

# Ph.D. General Examinations Proposal

*Topics in Social Information*

---

Cameron Marlow  
Electronic Publishing Group  
MIT Media Laboratory  
May 31, 2002 (last updated September 17, 2002)

## **Introduction:**

Internet communication technologies, such as email, instant messaging, and weblogs, have reached threshold levels, making computer mediated interactions inextricably entangled with our lives. Instead of physically conversing or telephoning with neighbors, coworkers, friends and family, we are increasingly finding Internet-based protocols as our sole form of communication.

Our personal exchanges and public announcements have become digitally manifested, and as such are persistently archived and potentially available to a global audience. In this environment, computer analysis is beginning to play an intermediary role between unknowing parties, linking people who have no prior contact.

This general proposal has been arranged around the general topics related to social information, and the function technology can play in understanding these data and relating people to each other. The main area, *identity* addresses the core question of how people cognitively represent one another, and the ways in which our individual differences are portrayed and perceived. The structural theories of *social networks* provide a contextual backdrop for interpreting how individuals can be tied together into a cohesive society. The technical supporting area of *textual information processing* is an overview of the different techniques and methods used by computers to extract and organize meaningful representations from written natural language.

## **Outline of Areas:**

*Identity*

### **Main Area**

Judith Donath, Assistant Professor, MIT Media Laboratory

*Social Networks*

### **Contextual Area**

Keith N. Hampton, Assistant Professor, MIT Department of Urban Studies and Planning

*Textual Information Processing*

### **Technical Area**

Walter Bender, Senior Research Scientist, MIT Media Laboratory

# Main Area

## *Identity*

---

**Examiner:** Judith S. Donath, Assistant Professor, MIT Media Laboratory

**Description:** In any social exchange we encounter, knowing the identity of other individuals is essential to the interpretation and understanding of the interaction. The main area collects readings from a number of disciplines around the topic of the role identity plays in our lives. The first area, *cognitive representations*, is a brief introduction to cognitive aspects of classification and category construction necessary to understand the other areas. *Perception of self and others* concentrates on the process by which we perceive and classify ourselves and others, while the flip side of the coin, how we expose our identity through cultural exchange, is covered in *presentation of identity*. The final area addresses the affordances and constraints of identities involved in computer mediated interactions.

**Requirements:** Assessment in this area will be in the form of a publishable-quality paper.

**Signature:** ..... **Date:** .....

## **Reading List:**

### *Theory*

1. H. Arendt. 1959. *The human condition*. Garden City, N. Y.,: Doubleday.
2. E. Aronson, T. D. Wilson and R. M. Akert. 1994. *Social psychology : the heart and the mind*. New York, NY: HarperCollins College Publishers.
3. P. L. Berger and T. Luckmann. 1966. *The social construction of reality; a treatise in the sociology of knowledge*. Garden City, N.Y.,: Doubleday.
4. C. Geertz. 1973. *Thick description: Toward an interpretive theory of culture*. Pp. ix, 470 in *The interpretation of cultures; selected essays*. New York,: Basic Books.
5. E. Goffman. 1956. *The presentation of self in everyday life*. Edinburgh,: University of Edinburgh Social Sciences Research Centre.
6. E. Goffman. 1961. *Encounters; two studies in the sociology of interaction*. Indianapolis,: Bobbs-Merrill.
7. W. James. 1950. *The principles of psychology*. New York: Dover Publications.
8. E. Prelinger. 1959. *Extension and Structure of the Self*. *Journal of Psychology* 47:13-23.
9. G. Simmel. 1971. *On individuality and social forms; selected writings*. Chicago: University of Chicago Press. Models of Identity

### *Aggregative Models of Identity*

10. R. P. Abelson. 1986. *Beliefs are like possessions*. *Journal for the Theory of Social Behavior* 16:223-250.
11. R. W. Belk. 1988. *Possessions and the Extended Self*. *The Journal of Consumer Research* 15:139-168.

12. M. Csikszentmihalyi. 1993. *Why we need things*. Pp. xvii, 300 in *History from things : essays on material culture*, edited by S. D. Lubar and W. D. Kingery. Washington: Smithsonian Institution Press.
13. M. Csikszentmihalyi and E. Rochberg-Halton. 1981. *The meaning of things : domestic symbols and the self*. Cambridge [Eng.] ; New York: Cambridge University Press.
14. F. Davis. 1992. *Fashion, culture, and identity*. Chicago: University of Chicago Press.
15. S. D. Feldman. 1979. *Nested Identities*. *Studies in Symbolic Interaction* 2:399-418.
16. K. Gelder and S. Thornton. 1997. *Subcultures Reader*. London ; New York: Routledge.
17. G. D. McCracken. 1988. *Culture and consumption : new approaches to the symbolic character of consumer goods and activities*. Bloomington: Indiana University Press.
18. S. Thornton. 1996. *Club cultures: Music, media and subcultural capital*. Hanover, NH: Wesleyan University Press. H. Arendt. 1959. *The human condition*. Garden City, N. Y.,: Doubleday.

