Conceptualizing and Advancing Research Networking Systems

TITUS SCHLEYER, BRIAN S. BUTLER, MEI SONG,
and HEIKO SPALLEK, University of Pittsburgh

Science in general, and biomedical research in particular, is becoming more collaborative. As a result, collaboration with the right individuals, teams, and institutions is increasingly crucial for scientific progress. We propose Research Networking Systems (RNS) as a new type of system designed to help scientists identify and choose collaborators, and suggest a corresponding research agenda. The research agenda covers four areas: *foundations, presentation, architecture*, and *evaluation*. Foundations includes project-, institution- and discipline-specific motivational factors; the role of social networks; and impression formation based on information beyond expertise and interests. Presentation addresses representing expertise in a comprehensive and up-to-date manner; the role of controlled vocabularies and folksonomies; the tension between seekers' need for comprehensive information and potential collaborators' desire to control how they are seen by others; and the need to support serendipitous discovery of collaborative opportunities. Architecture considers aggregation and synthesis of information from multiple sources, social system interoperability, and integration with the user's primary work context. Lastly, evaluation focuses on assessment of collaboration decisions, measurement of user-specific costs and benefits, and how the large-scale impact of RNS could be evaluated with longitudinal and naturalistic methods. We hope that this article stimulates the human-computer interaction, computer-supported cooperative work, and related communities to pursue a broad and comprehensive agenda for developing research networking systems.

Categories and Subject Descriptors: K.4.3 [**Computers and Society**]: Organizational Impacts—*Computer-supported collaborative work*; H.5.3 [**Information Interfaces and Presentation**]: Group and Organization Interfaces—*Collaborative computing*; J.4 [**Computer Applications**]: Social and Behavioral Sciences

General Terms: Design, Human Factors, Management, Theory

Additional Key Words and Phrases: Online knowledge communities, social networks, social software, Web 2.0, Web 2.0 applications, Web collaborative software

ACM Reference Format:
Schleyer, T., Butler, B. S., Song, M., and Spallek, H. 2012. Conceptualizing and advancing research networking systems. ACM Trans. Comput.-Hum. Interact. 19, 1, Article 2 (March 2012), 26 pages.
DOI = 10.1145/2147783.2147785 http://doi.acm.org/10.1145/2147783.2147785

B. S. Butler is currently affiliated with the Robert H. Smith School of Business, University of Maryland, College Park.
This project was, in part, supported by a grant (1 U54 RR023506-01) from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH) and NIH Roadmap for Medical Research. Addition funding was provided by the National Science Foundation (OCI-0951630).
Various parts of this article are built on discussions and findings in Spallek et al. [2008] and Schleyer et al. [2008a; 2008b].
Authors' addresses: T. Schleyer (corresponding author), Center for Dental Informatics School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA; email: titus@pitt.edu; B. S. Butler, School of Information Studies, Robert H. Smith School of Business, University of Maryland, College Park; Joseph M. Katz Graduate School Business and College of Business Administration, University of Pittsburgh, Pittsburgh, PA; M. Song, H. Spallek, Center for Dental Informatics School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA.

DOI 10.1145/2147783.2147785 http://doi.acm.org/10.1145/2147783.2147785

1. INTRODUCTION

Over the past several decades, science has become significantly more collaborative [Adams et al. 2002; Arzberger and Finholt 2002; Katz and Martin 1997; Rhoten 2007; Zerhouni 2003]. Increases in the number of international collaborations, coauthored papers, and multi-investigator grant proposals are evidence for this trend [Olson et al. 2008a], as is the rising frequency of terms such as "interdisciplinarity" and "multidisciplinarity" in the literature [Braun and Schubert 2003]. Olson et al. cite multiple reasons for this development: "the urgency, complexity and scope of unsolved scientific problems; the need for access to new, and often expensive, research instruments and technologies; pressure from funding agencies; and information and communication technologies that facilitate interaction and sharing" [Olson et al. 2008a]. Therefore, collaboration among the right individuals, teams, and institutions is becoming ever more crucial for progress in science.

Finding "optimal" (regardless of how one defines the term) collaborators, however, is difficult, and becoming more so [Schleyer et al. 2008a; 2008b; Spallek et al. 2008]. Establishing collaborations is a labor-intensive and risky process, especially when multiple disciplines are involved. Collaboration seekers often struggle with the target disciplines' terminology, have difficulty identifying true experts, and lack relevant social contacts. In addition, they must assess potential collaborators in light of many criteria [Schleyer et al. 2008b], a process impeded by incomplete, fragmented information. Finally, reviewing potential collaborators does not scale well. Assessing the n^{th} candidate takes as much work as assessing the first. At the same time, the universe of collaborative opportunities continues to expand as information about researchers becomes more accessible and remote collaborations become more feasible [Katz and Martin 1997].

