information. The project is using computational methods developed for modeling cascades on large networks to explore how the structure of social ties can affect the spread of participation in collective action in heterogeneous populations of boundedly rational actors.

A core component of the project is an international workshop designed to promote the exchange of theoretical ideas, skills, and methodologies between social, natural, and information scientists working on similar problems but with minimal mutual awareness or appreciation of potential synergies and no common language for communicating ideas or results. (A list of participants is included in section **III-C**.) The

workshop will take place during the first year of our theme project, and will provide the springboard to launch the workshop series that will take place during the second year.

In sum, there is growing recognition, by NSF and by the scientists they support, of the need for collaborations that extend beyond the social sciences. This is reflected in many small steps in this direction, such as the IGERT and HSD grants to Cornell. In pointing to this progress, we are concerned that readers may conclude that the problem we have identified is solving itself, and that the activities we propose will take place anyway, even without support from the ISS. Nothing could be further from the reality we face. These emerging collaborations provide ISS with a unique and remarkable opportunity to leverage modest resources to nudge strategically placed dominoes that have the potential to set into motion far-reaching changes in the configuration of the social sciences at Cornell.

II. PROJECT DESIGN

II.A. Structure of the Project Activities across the Three Years

The first year is a planning year in which the project team will meet on at least a monthly basis to develop plans for the specific workshop topics and other activities in the second year. As part of the planning year, four scholars will be invited to Cornell to help us refine our theme project, determine workshop topics, and aid in selecting two eminent scholars to each spend a sabbatical semester at Cornell during the second year. All six visitors (four in the first year and two in the second) will give public lectures and seminars as part of their visits. The project leader will also give a public lecture during the first year, to introduce the project to the Cornell community. Towards the end of the first year, we will interview applicants for the two post-doc positions.

During the second year of the project we will hold a series of six intensive workshops to be attended by scholars from outside Cornell, as well as Cornell faculty, researchers, and students recruited from a broad range of disciplines and research fields. Topics and distinguished participants will be selected by the project team during

the first year, with additional participants selected by the visiting scholars and post-docs.

The workshop format we propose is designed to be both broad (not the same group every time) and deep (the week-long format allows for intense exploration of a problem.) As to breadth, each workshop will have about 25 to 30 participants, with about 16-20 from Cornell, the ten team members plus local specialists whose fit with a particular workshop topic is primarily substantive, rather than methodological. For example, in a workshop on peer networks, we would invite people studying deviant adolescent behavior (drugs, gangs, teen pregnancy, school dropouts), group dynamics, life course, Weblogs, etc., even though they might have little or no experience with the uses of network analysis to study these questions. These are the people who are likely to benefit most. At the same time, computer and information scientists will benefit from exposure to novel technical challenges posed by unfamiliar applications.

Most of the teaching and outreach activities will also take place during the second year. One of the afternoon sessions of each workshop will be replaced by a public lecture by one of the external workshop participants, in a different venue and on a topic that is broadly engaging for non-specialists and undergraduates. We plan to offer at least one specialized seminar-type course each semester of the second year (at either the graduate or undergraduate level). Again the specific topics of these courses will be decided during the first year, but we believe that several of the proposed workshop topics are also excellent candidates for reading seminars and courses. These courses might be associated with IGERT seminars where appropriate.

II.B Sample Questions to be Addressed in the Seminar Series

In this section we present some potential workshop topics in order to illustrate the kinds of workshops that we envision as part of the theme project, though the final selection of topics would be made during the first year by the full project team. For each topic we have put in parentheses the names of some of the Cornell faculty who we believe would be natural participants in each workshop, though these lists are by no

means complete. In addition, we have highlighted the members of the core group, to illustrate the complementarities and synergies in the skills and interests that each member brings to the team.

1. <u>Diffusion of innovations</u>. This project goes beyond the study of the Web itself, to the underlying social processes that are reflected in its pages. With funding from the National Science Foundation, Cornell's Faculty of Computing and Information Science (CIS) and the Cornell Theory Center (CTC) are building the "Cornell Web Lab" (CWL). This new computing facility will give researchers at Cornell local and unrestricted access to the Internet Archive (www.archive.org), which has been capturing and storing snapshots of the Web every two months since 1996. These snapshots show the contents of every page on the Web, including the links between pages. When CWL comes on line next year, social scientists will be able to study the diffusion (as well as decline) of innovations, rumors, opinions, fads, urban legends, social movements – anything that might be found on a Web page – in ways we have never done before, given the difficulty collecting worldwide serial data on adoption (and abandonment) at the micro level.