### *Identity online*

19. P. Curtis. 1992. *Mudding: Social phenomena in text-based virtual realities*. Proceedings of Directions and Implications of Advanced Computing (DIAC '92). Berkeley, CA.
20. J. Donath. 1998. *Identity and Deception in the Virtual Community*. in *Communities in Cyberspace*, edited by M. Smith and P. Kollock. London: Routledge.
21. J. Donath. 2000. *Being Real*. edited by K. Goldberg. Cambridge, Mass.: MIT Press.
22. S. Herring. 1999. Bringing familiar baggage to the new frontier: Gender differences in computer-mediated communication. Pp. 144-154 in *CyberReader*. Boston: Allyn and Bacon.
23. J. Hollan and S. Stornetta. 1992. *Beyond being there*. Proceedings of ACM Conference on Human Factors in Computing Systems (CHI '92). Monterey, CA.
24. D. Jacobson. 1999. Impression formation in cyberspace: Online expectations and offline experiences in text-based virtual communities. *Journal of Computer Mediated Communication* 5.
25. B. Miller. 1994. *Vital signs of identity*. *IEEE Spectrum* 83:22-30.
26. S. Turkle. 1995. *Life on the screen : identity in the age of the Internet*. New York: Simon & Schuster.

# Contextual Area

## *Social Networks and Information*

---

**Examiner:** Keith N. Hampton, Assistant Professor, MIT Department of Urban Studies and Planning

**Description:** Social network analysis is a growing subfield of sociology dedicated to the viewpoint that the structure of social interactions plays a fundamental role in defining society. This contextual area is meant to provide a basis for understanding how localized interactions give rise to large-scale social phenomena, extending the identity of individuals to the identity of a culture or society. The first section, *background and theory*, is a collection of fundamental and defining papers in the area, while the second section, *methodology* focuses more slightly on the sampling and analysis that underpins this theory. The third and fourth areas focus the subject on information and communication: *search and diffusion* considers empirical studies of networks and information flow, while *technology and social networks* looks at how communication technologies are changing our social networks, and by extension, our society.

**Requirements:** This area will be examined by a 24-hour take home examination administered by Professor Hampton.

**Signature:** ..... **Date:** .....

### Reading List:

#### *Background and Theory*

1. Bott, Elizabeth. 1972. *Family and Social Network*. London: Tavistock.
2. Burt, Ronald S. 1993. "The Social Structure of Competition." in *Explorations in Economic Sociology*, edited by Richard Swedberg. New York: Sage.
3. Feld, Scott. 1982. "Social structure determinants of similarity among associates." *American Sociological Review* 47:797-801.
4. Fischer, Claude. 1982. *To dwell among friends*. Berkeley: University of California Press.
5. Granovetter, Mark S. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78:1360-1380.
6. —. 1982. "The strength of weak ties: A network theory revisited." in *Social Structure and Network Analysis*, edited by Peter Marsden and Nan Lin. Beverly Hills: Sage.
7. Jacobs, Jane. 1992. *The death and life of great American cities*. New York: Vintage Books.
8. Marsden, Peter. 1987. "Core discussion networks of Americans." *American Sociological Review* 52:122-131.
9. McPherson, Smith-Lovin, and Cook. 2001. "Birds of a feather: Homophily in Social Networks." *Annual Review of Sociology* 27:415-44.
10. Milgram, Stanley. 1967. "The small-world problem." *Psychology Today* 1:62-67.
11. Portes, Alejandro, and Patricia Landolt. 1996. "The downside of social capital." *The American Prospect* 26:18-21.

12. Putnam, Robert D. 2000. *Bowling alone : the collapse and revival of American community*. New York: Simon & Schuster.
13. Tilly, Charles. 1988. "Misreading, then rereading, nineteenth-century social change." Pp. 332-358 in *Social Structures: A Network Approach*, edited by Barry Wellman and S. D. Berkowitz. Cambridge: Cambridge University Press.
14. Wellman, Barry (Ed.). 1999. *Networks in the global village: Life in the contemporary communities*. Boulder, CO: Westview Press.
15. Wellman, Barry, and Scot Wortley. 1990. "Different strokes from different folks: community ties and social support." *American Journal of Sociology* 96:558-588.
16. Woolcock, Michael. 1998. "Social capital and economic development: Toward a theoretical synthesis and policy framework." *Theory and Society* 27:151-208.