Recently, the term "Research Networking Systems" (RNS) became popular to describe electronic systems designed to help researchers find collaborators. "Research networking system" emerged as an alternative to "research collaborator discovery system," "expertise location system," and other terms after the National Center for Research Resources awarded a $12m grant to the University of Florida to develop a national prototype system[1]. The request for applications solicited proposals to develop "infrastructure for connecting people and resources to facilitate national discovery of individuals and of scientific resources by scientists and students to encourage interdisciplinary collaboration and scientific exchange" [National Center for Research Resources 2009].

In light of this goal, we propose the following definition for RNSs.

> Research Networking Systems (RNS) are systems which support individual researchers' efforts to form and maintain optimal collaborative relationships for conducting productive research within a specific context.

Several aspects of the definition are noteworthy. While RNSs can serve other purposes, such as managing a university's research portfolio, the primary users whose needs must be met are "individual researchers." RNSs are intended to help "form and maintain" relationships, not complete collaborative tasks. "Collaborative relationships" refer to the interpersonal ties that support successful research collaborations. While the nature of these relationships is subject to ongoing debate, our definition assumes that they involve shared, two-way interests; ongoing, often sporadic, interaction; and the creation of joint work products. "Optimal" is a subjective and situational measure, yet searching for the best possible opportunities is central to RNSs. The aspect of "productive" research speaks to the collaboration outcomes. While papers, presentations,

[1]See http://www.vivoweb.org

and other scientific artifacts are generally accepted metrics of research productivity, they are arguably imperfect. Lastly, "context" is included in the definition of RNSs because of its importance in shaping research collaborations. Context includes factors such as researchers' needs and goals, project characteristics, organizational policies, disciplinary norms, and institutional constraints. Successful RNSs provide the information individual researchers need to develop and maintain contextually embedded collaborative relationships.

The goal of this article is to stimulate foundational research on research networking systems that takes into account what is known about collaboration, expertise location, and social networking. We hope to challenge researchers in multiple fields by proposing claims and corresponding research questions that can be tested and/or investigated.

Two considerations must be mentioned to put the proposed research agenda in context. First, in reviewing the literature, we draw from studies of many scientific disciplines, including computer science and biomedicine. Disciplinary culture, values, and norms have a significant impact on collaborative relationships. In order to frame the discussion, this article uses examples from biomedical research. In consequence, the relative importance of the issues we identify may vary in other disciplines. Second, we discuss RNSs primarily in the context of academic research. While the proposed research agenda may be applied to corporate research and development environments, academic research domains are complex and distinct enough to merit separate consideration.

2. RESEARCH NETWORKING AND COLLABORATOR DISCOVERY

Literature relevant to RNSs includes topics such as expertise location systems, formation of scientific collaborations, and the use of technology in research collaborations and social networking. In particular, research on expertise location and sharing [Ackerman et al. 2003; McDonald and Ackerman 1998] informs the discussion of RNSs because finding collaborators involves searching for individuals with specific expertise. We therefore review prior work on expertise location systems before discussing existing RNSs.

2.1. Expertise Location vs. Research Networking

Expertise location is a concern in several contexts, including "expertise locating systems" [McDonald and Ackerman 2000], "knowledge communities" [de Vries and Kommers 2004; Erickson and Kellogg 2003], and "communities of practice" [Johnson 2001; Millen et al. 2002]. Zhang et al [2007] defined Expertise Locator Systems (ELS) as "CSCW systems that help find others with the appropriate expertise to answer a question." In a review of contemporary ELSs, Becerra-Fernandez [2006] described them as knowledge sharing systems that "point to experts, those that have the knowledge." Others have defined ELSs in terms of the functions they perform. For example, ELSs can "connect people to people; link people to information about people; identify people with expertise and link them to those with questions or problems; identify potential staff for projects requiring specific expertise; assist in career development; and provide support for teams and communities of practice"[2].

The CSCW literature contains numerous references to expertise location and the design of expertise location systems [Ackerman and Palen 1996; Ehrlich et al. 2007; Friedman et al. 2000; Jacovi et al. 2003; Mattox et al. 1999; McDonald and Ackerman 2000; Mockus and Herbsleb 2002; Streeter and Lochbaum 1988]. This body of work can help us compare and contrast ELSs and RNSs.

[2]See http://www.kmdedge.org

First, locating an expert and establishing a research collaboration both involve looking for and discovering expertise. The focus of expertise location is finding an answer, a solution, or a person with whom details of a problem can be discussed [Ehrlich et al. 2007]. The need is largely determined by the task at hand. This emphasis is reversed when forming research collaborations. Researchers looking for collaborators primarily seek a person to establish a relationship with. The specific task or problem is secondary to forming and maintaining this relationship.

Second, the comparatively shorter time horizon of interaction in expertise location allows for benefits which are more asymmetrically distributed. Individuals looking for an answer often stand to gain more than the experts providing it [Lakhani and von Hippel 2003]. In research collaborations, on the other hand, benefits must be more evenly distributed because they often span multiple collaborative tasks and projects, and extended time frames.