Yet few social scientists even know this gigantic evolving archive exists, much less how to use it. Because expertise in the use of these new tools is currently limited to computer scientists, the full potential for social science has barely begun to be tapped. Meanwhile, the computer and information scientists who are building and beginning to use this powerful new tool are not entirely sure what to do with it, as they have explicitly noted. Social scientists who specialize in diffusion and social influence can bring to bear a better appreciation of the key theoretical questions that have resisted conventional research methods, such as the effects of early adopters on the legitimacy of late adoption, the effects of spatial and small world network structures, and the relationship between polarization of opinion (positions become extreme) and polarization of "camps" (positions become correlated across issues). Social scientists are also better able to identify the full range of possible applications for these new data and to frame new research results against the backdrop of what is already known from past

research. Do sites embedded in dense clusters adopt innovations more slowly, compared to those with fewer or more random links? Are early adopters more or less likely to abandon early, and does this depend on the type of innovation (technical, political, fashion, etc.)? Do abandonments diffuse in patterns similar to those observed for adoptions? Have Weblogs attenuated political and social moderation in favor of more extreme positions on issues? Have they also promoted polarization of opinion (defined as the tendency for positions on substantively unrelated issues to become correlated over time)? Do links form between hostile pages, and if so, does this in turn intensify or attenuate the hostility over time? Are innovations "homogenized" as diffusion progresses, or do later adopters become bolder as the innovation gains legitimacy?

To pursue these and related questions, we will recruit a team of specialists on the diffusion of innovation, from inside and outside Cornell, and use one or more of the week-long workshops to train participants in the use of CWL and to identify the most urgent hypotheses that CWL data are ideally suited to address. This research project will give an expanding group of social scientists access to powerful computational tools that few of us even know exist, much less how to use. At the same time it will guide computer and information scientists in building systems that are better suited to answering questions posed by social scientists.

2. Hidden role of networks in empirical studies of peer influence. Research on adolescent behavior has established the powerful effects of peer influence for a wide range of individual outcomes, such as achievement in school, committing a crime or becoming pregnant. Ethnographic studies confirm, for instance, how peers support the sexual decisions of both males and females that lead to teen pregnancy (Kaplan 1997). Yet these effects are typically very difficult to detect in population-level data on social interactions. Scholars' responses have been to look for problems in the statistical methods for estimating social effects. We are instead interested in investigating an alternative hypothesis that has been largely overlooked in the social science literature: The paradox of peer effects that are reported at the individual level but that are missing

at the aggregate level may be indicative of social network structure. Although peers are influential in individual decision-making, as seen in the ethnographic studies, the network of peer interactions may prevent peer effects from propagating. A suggestive example is how the predictions of standard epidemiological models in fully connected populations are modified when network structure is taken into account. Topology matters, and its statistical representation may determine the differential local and global peer effects. The larger research issue is to understand how the mechanisms of peer interaction and the topology of peer relations interact to determine aggregate social outcomes.

      3. <u>Dynamic networks and individual personality, opinions, and beliefs.</u> Traditional social network analytic studies tend to take network structure as a given and disregard how attributes and characteristics of agents within a network impact an evolving network structure as well as each agent's emergent position within those structures. At the same time, research on individual differences (e.g., personality psychology or opinion research) tends to ignore the larger social network within which an agent is situated. However, research in Cornell's HCI lab has begun to uncover important effects on an agent's location in a network (e.g., central or peripheral, brokerage positions, etc.) that flow from the specific characteristics and attributes of the agent. Other studies have shown how homophily and influence (the tendency for likes to interact and for interactants to become more alike) generates tendencies toward polarization and differentiation in dynamic networks.

II.C Personnel

II.C.1 The Project Team

    We have assembled a core group of network specialists that span multiple disciplines and colleges, both within and beyond the social sciences. Despite its diversity, the group is highly cohesive and complementary. We share a common interest in network analysis, yet we are each well connected into very different scholarly niches. The section that follows describes each of our backgrounds and the

specific skills that each brings to the project, reflected as well in the associations with various workshop topics listed above.

David Easley is a Professor of Economics at Cornell University and holds the Henry Scarborough Chair of Social Science. He received his bachelor's degree from the University of Kentucky and his Ph.D. from Northwestern University. His research interests are in learning and evolution in markets, decision theory and market microstructure. He is the Director of the Solomon Interdisciplinary Social Science Program at Cornell University. He has been Chair of the Economics Department and Chair of the Social Science Advisory Council at Cornell. He has been a Fellow of the Econometric Society since 1997. He is the author of over 40 published papers and has been the recipient of 11 National Science Foundation grants.

Geri Gay is Professor and Chair of Communication. She is also director of the Human Computer Interaction (HCI) Lab at Cornell. She received her bachelor's degree from the University of Maine and her Ph.D. from Cornell. Professor Gay's research interests are in social networks, HCI, information visualization, computer mediated communication, and design. She brings to this project special expertise and interests in issues of social mapping, social translucence, and social psychological aspects of networking. She recently co-authored Activity Centered Design (MIT Press, 2004). She has won several teaching awards including the Chancellor's Award for Excellence in Teaching in 2001 and was named a Stephen H. Weiss Fellow in 2004. Professor Gay has received funding for her research and design projects from NSF, Intel, Microsoft, the Mellon Foundation, GE Foundation, AT& T Foundation, IBM, Getty, Apple and several private donors. She teaches courses in human computer interaction and the social design of communication systems. Geri will be the project coordinator in year 3.

Daniel Huttenlocher is a Professor of Computer Science and of Management at Cornell University, and holds the John P. and Rilla Neafsey Chair of Computing, Information Science and Business. He received his bachelor's degree from University of Michigan and his masters and doctorate from MIT. Prof. Huttenlocher's research interests are in computer vision, geometric algorithms, interactive document systems,

financial trading technology, and IT strategy. He has 24 U.S. patents, has published more than 75 technical papers, and has been recognized on several occasions for his teaching and research, including being named an NSF Presidential Young Investigator in 1990, the New York State Professor of the Year in 1993, and a Stephen H. Weiss Fellow in 1996. In addition to his academic work, Huttenlocher has served as CTO of Intelligent Markets and was on the senior management team at Xerox PARC.