### *Methodology*

17. Freeman, Linton C. 1979. "Centrality in Social Networks: Conceptual Clarification." *Social Networks* 1:215-239.
18. Killworth, Peter, Eugene Johnsen, H. Bernard Russell, Gene Ann Shelley, and Christopher McCarthy. 1990. "Estimating the size of personal networks." *Social Networks* 12:289-312.
19. Marsden, Peter. 1984. "Measuring tie strength." *Social Forces* 63:482-501.
20. van der Poel, Mart G. M. 1993. "Delineating personal support networks." *Social Networks* 15:49-70.
21. Wellman, Barry. 1997. "Structural analysis: From method and metaphor to theory and substance." Pp. 19-61 in *Social Structures: A Network Approach*, edited by Barry Wellman and S. D. Berkowitz. Greenwich, CT: JAI Press.

### *Search and Diffusion*

22. Burt, Ronald S. 1987. "Social contagion and innovation: Cohesion versus structural equivalence." *American Journal of Sociology* 92:1287-1335.
23. Lee, Nancy Howell. 1969. *The search for an abortionist*. Chicago,: University of Chicago Press.
24. Rogers, Everett M. 1995. *Diffusion of innovations*. New York: Free Press.
25. Tepperman, Lorne. 1975. "Deviance as a search process." *Canadian Journal of Sociology* 1:277-294.
26. Valente, Thomas W. 1995. *Network models of the diffusion of innovations*. Cresskill, N.J.: Hampton Press.
27. Weimann, Gabriel. 1982. "On the Importance of Marginality: One More Step into the Two-Step Flow of Information." *American Sociological Review* 47:764-773.

### *Technology and Social Networks*

28. Haythornthwaite, Caroline, and Barry Wellman. 1998. "Work, friendship and media use for information exchange in a networked organization." *Journal of the American Society for Information Science* 49:1101-1114.
29. Philip, Howard, Lee Rainie, and Steve Jones. 2001. "Days and nights on the Internet: The impact of a diffusing technology." *American Behavioral Scientist* 45:404-419.
30. Rhinegold, Howard. 2000. *The Virtual Community: Homesteading on the Electronic Frontier*. Cambridge: MIT Press.

31. Smith, Marc. 1999. "Invisible crowds in cyberspace: mapping the social structure of the Usenet." in *Communities in Cyberspace*, edited by Mark Smith and Peter Kollock. London: Routledge.
32. Wellman, Barry. 2001. "Physical Place and Cyper Place: The Rise of Personalized Networking." *International Journal of Urban and Regional Research* 25:227-252.
33. Wellman, Barry, and Milena Gulia. 1999. "Net-surfers don't ride alone: Virtual communities as communitites." in *Networks in the Global Village*, edited by Barry Wellman. Boulder, CO: Westview Press.
34. Wellman, Barry, and Keith N. Hampton. 1999. "Living networked on and off line." *Contemporary Sociology* 28:648-654.

# Technical Area

## *Textual information processing*

---

**Examiner:** Walter R. Bender, Senior Scientist, MIT Media Laboratory

**Description:** This area is intended to represent the wide-ranging approaches to the computer representation of textual information, considering the tasks of representation, retrieval, and navigation. The first reading group is a general overview of the classical methodology and algorithms for indexing information, while the following two focus on the more specific tasks of classification and natural language understanding. The fourth section outlines the relationship between people and information systems, while the final group considers groups of users working in collaboration.

**Requirements:** This area will be examined by a 24-hour take home examination administered by Mr. Bender.

**Signature:** ..... **Date:** .....

### **Reading List:**

#### *Information Retrieval Algorithms*

1. Agrawal, Rakesh, and Ramakrishnan Srikant. 1994. "Fast algorithms for mining association rules." in *20th International Conference on Very Large Databases*. Santiago, Chile.
2. Baeza-Yates, R., and Berthier de Ara jo Neto Ribeiro. 1999. *Modern information retrieval*. New York: Addison-Wesley.
3. Deerwester, Scott, Susan T. Dumais, George W. Furnas, Thomas K. Landauer, and Richard Harshman. 1990. "Indexing by Latent Semantic Analysis." *Journal of the American Society of Information Science* 41:391-407.
4. Lenat, Douglas B., and V. Ramanathan. 1990. *Building large knowledge-based systems*. Reading, Mass.: Addison-Wesley.
5. Mani, Inderjeet, and Mark T. Maybury. 1999. *Advances in automatic text summarization*. Cambridge, Mass.: MIT Press.
6. Rowley, J. E. 1995. *Organizing knowledge: an introduction to information retrieval*. Aldershot, Hants, England; Brookfield, Vt.: Gower.
7. Salton, Gerard. 1968. *Automatic information organization and retrieval*. New York,: McGraw-Hill.
8. Van Rijsbergen, C. J. 1979. *Information retrieval*. London; Boston: Butterworths.