Third, ELSs are designed for situations where the goal is defined but needed knowledge is "hidden." To succeed, individuals must extract answers from the set of available experts. In contrast, scientific researchers often work with ill-defined questions and objectives that shift over time. These collaborative relationships reflect the nature of scientific inquiry in which large problems are pursued incrementally in a meandering, exploratory fashion. The query-driven approach is complemented by an opportunity-driven one, with new directions emerging serendipitously as methods and concepts developed in one area find novel uses in another.

Last, in industry, where most ELSs are deployed, individuals typically work within a single organization. Project assignments, team memberships, and immediate colleagues are determined by management. In academia, scientists often work across institutional boundaries [Cummings and Kiesler 2005; Olson et al. 2008b] and have significant autonomy when selecting their projects, affiliations, and collaborators.

In summary, while ELSs and RNSs have common functions, they also differ significantly with respect to user characteristics, organizational context, and the goals they serve. Thus, while prior work on ELSs provides a useful starting point for discussions of RNSs, we must also consider systems specifically designed for supporting research networking.

2.2. Current Research Networking Systems

While there is a relatively large body of literature on expertise location systems [Ackerman et al. 2003; Becerra-Fernandez 2006], studies of research networking systems are rare. A focused literature search identified descriptions of only five systems which have been tested and/or implemented. Several other recently developed systems have not been described in the literature.

At the University of Pittsburgh, an application called Faculty Research Interests Project (FRIP) helps faculty establish collaborations [Friedman et al. 2000]. FRIP indexes faculty research interests using Medical Subject Headings (MeSH) [Coletti and Bleich 2001] and draws on MEDLINE-indexed publications to populate its database. In 2000, FRIP indexed 1,925 research faculty at the six schools of the University of Pittsburgh's Health Sciences Center. FRIP's functionality is currently being replaced by Pitt's Digital Vita (see the following) system.

A second recently developed tool for helping connect researchers with shared interests is a Facebook application called MEDLINE Publications (MP) [Bedrick and Sittig 2008]. The system uses the PubMed database to automatically create user-customizable lists of publications. The system includes a rudimentary recommendation algorithm to identify other users with similar publication profiles. Like FRIP, MP uses MeSH as the controlled vocabulary for specifying research interests. MP has attracted a reasonable user base, and anecdotal evidence suggests that it has been useful to some.

A third research networking system is Searchable Answer Generating Environment (SAGE), a searchable repository of funded research information for all universities in Florida [Becerra-Fernandez 2006]. This system implements a distributed database schema that can be searched by criteria such as research topic, investigator name, funding agency, and university. To keep the data repository current, participating institutions must provide funding data on an ongoing basis. Researchers across Florida benefit from SAGE increasing their visibility and facilitating efforts to locate potential collaborators at other universities, in industry, and in federal agencies. SAGE has also been used by NASA and small businesses to identify university researchers for collaboration. As of 2006, the SAGE database included about 7,817 researchers and 53,124 projects from fourteen institutions throughout Florida.

Liu et al. [2005] described a system that uses RDF (Resource Description Framework) for expertise matching by integrating data from multiple, heterogeneous sources and making them available through concept-based searches. An initial prototype system was evaluated in the School of Computing at the University of Leeds. Results indicate that the RDF-based expertise matching system outperforms traditional DBMS techniques because it improves match accuracy and facilitates expertise selection.

Last, Schleyer and colleagues [2008a] proposed the Digital Vita system as a prototypical design and architecture responsive to initial requirements for research networking [Schleyer et al. 2008b]. Digital Vita includes four main functions: maintaining, formatting, and semiautomated updating of biographical information; searching for researchers; building and maintaining social networks; and managing document flow. The system departs from other approaches for representing researchers in that it is built around a researcher's academic *Curriculum Vitae* (CV). While not perfect, the CV is often the most up-to-date and comprehensive document describing a scientist's accomplishments and activities. With its focus on CV maintenance, integration with the local context, and provision of benefits for individual researchers, Digital Vita has the potential to reduce adoption barriers, represent researchers more comprehensively than keyword-based profiles, and achieve ongoing system utilization.

In addition to the five systems described in the literature, several other research networking systems exist in academia and industry. Academic systems include the University of Florida's VIVO[3] project [Gewin 2010], Harvard's Catalyst Profiles[4], and the University of Iowa's Loki[5]. The Distributed Interoperable Research Experts Collaboration Tool (DIRECT)[6] is a recent initiative to allow users to search for experts across these systems. Commercial systems include the Community of Science (http://www.cos/com), Index Copernicus Scientists (http://scientists.indexcopernicus.com/), Research Crossroads (http://www.researchcrossroads.com/), BiomedExperts (http://www.biomedexperts.com/), and Epernicus (http://www.epernicus.com).

Each of these systems has a different approach for creating searchable directories of researchers. As a result, they provide useful insights into the architectural and data management problems associated with gathering and storing researcher profiles. However, as with expertise location systems, the research networking systems described in the literature only partially address the requirements of research networking.