Jon Kleinberg received his A.B. from Cornell in 1993 and his Ph.D. in computer science from MIT in 1996. He subsequently spent a year as a Visiting Scientist at the IBM Almaden Research Center, and is now an Associate Professor in the Department of Computer Science at Cornell University. His research interests are centered around algorithmic issues at the interface of networks and information, with an emphasis on the network structures that underpin the Web and other on-line media. Given Kleinberg's interests in the interplay between information resources and the networks -- both human and technological -- that shape and are shaped by them, his work will fit closely with the proposal's focus on the interaction between social and information networks. In particular, through his current research and through collaboration with the other team members, he hopes to expose deeper connections between emerging on-line media and the social systems that are mediated by them. He is the recipient of an NSF Career Award, an ONR Young Investigator Award, an Alfred P. Sloan Foundation Fellowship, a David and Lucile Packard Foundation Fellowship, best-paper awards from the ACM conferences on data mining and on database theory, teaching awards from the Cornell Engineering College and Computer Science Department, and the 2001 National Academy of Sciences Award for Initiatives in Research, given annually to one U.S. scientist under the age of 35.

Michael Macy is Professor and Chair of Sociology. He has a BA and PhD from Harvard University and an MA from Stanford. Prof. Macy's research uses computational models and laboratory experiments with human subjects to explore how threshold effects in local interactions might generate familiar but enigmatic social phenomena, such as the emergence and collapse of fads, collective actions, and most

recently, the adoration of "naked emperors." He brings to the proposed theme project skills and research interests in computational modeling of dynamic and complex networks. In addition to numerous contributions to encyclopedias and edited volumes, he has published thirty peer-reviewed publications, including seven papers in the flagship sociology journals, three papers in the *Annual Review of Sociology*, and two articles in the *Proceedings of the National Academy of Sciences*. He was also a fellow at the Netherlands Institute for Advanced Study and recipient of two best-paper awards, from the American Sociological Association and the Academy of Management. He is a recipient of seven awards from the National Science Foundation.

II.C.2 Cornell Affiliates

Cornell faculty whose research interests intersect with the study of social and information networks span the social and information sciences at Cornell:

NAMES DELETED FROM PUBLIC PROPOSAL

II.C.3 Visiting Scholars

Below we list prominent scholars who might be invited to come to Cornell as speakers, visiting scholars, and workshop participants. This group not only spans multiple disciplines but also multiple countries. Those names with an asterisk have already agreed to participate in the HSD workshop in October, 2005 that will launch the proposed three-year series of meetings.

US:
- Robert Axelrod, Pol Science, Michigan
- Rob Axtell, Economics, Brookings*
- Peter Bearman, Sociology, Columbia
- Chris Dellarocas, Business, U. Maryland
- Joshua Epstein, Economics, Brookings*
- Noah Friedkin, Sociology, UCSB
- Mark Granovetter, Sociology, Stanford
- Matt Jackson, Economics, Cal Tech*
- David Lazer, Political Science, Harvard
- Mark Newman, Physics, Michigan*
- Paul Resnick, Info Science, Michigan
- Brian Skyrms, Phil of Science, UC Irvine*
- Duncan Watts, Sociology, Columbia
- Peyton Young, Economics, Brookings, Hopkins, Oxford*
- Chuck Manski, Economics, Northwestern
- Steven Durlauf, Ecomomics, isconsin
- Gary Solon, Economics, Michigan
- Chris Udry, Economics, Yale
- Ed Lazer, Economics, Stanford

- Peter Bearman, Sociology, Columbia
- Noah Friedkin, Sociology, UCSB

Europe:
- Marcel van Assen, Sociology, University*
- Vincent Buskens, Sociology, Utrecht*
- Victor Eguiluz, Physics, Mediterranean Institute for Advanced Studies*

- Andreas Flache, Sociology, Groningen*
- Peter Hedstrom, Sociology, Oxford*
- Arno Riedl, Economics, Amsterdam*
- Maxi San Miguel, Physics, Mediterranean Institute for Advanced Studies*
- Stephane Rosenkranz, Econ, Utrecht*
- Chris Snijders, Sociology, Utrecht
- Fernando Vega Redondo, Econ, Alicante*

II.D Outreach to the Broader Cornell Community and Beyond

Scholars from other institutions will be invited to Cornell to help the Cornell team select, design, develop, and refine our theme projects, workshops and other products from this project. In addition, four visitors and the two resident scholars from outside Cornell will give public lectures and seminars as part of their visits. These public events are intended to promote and coalesce interest around computational social science themes and topics not only at Cornell but also at other institutions.

We expect to run six intensive workshops, as detailed in section **II.A**. Scholars from outside Cornell, as well as Cornell social science faculty, researchers, and students, from a broad range of disciplines and research areas, will participate in the workshops and seminars. Multimedia "whitepapers" will be developed and posted on our project website.

Officers and researchers from foundations, government agencies, and businesses will be invited to attend sessions. It is our intent to partner with other institutions and funding agencies.