#### *Classification*

9. Borgida, Alexander, and Sowa John F. 1991. *Principles of semantic networks: explorations in the representation of knowledge*. San Mateo, Calif.: Morgan Kaufmann.
10. Clark, Herbert H., and Eve V. Clark. 1977. *Psychology and language: an introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.
11. Kaufman, Leonard, and Peter J. Rousseeuw. 1990. *Finding groups in data: an introduction to cluster analysis*. New York: Wiley.

12. Riloff, Ellen, and Wendy Lehnert. 1994. "Information extraction as a basis for high-precision text classification." *ACM Transactions on Information Systems* 12:296-333.
13. Svenonius, Elaine. 1986. "Unanswered questions in the design of controlled vocabularies." *Journal of the American Society of Information Science* 37:331-340.
14. Yang, Yiming. 1999. "An evaluation of statistical approaches to text categorization." *Information Retrieval* 1:69-90.

### *Linguistics and Natural Language Processing*

15. Fellbaum, Christiane. 1998. *WordNet: an electronic lexical database*. Cambridge, Mass: MIT Press.
16. Grefenstette, Gregory. 1994. *Explorations in automatic thesaurus discovery*. Boston: Kluwer Academic Publishers.
17. Gumperz, John Joseph. 1982. *Discourse strategies*. Cambridge [Cambridgeshire]; New York: Cambridge University Press.
18. Schank, Roger C., and Robert P. Abelson. 1977. *Scripts, plans, goals and understanding: an inquiry into human knowledge structures*. New York: L. Erlbaum Associates
19. Soergel, Dagobert. 1974. *Indexing languages and thesauri: construction and maintenance*. Los Angeles: Melville Pub. Co.

### *Information Users*

20. Bates, Marcia J. 1979. "Information search tactics." *Journal of the American Society of Information Science* 30:205-214.
21. —. 1989. "The design of browsing and berrypicking techniques for the on-line search interface." *Online Review* 13:407-431.
22. —. 1988. "How to use controlled vocabularies more effectively in online searching." *Online* 12:45-56.
23. Belew, Richard K. 2000. "Finding out about: a cognitive perspective on search engine technology and the WWW." Cambridge, U.K.; New York: Cambridge University Press.
24. Furnas, G. W., T. K. Landauer, L. M. Gomez, and S. T. Dumais. 1987. "The vocabulary problem in human-system communication." *Communications of the Association for Computing Machinery* 30:964-971.
25. O'Day, Vicki L., and Robin Jeffries. 1993. "Orienteering in an information landscape: How information seekers get from here to there." Pp. 438-445 in *ACM SIGCHI '93*.
26. Russell, Daniel M., Mark J. Stefik, Peter Pirolli, and Stuart K. Card. 1993. "The cost structure of sensemaking." Pp. 269-276 in *Conference on Human Factors in Computing Systems (INTERCHI '93)*.
27. Salton, Gerard, and Christopher Buckley. 1990. "Improving retrieval performance by relevance feedback." *Journal of the American Society of Information Science* 41:288-297.

### *Collaboration and Information Design*

28. Berlin, Lucy M., Robin Jeffries, Vicki L. O'Day, Andreas Paepcke, and Cathleen Wharton. 1993. "Where did you put it? Issues in the design and use of a group memory." in *Conference on Human Factors and Computing Systems*. Amsterdam, The Netherlands.
29. Goldberg, David, David Nichols, Brian M. Oki, and Douglas Terry. 1992. "Using collaborative filtering to weave an information tapestry." *Communications of the ACM* 35:61-70.



30. Newman, Mark W., and James A. Landay. 2000. "Sitemaps, storyboards and specifications: A sketch of web site design practice." in the *Proceedings of Designing Interactive Systems (DIS 2000)*. New York City.
31. Resnick, Paul, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom, and John Riedl. 1994. "Grouplens: An open architecture for collaborative filtering of netnews." Pp. 175-186 in the *Proceedings of Computer Supported Cooperative Work (CSCW '94)*. Chapel Hill, NC.
32. Rosenfeld, Louis. 1998. *Information Architecture for the World Wide Web*. Sebastopol: O'reilly & Associates.
33. Shardanand, U., and P. Maes. 1995. "Social information filtering: Algorithms for automating 'word of mouth'." in the *Proceedings of ACM CHI-95*. Denver, CO: ACM Press.