2.3. Research Networking Challenges in Biomedical Sciences

While the marketplace and academic institutions have begun implementing expertise-focused research networking systems, there is a need for theories and models to

[3]http://www.vivoweb.org/
[4]http://connects.catalyst.harvard.edu/Profiles/search
[5]http://www.icts.uiowa.edu/Loki/
[6]http://www.direct2experts.org

inform RNS design, implementation, and evaluation. No extant studies directly consider RNSs. Nonetheless, the literature on scientific collaboration and collaboration formation provides some insight into the problems that RNSs are intended to address.

A recent study by Weng et al. [2008] showed that collaboration on cross-cutting research topics such as obesity is not well served by the traditional organization of biomedical research institutions. The authors identified obesity researchers using several search strategies (Google, PubMed, and snowball sampling) and surveyed them to determine departmental/center affiliation, collaborators, and research interests. Participants were distributed over multiple departments and often affiliated with more than one research center. Respondents who collaborated with others had 8.8 collaborators on average, indicating a relatively active community. Some research groups, however, were only connected by a single pair of individuals. Institution-level success factors for interdisciplinary collaboration suggested by the study included "(1) establishment of interdisciplinary research centers; (2) identification of boundary spanners who link dispersed research communities; and (3) creation of scientific journals that publish transdisciplinary research results." The findings of this study suggest that interdisciplinary collaborations could be organized as "virtual teams" [Hinds et al. 2002].

In a more general attempt to understand how research collaborations are formed in the health sciences, Spallek and colleagues [Schleyer et al. 2008a; 2008b; Spallek et al. 2008] conducted semistructured interviews with 27 biomedical scientists at the University of Pittsburgh. The study focused on general aspects of subjects' collaboration activity, such as who they were currently collaborating with, what motivated them to seek collaborators, and how they searched for them. Four main groups of factors were found to affect collaboration-seeking: motivation, evaluation, search and selection, and barriers. Participants who reported using directories such as FRIP or Community of Science noted that they were useful for people new to an institution and for finding individuals outside the immediate work context. However, researcher directories were seen as limited because of incomplete coverage of research domains; sparse, outdated researcher profiles; and lack of support for leveraging social networks. Although this study did not focus on research networking systems, its results suggest that developing and refining such systems would have significant practical utility.

In parallel, our research group also formulated an initial set of requirements for collaborator discovery systems in biomedical science [Schleyer et al. 2008b]. The study used affinity diagramming, literature reviews, contextual inquiries, and semistructured interviews to develop a list of requirements for systems for finding collaborators. The requirements include: a good cost/benefit ratio for the user when creating and updating online profiles; representation of researchers through rich, comprehensive, and up-to-date information; exploitation of social networks; assessment of potential collaborators' "soft" traits, such as personality and work styles; use of multiple indicators of past collaboration activity; user-modifiable preferences regarding privacy and public availability of profile information; effective cross-disciplinary search; and active highlighting of "nonintuitive" connections between researchers.

Existing studies show that RNSs must function in a complex socio-technical context. They are subject to multiple, sometimes conflicting, requirements that must be balanced carefully in order to maximize system utility for all user populations. While there is a growing body of work which examines the factors underlying effective research collaborations, many unanswered questions remain about how to best use information technology to facilitate research networking.

3. RESEARCH AGENDA FOR RESEARCH NETWORKING SYSTEMS

The following research agenda is organized around four areas that contribute to RNS success: *foundations, presentation, architecture,* and *evaluation*. Foundations addresses theoretical models, core principles and general factors that underlie the design of effective RNSs. Presentation examines issues concerning user interfaces, interaction design, and system functionality. Architecture discusses the internal design of RNSs, how they interact with external information sources, and interoperability. Finally, evaluation is concerned with how RNS outcomes can be framed and measured.

While the proposed categorization of the particular claims and research questions may be debated, the four areas are critical aspects of RNS design and implementation. They support both targeted investigation of issues and identification of useful links to the diverse body of existing research. In each area, we posit claims regarding the nature of collaborative relationships and RNSs. Each claim is followed by a brief review of the relevant literature and a list of open questions which, if addressed, would significantly improve our ability to design, implement, and evaluate research networking systems. The goal of this research agenda is to advance the study and development of RNSs, and to make them a useful part of the scientific enterprise. Hence, the open questions were selected to focus attention on issues particular to RNSs as opposed to related systems, such as virtual communities, expertise location, and cooperative work.

3.1. Foundations

While it may be convenient from a systems design perspective to conceptualize research networking as a search or information display problem, RNSs must support a more complex set of social behaviors. In this section we describe three foundational perspectives on collaborative relationships, and examine their implications for the design and evaluation of RNSs.

Claim 1. To form collaborative relationships, individuals must balance the different motivations of potential collaborators in the context of projects, institutions, and disciplines.

Many researchers have proposed models for describing effective collaborations [Suchman and Trigg 1986]. Existing frameworks focus on various aspects of collaboration, including key concepts/variables at work in research collaborations [Katz and Martin 1997; Larson 2003; Melin 2000; Suchman and Trigg 1986], participants in a collaboration and the division of labor [Jenerette et al. 2008; Kouzes et al. 1996], and the process of collaboration and activities involved at each stage [Gitlin et al. 1994; Kraut et al. 1987]. In part, this body of work has also explored the motivations and mechanisms underlying collaboration formation.