II.E Outcomes

The outputs of each workshop will be several-fold. First, a web site with all the presentation materials, notes and videos will be created for each workshop, in order to publicize the presentations and discussions. Second, each workshop will have two co-organizers whose responsibility will be to identify possible multi-disciplinary research collaborations that could yield proposals to funding agencies (to be developed during the third year of the theme project). These co-organizers will play a coordinating role, but may or may not actually participate in the submission of proposals, depending on

the specific topic areas. Furthermore, we expect research teams to self-organize out of workshop discussions, with each team then figuring out their own schedules for on-going research collaborations.

During the third year we will develop and submit proposals to funding agencies such as NSF (e.g., HSD, DST, and/or MSBS) and ARDA. We expect that one of these proposals will be a center-scale project that covers a broad range of topics in social and information networks. The remainder will be smaller projects with fewer co-investigators. We may also offer a "topics course" in the third year, in which several faculty members give lectures describing their research in a particular focus area.

In sum, we propose to take the emerging cross-college partnership at Cornell to the next level with intensive workshops, and with the new Cornell Web Lab as an essential research tool. Research on topics ranging from adolescent behavior to the diffusion of innovations would clearly benefit from the infusion of skills, tools, methods, theories, and findings that are widely dispersed across disciplines and colleges. Our primary goal, however, goes beyond this. By bringing together social and information scientists inside and outside Cornell, across many departments and colleges, with shared interests and complementary skills, we aim to build a culture for collaborative research that will have a lasting impact on social science at Cornell.

REFERENCES

Elaine B. Kaplan. *Not Our Kind of Girl: Unraveling the Myths of Black Teenage Motherhood.* University of California Press, Berkeley CA, 1997.

L. M. Sander, C. P. Warren, I. Sokolov, C. Simon, and J. Koopman. "Percolation on disordered networks as a model for epidemics." *Math. Biosci.*, 180:293–305, 2002.

R. Axelrod. *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration.* Princeton, NJ: Princeton University Press, 1997.

D. Easley and M. O'Hara. "Asset pricing and market microstructure." In M. Harris G. Constantinides and R. Stulz, editors, *Handbook of the Economics of Finance.* Elsevier, 2003.

T. Joachims. *Learning to Classify Text using Support Vector Machines*. Kluwer, 2002.

B. Pang, L. Lee, and S. Vaithyanathan. "Thumbs up? Sentiment classification using machine learning techniques." In *Proceedings of EMNLP-2002*, pages 79–86, 2002.

APPENDIX 1. BUDGET AND JUSTIFICATION

DELETED FROM PUBLIC PROPOSAL

APPENDIX 2. ABBREVIATED CV'S OF TEAM MEMBERS

Biographical Sketch: David Easley

Professional Preparation

University of Kentucky, Economics, B.A., 1974

Northwestern University, Economics, M.A., 1975

Northwestern University, Economics, Ph.D., 1979

Appointments:

Henry Scarborough Professor of Social Science, Cornell University, 1996-

Professor of Economics, Cornell University, 1989-

Overseas Fellow, Churchill College, Cambridge University, 1993-94

Visiting Associate Professor, California Institute of Technology, 1985-86

Assistant Professor, Cornell University, 1979-84

Instructor, Northwestern University, 1978-79.

Publications Most Closely Related to the Proposed Project:

1. "If You're So Smart, Why Aren't You Rich? Belief Selection in Complete and Incomplete Markets," with L. Blume, forthcoming, *Econometrica*.

2. "Asset Pricing and Market Microstructure," with M. O'Hara, in *Handbook of the Economics of Finance*, Elsevier Scientific Press, 2003.

3. "Optimality and Natural Selection in Markets," with L. Blume, *Journal of Economic Theory*, Vol. 107, No. 1, 2002.

4. "Is Information Risk a Determinant of Asset Returns?," with S. Hvidkjaer and M. O'Hara, *Journal of Finance*, Vol. 57, No. 5, 2002.

5. "Price, Trade Size and Information in Securities Markets," with M. O'Hara, *Journal of Financial Economics*, Vol. 19, 1987.

Other Significant Publications:

1. "Choice Without Beliefs," with A. Rustichini, *Econometrica*, Vol. 67, No. 5, 1999.

2. "One Day in the Life of a Common Stock," with N. Kiefer and M. O'Hara, *Review of Financial Studies*, Vol. 10, No. 3, 1997.

3. "Evolution and Market Behavior," with L.Blume, *Journal of Economic Theory*, Vol. 58, No. 1, 1992.

4. "Implementation of Rational Expectations Equilibrium," with L. Blume, *Journal of Economic Theory*, Vol. 51, No. 1, 1990.

5. "Controlling a Stochastic Process with Unknown Parameters," with N. Kiefer, *Econometrica*, Vol. 56, No. 5, 1988.

Synergistic Activities

Larry Blume, Joe Halpern and I teach a course that we created on Decision Theory. This course is cross-listed between Computer Science and Economics.

Larry Blume and I taught a mini-course on learning and evolution in markets at the Stockholm School of Economics in May, 2000 to an international group of Ph.D. students and faculty.

I gave a series of lectures on market microstructure and asset pricing at the Jerusalem Summer School in Economic Theory in June, 2000 to a large group of Ph.D. students and faculty.