At societal level, researchers have examined the transformation of modern science and the social, cultural, and technological factors that drive collaboration [Börner et al. 2010]. These factors include use of expensive, sophisticated instrumentation [Olson et al. 2008a]; more emphasis on application; greater specialization and concentration of resources [Ziman 1994]; changing patterns and levels of funding; and the growing professionalism of science [Katz and Martin 1997]. However, the move towards a greater degree of collaboration in science is not without problems [Cummings and Kiesler 2007]. Multi-university collaborations face significant coordination challenges which, if not addressed, can lead to suboptimal project outcomes [Finholt and Olson 1997].

At project level, factors that affect collaboration include problem complexity and scale, division of labor, and degree of specialization [Laudel 2002; Rhoten 2007]. Institutional factors that influence collaboration activity include role specialization [Madanmohan and Navelkar 2004], the nature of the work [Birnholtz 2007], the radicalness of the research [Belkhodja and Landry 2007], access to particular resources [Mattessich and Monsey 1992], structural characteristics of organizations [Walsh and Maloney 2007], organizational processes [LeGris et al. 2000], organizational management and support [Millen et al. 2002], and funding contingencies [Bos et al. 2007]. Many things which motivate individual scientists to collaborate, such as the need for knowledge, expertise, and skills [Beaver 2001]; access to special equipment and funding [Melin 2000]; the desire for social relationships [Fox and Faver 1984; Terveen and McDonald 2005]; and the need to educate and mentor students [Katz and Martin 1997; Melin 2000], are directly linked with these project and institutional factors.

As with any long-term relationship, collaborations can only be maintained if the work and incentive structures are aligned so that all of the involved individuals benefit from participation [Numprasertchai and Igel 2005]. Successful RNSs must support individuals' efforts to identify potential collaborators whose needs and incentives complement their own. Through rich user models and appropriately designed profiles, RNSs can leverage information about institutional, project, and individual factors to help collaboration seekers' detect when and where collaboration is useful and feasible. This suggests that the following questions are central to the design and creation of effective RNSs.

—How should RNSs model user characteristics known (or hypothesized) to affect willingness of individuals to engage in collaborations? While a variable such as "age" is easy to model, others like "seniority" or "technical competence" are more difficult to represent.
—How should RNSs incorporate project-related, institutional, social, structural, and cultural characteristics which affect individuals' motivation to participate in collaborative relationships?
—How should RNSs model the conditions under which researchers start looking for a collaborator? Is a single model sufficient? How should the model evolve over time as careers, accomplishments, and interests change?

Claim 2. Exploiting social networks is essential for efficient and effective research networking.

People work within social networks. Although these networks may cross organizational boundaries and span geographic distance, individuals are still constrained by who they know and what they know about them. As a result, many expertise location systems developed in recent years have integrated social network information to help evaluate potential experts and facilitate communication with them [Kautz et al. 1997a; Ogata et al. 2001]. McDonald [2003] compared two different social networks as alternative bases for recommending experts within a medical software company. The first network, based on shared work contexts, captured network ties arising from work arrangements. The second, the socializing network, linked individuals who interacted socially. The results illuminated a number of critical issues to consider in development of RNSs. Using network information forced a trade-off between finding the most knowledgeable person and finding the person with whom the searcher could most easily interact. Also, users still sometimes desired broader recommendations even if the system's recommendations were appropriate. Lastly, users often preferred their own egocentric social network over the one generated and recommended by the system.

In another system, Yang and Chen [2008] developed a mathematical model of a three-layer social network to support interactive collaboration, taking into account the knowledge relationship and social relationship ties of potential collaborators. In this system, a peer-to-peer knowledge net is overlaid with the peer-to-peer social net. An Instant Messaging (IM) system helps individuals communicate with peers identified through the social network. Preliminary evaluation of this system with student users showed that most were willing to use this system to find others open to sharing their knowledge.

Many methods for gathering social network information have been suggested. Social networks have been constructed based on email exchanges among individuals [Ogata et al. 2001], Web pages related to a person, the Database systems and Logic Programming (DBLP) bibliographic information service for computer science, and the publication ranking list from Citeseer [Li et al. 2007]. Pavlov and Ichise [2007] built link predictors which identify potential collaboration opportunities using the structural information in coauthorship networks. However, social networks derived from coauthorship are likely to be imperfect representations of a researcher's collaborative relationships [Katz and Martin 1997]. To overcome this problem, McDonald and Ackerman [2000] used participant observation, formal and informal interviews, and pile sorts. Yang and Chen [2008] had users fill out forms and answer questions about peers' knowledge and social ties. The Digital Vita system blends the two strategies, allowing researchers to specify collaborative relationships explicitly through "colleague requests" (equivalent to "friend requests" in Facebook) [Schleyer et al. 2008a] while also deriving implicit ties such as coauthorship and shared department membership from CVs.

In traditional social networks, individuals rely on their contacts to provide access to a wide range of information and opportunities [Adler and Kwon 2002]. Supporting searches within a network is an important part of facilitating collaboration formation. Previous research on search strategies in social networks has identified two main approaches. The first is automation of the small-world approach, where the target is known by name or a unique identifier [Adamic and Adar 2005; Yang and Garcia-Molina 2002]. Adamic and Adar [2005] simulated small-world experiments on an email network in an organization and a student social networking system Web site. They found that small-world search strategies using a contact's position in physical space or an organizational hierarchy could effectively locate the most appropriate individuals. However, in a social network where hierarchical structures were not well defined, local search strategies were less effective.

A second approach for searching within a social network focuses on locating a person with specific expertise or knowledge. Zhang and Ackerman [2005] evaluated three families of strategies for searching using social network information. These strategies were based on computation, network structure, or individual similarity. The computational approach, for instance, used breadth-first Search to broadcast a query to a person's neighbors. Information scent search, on the other hand, selected the person with the highest match score between the query and his profile [Yu and Singh 2003]. In a simulation on an organization's email dataset, the different strategies affected the search process in important ways. For example, weak ties [Granovetter 1973] appeared more effective for seeking new information, but the relative rank of different algorithms changed little when examining social costs.

The importance of existing network structures in formation of collaborations suggests that the following questions are critical for design of effective RNSs.

—How can information about researchers' social and collaborative networks be gathered and maintained efficiently? How can implicit relationships, such as coauthorship, be refined and/or augmented to serve as a basis for constructing social networks?
—How can explicit relationship identification be applied in RNSs? Should network size be limited to avoid "colleague inflation"?
—How should social network data be used to support collaboration seeking? Should users be encouraged to focus on relatively small social distances [Schleyer et al. 2008a] or explore lengthy referral chains [Kautz et al. 1997b]?
—How should existing and potential collaborative relationships be represented? Should weak ties be distinguished from (and perhaps given priority over) strong ties? Should

potential collaborators be ranked based on the number of current collaborators (i.e., network degree)?

—How can boundary-spanning individuals be identified and leveraged in order to generate collaboration opportunities?

Claim 3. Establishing collaborations requires individuals to form impressions of and evaluate potential collaborators based on information beyond expertise and interests.

In the research collaboration literature, few studies have focused on the initiation of collaborations. Kraut et al. [1987] suggest that collaboration formation is more a process than an event. The initiation stage involves both relationship- and task-related activities. For the relationship, the essential activity is determining whether potential collaborators are acceptable partners. At task level, participants must identify collective research objectives and formulate specific work plans. If a collaboration is to succeed, researchers must develop from mere acquaintances to committed partners. Kraut identified two paths for this process. For some researchers, the initial contact evolves into joint commitment the way a bilateral friendship develops. For others one partner proposes collaboration just like during an asymmetric courtship ritual. Whichever way collaborations develop, serendipitous and informal conversations are an important early step.

Prior studies have examined a variety of factors that affect how prospective collaborators are evaluated, including competence, complementarity, work and collaboration styles [Axelrod 1984], personality, and physical proximity [Kraut et al. 1987]. Schleyer et al. [2008b] identified support for compatibility assessment as a key element of RNS requirements. For instance, the system should enable users to find collaborators compatible in personality, work style, and other factors. An individual's likely availability, accessibility, and willingness to engage in collaboration also might impact her selection as a potential collaborator. Several studies found that researchers trust personal recommendations when assessing compatibility with potential collaborators [Beaver 2001; Flynn 2005]. Interaction with potential collaborators is another way researchers gather information about their compatibility. Face-to-face interaction seems to produce the highest trust among unfamiliar collaborators [Moore et al. 1999]. In the absence of face-to-face opportunities, strategies such as chat sessions and exchange of personal information can also help to overcome limited availability of information [Zheng et al. 2002]. While there are a range of evaluation criteria, their relative importance appears to be context- and situation-dependent. For instance, work style compatibility may only be a minor constraint for a collaboration based on sharing equipment or other scarce physical resources [Spallek et al. 2008].

While much research in the expertise location system literature has focused on expertise representation, there is evidence from studies of work relationship formation that expertise may sometimes be a secondary concern when selecting collaborators [Casciaro and Lobo 2005; 2008]. Studies of social matching, such as the work of Terveen and McDonald [2005], show that personal characteristics must be taken into account during the matching process. This suggests that providing information, either directly or indirectly, about traits such as personality, friendliness, character, trustworthiness, sense of humor, and work style may be relevant in the design of RNSs. The apparent difficulty of obtaining information about these traits is one reason why social connections are so important in collaborator discovery: they can be a source of information about personal traits. The importance of this information in collaborative relationship formation suggests that the following research questions are central to the study of RNSs.

—What collaborator traits, other than expertise and interests, are useful in making collaboration decisions?