Larry Blume and I are writing a book on the topic of "Learning and Evolution in Markets" with economics and finance Ph.D. students and faculty as the target audience.

Collaborators

Lawrence Blume (Cornell University), Tarek Coury (Oxford University), Robert Engle (New York University), Joseph Halpern (Cornell University), Soeren Hvidjkaer (Cornell University), Maureen O'Hara (Cornell University), Aldo Rustichini (Tilburg University), Liuren Wu (Baruch College).

Graduate Advisors:

John Ledyard (Cal Tech), Dale Mortensen (Northwestern University), Stanley Reiter (Northwestern University)

Ph.D. Students

Tarek Coury, Pablo Beker, Parimal Bag, Sanjeev Goyal, Kaushik Mitra.

GERALDINE K. GAY

Department of Communication
339 Kennedy Hall
Ithaca, NY 14853
p) 607.255.7737
w) www.comm.cornell.edu
e) gkg1@cornell.edu

Human Computer Interaction Group
301 College Avenue
Ithaca, NY 14850
p) 607.255.5530
w) www.hci.cornell.edu
e) gkg1@cornell.edu

EDUCATION

University of Maine	English & Biology	B.A., 1972
Cornell University	Communication	M.P.S., 1980
Cornell University	Education	Ph.D., 1985

PROFESSIONAL HISTORY

Chair, Communication, Cornell University   (March 2004 – present)

Director, Information Science Ph.D. program, Cornell University   (2003 – 2004)

Professor with Tenure, Communication/Information Science, Cornell University (2001 – present)
Responsible for designing and researching the use of computer mediated communication environments. Research focuses on cognitive an social issues for the design and use of interactive communication and mobile technologies. Past research has explored navigation issues, knowledge management, mental models and metaphors, knowledge representations, collaborative work and learning, and system design. Funding sources for research includes the National Science Foundation, the Mellon Foundation, the Intel Corporation, Microsoft, the AT&T Foundation, IBM, the Getty Foundation, NASA, and private donors. Field Membership: Cognitive Studies and Educational Psychology.

Director, Human Computer Interaction Group, Cornell University   (2001 – present)

Associate Professor, Communication/Information Science, Cornell University   (1994 – 1999)
Responsible for developing and conducting a research, teaching and design program in human computer interaction technologies, including digital library projects and computer mediated communication, cognition and social design issues. Recent sponsors include the GE Foundation, IBM, the Mellon Foundation, the National Science Foundation, DARPA, the National Endowment for Humanities, and private donors.
Field Membership: Cognitive Studies and Educational Psychology

Assistant Professor, Communication, Cornell University   (1988-1994)
Responsible for developing and conducting a research, teaching and design program in interactive multimedia technologies, cognition and social design issues. Research sponsors included the Mellon Foundation, Apple Computers, IBM, and the National Science Foundation.

SABBATICALS

July – December, 2002
Scotland; England; Hillsboro, OR; & Washington, DC
Intel Wireless Computing Study

July – December, 1995
Tokyo, Japan (IBM)
Global Digital Museum Project

RECENT AWARDS & GRANTS

Co-Principal Investigator, AT&T, NASA, and NYS ($4,500,000). *An Advanced Interactive Discovery Environment for Engineering Education*. (2001-2006). Award: #3534168.

NSF Subcontract ($300,000). *NSDL Training Digital Library Interface*. (2003-2006).

NSF Co-Principal Investigator ($98,000). Collaborative Project - *Core Integration: Leading NSDL Toward Long Term Success*. (2003-2006).

NSF Subcontract ($57,000). *Evaluation, Usability, and Learning Using K-MODDL Digital Library*. (2003-2004).

NSF Co-Principal Investigator with Elizabeth Liddy, Syracuse University ($374,000). *MetaTest: Evaluating the Quality and Utility of MetaData*. (2002-2004).

Intel-CIMI ($225,000). *Wireless Computing in Museums*. (2001-2004).

Principal Investigator, anonymous gift ($347,000). *Wireless Computing in Education*. (2000-2006).

Co-Principal Investigator with William Arms, Cornell University & the Intel Corporation ($300,000). *Evaluating Student Collaboration and Team-Based Development in a Nomadic Computing Environment*. (1999-2003).

NSF Co-Principal Investigator with Carl Lagoze, Cornell University ($2,000,000). Digital Library II. (1999-2002).

PUBLICATIONS

2003-2004
Gay, G., and Hembrooke, H. (2004). *Activity-Centered Design: An Ecological Approach to Designing Smart Tools and Usable Systems.* Cambridge, MA: MIT Press.

Boehner, K., Thom-Santelli, J., Zoss, A., Gay, G., Barrett, T., and Hall, J. (2005). Imprints of Place: Creative Expressions of the Museum Experience. Submitted to CHI 2005.

Walther, J., Gay, G., and Hancock, J. (2005). How do communication and technology researchers study the Internet. *Journal of Communication*.

Cho, H.C., Gay, G., Davidson, B.D., and Ingraffea, A., (in press). Communication Styles, Social Networks, and Learning Perfomance in a CSCL. *Journal of Computers and Education*.