—How can traits such as productivity, work style, adherence to deadlines, organization, communication style, conflict resolution skills, and personality be assessed, modeled, and presented? Which traits should be highlighted in interfaces designed to support evaluation of potential collaborators?
—What features and technologies are best suited for supporting joint exploration of relationship and task issues during the initial stages of collaboration formation?

While they are not the only possible characterizations of collaborative relationships, the three theoretical perspectives presented here (collaborative relationships as balanced incentive structures; embedded social ties; and the result of impression formation) provide a foundation for defining requirements for RNSs.

3.2. Presentation

At their core, RNSs are systems that capture, store, and present data about people and relationships. The interface used interact with these data significantly affects how an RNS influences users' efforts to form and maintain collaborations. In this section, we consider aspects of presentation and representation that theory and prior work suggest will be critical for the creation of successful RNSs.

Claim 4. RNS must describe potential collaborators' expertise and interests in a comprehensive and up-to-date manner.

Early attempts at compiling representations of expertise relied on data provided by the user, typically in the form of profiles. HelpNet, for instance, asked users to fill in and maintain profiles [Maron et al. 1986]. This approach, however, often suffers a lack of compliance [Ehrlich 2003]. As a result, much attention has been devoted to the automated acquisition of expertise information. Sources used include published documents such as resumes [Becerra-Fernandez 2006]; Wikipedia content, discussions, and user data [Demartini 2007]; literature databases [Friedman et al. 2000]; newsgroup postings [Terveen et al. 1997], and online community site data [Bojars et al. 2008].

A key limitation of these approaches is that they conflate an individual's credentials, expertise, and interests, each playing a different role in the evaluation of a potential collaborator. Credentials project an image of general competence in a domain, such as medicine or law. Expertise specifies knowledge and prior experience in one or more topics in that domain. Statements of interest provide information about current motivations. Collaboration seekers typically examine all three areas when assessing potential matches. For instance, a researcher's publications provide a historical record of performance which is only useful in the context of current interests. If current interests do not match those of the collaboration seeker, even a highly productive publication record is irrelevant.

Derivation, representation, and presentation of potential collaborators' expertise and interests are critical to the design of effective, sustainable RNSs. In light of this, we propose the following research questions.

—How should type and extent of expertise and interests be represented to help researchers make nuanced and valid collaboration decisions?
—Can researcher interests be inferred computationally or do they have to be specified by the user? Should both methods be used together?
—How can the representation of a researcher's expertise and research interests be kept up-to-date with minimal user effort? To what degree can current activity be inferred computationally, for instance, through the semantic Web [Schleyer et al. 2008b]? How should current and past activities be summarized and displayed to support identification and evaluation of collaboration potential?

Claim 5. RNS must represent individuals' expertise, interests, and activities using controlled terminologies.

Some fields, such as biomedicine, have a strong tradition of using controlled terminologies [Coletti and Bleich 2001]. Others, such as computer science, do not. Folksonomies [Woolwine et al. 2011] have multiple advantages and benefits for indexing documents and people, including their authentic use of language and multiple potential interpretations. However, they also create problems for representing concepts in ways that are commonly understood [Peters and Stock 2007]. Two approaches have been proposed to address the limitations of user-created tags. One is to improve users' "tag literacy" [Guy and Tonkin 2006], while the other considers tags as natural language elements amenable to automatic NLP methods [Stock 2007].

A recent study by Lee and Schleyer [2012] found minimal overlap between social tags and controlled index terms for a sample of 231,388 biomedical research papers. A resulting challenge for RNSs is how to balance the effects of controlled and user-generated terminologies. Controlled terminologies are valuable because they provide high-quality information about a potential collaborator. They enable cross-disciplinary searches, support identification of synonyms and related terms, and facilitate automatic discovery of otherwise undetected similarities between individuals. Yet, controlled vocabularies necessarily place constraints on individuals' ability to describe their expertise, interests, experiences, and characteristics in their own terms. To the degree that research networking is a process of impression management and impression formation, use of controlled vocabularies may be perceived by RNS users and subjects as unnecessarily limiting. The potentially conflicting implications of controlled vocabularies in the context of RNS suggest the following research questions.

- Are existing controlled terminologies and taxonomies for indexing publications, such as the Medical Subject Headings [Coletti and Bleich 2001] and the ACM Computing Classification System, adequate for representing individuals' expertise, interests, and characteristics? If not, how should they be improved or expanded?
- How should expertise and interests be represented in domains which lack widely accepted controlled terminologies?
- What are the strengths of folksonomies and social tagging for representation of individual researchers? When and how should controlled and user-generated terms be combined in researcher profiles?
- How does use of controlled terminologies affect individuals' willingness to use, create and maintain profiles within an RNS?

Claim 6. RNSs must allow users to search and visualize researcher profiles in multiple ways.

RNSs are designed, in part, to make the large search spaces of potential collaborators tractable and accessible. A tension exists between focused result sets, in which the system provides a few, presumably high-quality, matches and broader ones, which require more user effort to explore. McDonald's work [McDonald 2003] suggests that RNSs should allow user experimentation and adaptation of the system for different purposes.