Cho, H., Lee, J., Stefanone, M., and Gay, G. (in press). Development of Computer-Supported Collaborative Social Networks in an Online Learning Community. *Behavior and Information Technology*.

Boehner, K., Sengers, P., and Gay, G. (in press). Affective presence in museums: Ambient systems for creative expression. *Journal of Digital Creativity*.

Boehner, K., Gay, G., and Larkin, C. (2005). Drawing Evaluation into Design for Mobile Computing: A Case Study of the Renwick Gallery's Handheld Education Project. *Journal of Digital Libraries, Special Issue.*

Hembrooke, H., Granka, L., Gay, G.K., and Liddy, E. (in press) The Effects of Expertise and Feedback on Search Term Selection and Subsequent Learning. *Journal of the American Society for Information Science and Technology.*

Lee, J. S., Cho, H., and Gay, G. (2004) Applying Network Analysis to the Analysis of Web Traffic. In the *Proceedings of Work with Computing Systems* (pp. 634-639). Kuala Lumpur: Damai Sciences.

Stefanone, M., Hancock, J., Gay, G., & Ingraffea, T. Emergent networks, locus of control, and the pursuit of social capital. *Proceedings of the 2004 Computer Supported Work Conference*, Chicago.

Pan, B., Gay, G., Saylor, J.M., Hembrooke, H. A., Henderson, D. (2004). Usability, learning and subjective experience: User evaluation of K-MODDL in an undergraduate class. In H. Chen, M. Christel, & E. Lim (Ed.), *Proceedings of the Fourth ACM/IEEE Joint Conference on Digitial Libraries* (JCDL '04) (pp. 188-189). New York: ACM.

Granka, L., Joachims, T., & Gay, G. (2004). Eye-Tracking Analysis of User Behavior in WWW Search. In *Proceedings of 28th Annual ACM Conference on Research and Development in Information Retrieval* (SIGIR '04), Sheffield, UK.

Pan, B., Hembrooke, H., Gay, G., Granka, G., Feusner, M., and Newman, J. (2004). The determinants of web page viewing behavior: An eye tracking study. In S.N. Spencer (Ed.), *Proceedings of Eye Tracking Research & Applications*, New York: ACM SIGGRAPH.

Sengers, P., Kaye, J. J., Boehner, K., Fairbank, J., Gay, G., Medynskiy, E., Wyche, S. (2004). Culturally embedded computing. *IEEE Pervasive Computing,* Special Issue on Art, Design and Entertainment in Pervasive Environments, 3, (1), 14-22.

Hembrooke, H., Gay, G., (2003). The lecture and the Laptop: Multitasking in wireless learning environments. *Journal of Computing in Higher Education*, 15(1), 46-65.

Liddy, L., Gay, G., Allen, E., Hembrooke, H., Finneran, T., Granka, L. (2003). MetaTest: Evaluation of metadata from generation to use. *Proceedings of the Joint Conference on Digital Libraries.*

Daniel Huttenlocher

a. Professional Preparation

Undergraduate: University of Michigan, double major in Computer Science and Experimental Psychology, B.Sc., 1980.
Graduate: Massachusetts Institute of Technology, M.Sc., 1984, Ph.D, 1988.

b. Appointments

John and Rilla Neafsey Professor of Computing, Information Science and Business, Cornell University, 2002-present.

Chief Technology Officer and Founder, Intelligent Markets, 2000-present.

Office of the Center Manager, Xerox Palo Alto Research Center (PARC), 1999.

Associate Professor of Computer Science, Cornell University, 1994-1999.

Principal Scientist, Xerox Palo Alto Research Center (PARC), 1995-1999.

Manager, Image Understanding Area, Xerox Palo Alto Research Center (PARC), 1992-1995.

Member of Research Staff, Xerox Palo Alto Research Center (PARC), 1988-1992.

Assistant Professor of Computer Science, Cornell University, 1988-1993.

c. (i) Publications most related to this research

J. Aizen, D.P. Huttenlocher, J. Kleinberg and T. Novak Traffic-Based Feedback on the Web, Proc. National Academy of Sciences, 6 January 2004.

P. F. Felzenszwalb, D.P. Huttenlocher and J. Kleinberg, Fast Algorithms for Large State Space HMM's with Applications to Web Usage Analysis, Advances in Neural Information Processing Systems (NIPS) 16, 2003.

P.F. Felzenszwalb and D.P. Huttenlocher, Efficient Matching of Pictorial Structures, Proceedings of the IEEE Computer Vision and Pattern Recognition Conference, pp. 66-73, 2000.

G. Gay, A. Sturgill, W. Martin and D.P. Huttenlocher, Document-centered peer collaborations: An exploration of educational uses of networked communication technologies, Journal of Computer Mediated Communication, vol. 4, no. 3, 1999.

J.R. Davis and D.P Huttenlocher, Shared Annotation for Cooperative Learning, Proceedings of the ACM Conference on Computer Supported Cooperative Learning, pp. 84-88, 1995.

c. (ii) Other selected publications

P.F. Felzenszwalb and D.P. Huttenlocher, Efficient Belief Propagation for Early Vision, Proceedings of the IEEE Computer Vision and Pattern Recognition Conference, vol. I, pp 261-268, 2004.

L.P. Chew, K. Kedem, D. Huttenlocher and J. Kleinberg Fast Detection of Common Geometric Substructure in Proteins, *Journal of Computational Biology*, Vol 6, No. 3, pp. 313-325, 1999.

D. Huttenlocher, G.A. Klanderman and W.J. Rucklidge, Comparing Images Using the Hausdorff Distance, IEEE Trans. on Pattern Analysis and Machine Intelligence, 1993.

W.E.L. Grimson and D. Huttenlocher, On the Verification of Hypothesized Matches in Model-Based Recognition, IEEE Trans. on Pattern Analysis and Machine Intelligence, 1991.

D. Huttenlocher and S. Ullman, Recognizing Solid Objects by Alignment with an Image, Intl. Journal of Computer Vision, vol. 5, no. 2, pp. 195-212, 1990.

d. Synergistic Activities

Chaired the committee that led to the establishment of the new Information Science Program at Cornell and serve on the undergraduate curriculum committee for that program.   These activities span information systems (primarily Computer Science), human-centered systems (Psychology, Cognition, Communications), and social systems (economics, sociology and business).

Developed a new course, starting in 1997, "Creating Interactive Web Documents". This introductory course, with no pre-requisites in mathematics or computer science, attracts a broad range of students. Students are introduced to programming and to quantitative reasoning skills, in the context of creating interactive Web pages on topics of interest to them and using their own designs. The course is specifically targeted at students who do not generally take "quantitatively oriented" classes.

Received several awards for outstanding teaching and advising, including New York State Professor of the Year (CASE) and Cornell's top honor the Weiss Presidential Fellowship.

e. (i) Collaborators

Jon Aizen (Internet Archive), Danny Bobrow (Xerox), John Seely Brown (consultant), Pedro Felzenszwalb (MIT), Eric Friedman (Cornell), Geri Gay (Cornell), Joe Halpern (Cornell), Mike Hopcroft (Microsoft), Klara Kedem (Ben Gurion), Jon Kleinberg (Cornell), Ryan Lilien (Dartmouth), Michael Macy (Cornell), Wendy Martin (Cornell), Tony Novak (Cornell), Clark Olson (Univ. Washington), William Rucklidge (Intelligent Markets), Phoebe Segners (Cornell), Amanda Sturgill (Baylor), Eva Tardos (Cornell), Phil Torr (Oxford Brookes), Peter Wayner (consultant), Ramin Zabih (Cornell).

e. (ii) Graduate advisor: Eric Grimson (MIT), Tomas Lozano-Perez (MIT), Shimon Ullman (Weizmann), Victor Zue (MIT).

e. (iii) Thesis advisees and post-doctoral researchers sponsored: Yuri Boykov (W. Ontario), Nick Howe (Smith), Anurag Mittal (Siemens).

Jon M. Kleinberg

Department of Computer Science
Cornell University
Ithaca NY 14853
Phone: (607) 255-3600; FAX: (607) 255-4428
http://www.cs.cornell.edu/home/kleinber/

Education

Massachusetts Institute of Technology.
   Ph. D., May 1996.
       Advisor: Michel Goemans
   S.M., May 1994.
Cornell University.
   A.B., May 1993.

Professional Experience

Cornell University.
   Assistant professor in Department of Computer Science, 1996-2001.
   Associate professor in Department of Computer Science, 2001-present.
IBM Almaden Research Center, San Jose CA. 1996–1997.
   Visiting scientist in CS Principles and Methodologies Group.
   Currently member of Visiting Faculty Program.

Selected Awards and Honors

National Academy of Sciences Award for Initiatives in Research, 2001.
David and Lucile Packard Foundation Fellowship, 1999-2004.
Alfred P. Sloan Fellowship, 1997-1999.
Office of Naval Research Young Investigator Award, 1999-2002.
NSF Career Award, 1997-2001.
Fiona Ip Li '78 and Donald Li '75 Cornell Engineering College Teaching Award, 2000.
IBM Outstanding Innovation Award, 2002.
Best research paper award, ACM Conf. Knowledge Discovery and Data Mining, 2003.
Best paper award, ACM Symposium on Principles of Database Systems, 2000.
Machtey Award, IEEE Symposium on Foundations of Computer Science, 1996.
MIT EECS George M. Sprowls Ph.D. Dissertation Prize, 1996.

Selected Publications

- J. Aizen, D. Huttenlocher, J. Kleinberg, A. Novak. Traffic-Based Feedback on the Web. Proc. Natl. Acad. Sci. USA, 6 January 2004.

- D. Kempe, J. Kleinberg, E. Tardos. Maximizing the Spread of Influence through a

Social Network. Proc. 9th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining, 2003.

- J. Kleinberg. Bursty and Hierarchical Structure in Streams. Data Mining and Knowledge Discovery 7(2003), 373.

- P. Felzenszwalb, D. Huttenlocher, J. Kleinberg. Fast Algorithms for Large-State-Space HMMs with Applications to Web Usage Analysis. Advances in Neural Information Processing Systems (NIPS) 16, 2003.

- J. Kleinberg, É. Tardos. Approximation Algorithms for Classification Problems with Pairwise Relationships: Metric Labeling and Markov Random Fields. Journal of the ACM 49:5(2002).

- J. Kleinberg. An Impossibility Theorem for Clustering. Advances in Neural Information Processing Systems (NIPS) 15, 2002.

- J. Kleinberg, S. Lawrence. 2001. The Structure of the Web. Science 294(2001), 1849.

- J. Kleinberg. Navigation in a Small World. Nature 406(2000), 845.

- S. Chakrabarti, B. Dom, D. Gibson, J. Kleinberg, S.R. Kumar, P. Raghavan, S. Rajagopalan, A. Tomkins. Mining the Web's Link Structure. IEEE Computer 32(8): 60-67 (1999)

- J. Kleinberg. Authoritative Sources in a Hyperlinked Environment. Journal of the ACM, 46:5(1999).

Current and former students

Elliot Anshelevich (Cornell), Debra Goldberg (Harvard Medical School),
David Kempe (USC), Amit Kumar (IIT Dehli), Mark Sandler (Cornell),
Alex Slivkins (Cornell).

Michael Walton Macy   (607) 255-3820 (voice)
Department of Sociology   (206) 666-6021 (fax)
Cornell University   email: mwm14@cornell.edu
Ithaca, NY 14853   URL: www.people.cornell.edu/pages/mwm14

Education:

1980-85: Ph.D. Sociology, Harvard University
1971-72: MA Education, Stanford University
1966-70: BA Government, Harvard University

Academic Experience:

2002- : Professor and Chair, Department of Sociology, Cornell University.
1997-02: Professor, Department of Sociology, Cornell University.
1992-96: Associate Professor, Department of Sociology, Brandeis University.
1993-94: Visiting Associate Professor of Social Studies, Harvard University.
1986-91: Assistant Professor, Department of Sociology, Brandeis University.
1985-86: Lecturer in Social Studies, Harvard University.
1981-85: Teaching Fellow in Social Studies and the Department of Sociology, Harvard University.

University Service (since 2000):

Department Chair, Sociology, Cornell University, 2002-
Einaudi Center Director Search Committee, Cornell University, 2002-03
Faculty Senate, Cornell University, 2000-02
    Dean's Advisory Committee on Promotions, Cornell University, 2000-03
Multidisciplinary Research Program on Organizations and Institutions, Cornell, 1999-2001.
Director of Graduate Studies, Field of Sociology, Cornell University 1997-2000

Recent and Related Publications:

Centola, Damon and **M. Macy. "**Social Life *In Silico*: the Science of Artificial Societies." *Handbook of Group Research and Practice*, edited by Susan Wheelan. Sage Publications.

Centola, D., R. Willer, and **M. Macy**. 2004. "The Emperor's Dilemma: A Computational Model of Self-Enforcing Norms." Forthcoming, *American Journal of Sociology* (projected for spring 2005).

Simpson, B. and **M. Macy**. 2004. "Coalitions in Network Exchange." *Social Forces*, July, 2004.

Macy, M., J. Kitts, A. Flache, and S. Benard. 2003. "Polarization in Dynamic Networks: A Hopfield Model of Emergent Structure." In R. Breiger and K. Carley, eds. *Dynamic Social Network Modeling and Analysis*. National Academy Press.

Vasi, I and **M. Macy**. 2003. "The Mobilizer's Dilemma." *Social Forces*. 81: 983-1002.

Flache, A. and **M. Macy**. 2002. "Stochastic Collusion and the Power Law of Learning." *Journal of Conflict Resolution*. 46:629-653.

Macy, M. and R. Willer. 2002. "From Factors to Actors: Computational Sociology and Agent-Based Modeling." *Annual Review of Sociology*, 28:143-66.

Macy, M. and A. Flache. 2002. "Learning Dynamics in Social Dilemmas." *Proceedings of the National Academy of Sciences*, 99: 7229-36.

Macy, M. and Y. Sato. 2002. "Trust, Cooperation, and Market Formation in the U.S. and Japan." *Proceedings of the National Academy of Sciences*, 99: 7214-20.

Strang, D. and **M. Macy**. 2001. "'In Search of Excellence': Fads, Success Stories, and Adaptive Emulation." 2001. *American Journal of Sociology*, 107: 147-82.

Recent Grants, Awards, and Recognition:

2004-05: Principal Investigator, National Science Foundation Program on Human and Social Dynamics. "Network Topology and the Dynamics of Collective Action."

2004-05: Principal Investigator, National Science Foundation Dissertation Improvement Fellowship for Robert Willer, for laboratory research to test a status theory of collective action.

2003-04: Principal Investigator, National Science Foundation two-year grant to study popular enforcement of unpopular norms ("The Emperor's Dilemma"), using agent-based models and laboratory experiments with human subjects.

2002-03: Research Fellowship, Netherlands Institute for Advanced Study, Wassenaar, NL, in conjunction with a Fellowship from the Fetzer Institute (supplement to NIAS fellowship).

2000-01: Principal Investigator, National Science Foundation one-year grant for web-based laboratory to study trust and cooperation in the U.S. and Japan (with Toshio Yamagishi).

1999: Academy of Management OMT Best Paper Award for "'In Search of Excellence': Fads, Success Stories, and Communication Bias" (with David Strang).

Courses Taught since 2000:

Graduate:

Identity and Interest in Collective Action
Artificial Societies: Computational Models of Social Order and Change

Undergraduate:

Group Solidarity
Introduction to Sociology