Previous work suggests that allowing users to apply different types of criteria may be beneficial. The Expertise Recommender [McDonald and Ackerman 2000] offers "Departmental" and "Social Network" as filters for system recommendations. The SmallBlue system implements a social-context-aware expertise search system that presents an unfiltered list of experts with information about the degree of separation, allowing the user to select the "right" person using social connection information [Ehrlich et al. 2007].

RNSs must incorporate and combine traditional methods of locating collaborators, such as social networks and expertise database searches. RNSs which treat collaborator identification as a decontextualized search process based on impersonal expertise profiles are unlikely to have much impact on users' relationship formation and maintenance activities. Yet, RNSs which only reveal opportunities in the user's immediate social context will overlook potentially fruitful chances for novel and interesting relationships. Taken together, these issues suggest the following research questions regarding the need for diverse presentation and discovery strategies in RNSs.

—What different strategies should RNSs support for locating collaborators? When are strategies based on information artifacts, general profiles, and/or existing network structures most effective?
—What types of filters and representation are most useful to users when navigating the research collaboration search space?
—Should the presentation of RNS information depend on user characteristics, project features, or disciplinary norms? What are the primary dimensions that can be varied to create high-impact, individualized representations?
—Which search algorithms minimize the user effort required to search efficiently and effectively for collaboration opportunities?

Claim 7. RNSs must balance the tension between seekers' need for comprehensive information and potential collaborators' desire to control how they are seen by others.

A collaboration seeker's desire for comprehensive information needs to be balanced with potential collaborators' requirements for privacy and access control [DiMicco and Millen 2007; Hewitt and Forte 2006]. Privacy is not as central in expertise location systems as it is in RNSs [Bellotti 1996; Fogel and Nehmad 2009] because expertise location focuses on task-oriented, episodic interactions. The long-term relationships that RNSs help establish, on the other hand, are central to an individual's professional identity, career success, and self-efficacy. As a result, how an individual is presented to and seen by others in an RNS is an important factor [Goffman 1959; Leary 1996; Schlenker 2003; Schlenker and Leary 1982].

Being visible and accessible in an RNS also carries different costs depending on individual characteristics. To some, the benefits of greater visibility outweigh any potential costs [Gross et al. 2005]. Others may find the loss of privacy and control unacceptable [Mann 2007; Rosenblum 2007]. A senior scientist with many existing collaborations may want to be less visible than a junior scientist for whom exposure can be advantageous. Thus, availability of privacy and access controls may be critical for an individual's willingness to participate in an RNS.

Taken together, these issues suggest a fundamental tension in RNS design. For individuals seeking to form collaborations, the value of an RNS increases if it can provide comprehensive information about potential collaborators. However, the individuals being profiled may be wary of a public presentation of their expertise, interests, past activities, and personal characteristics that encourage detailed comparisons with others. Effective RNSs must balance the needs of both collaboration seekers and potential collaborators. This requirement suggests the following questions.

—How does a researcher's willingness to share different types of profile information vary and what are implications for RNS design? For instance, while researchers are unlikely to object to sharing public information, under what conditions will they be willing to share information about current research projects?
—How do individuals react when information from many public sources about them is presented in one place?

—How does the willingness to share information vary with the personal, social, and organizational distance to others? For instance, are researchers more or less willing to share information with other researchers in their home discipline?
—How should RNS allow researchers to control privacy and public availability of information about themselves? How much control is reasonable without reducing the system's utility?

Claim 8. RNSs should support serendipitous discovery of collaborative opportunities.

While query-driven interfaces play an important role in supporting research networking, effective RNSs must also promote appropriate serendipitous discovery. Like successful entrepreneurs [Gaglio and Winter 2009], high-impact researchers are able to accomplish their goals in part because they can recognize and capitalize on emerging opportunities not obvious to them. Although deliberate planning and intentional search are an important part of forming collaborations, so too is the ability to identify and respond to unanticipated opportunities that emerge from the complex social, institutional, and intellectual environments in which research takes place. To fully support researchers' efforts to form collaborative relationships, RNSs must facilitate both the intentional and serendipitous discovery of potential collaborators.

Matching services have been used with success in many social contexts, but it is less clear how they would be applied to research collaborations. The literature on social matching and collaborative support contains a number of algorithms to match potential partners [Budzik et al. 2002; Pavlov and Ichise 2007; Terry et al. 2002; Zhang et al. 2007]. For example, Yenta is a distributed agent-based system that groups people with common interests by examining the content of their file systems [Foner 1996]. MEDLINE Publications, a scientific collaboration tool built on Facebook, offers a recommendation engine that helps connect a user with others who have similar publication profiles, thereby exposing him to new potential collaborators [Bedrick and Sittig 2008]. Active matching services, similar to the current awareness systems offered by many literature databases, might be set up to proactively notify users about potential collaboration opportunities. This RNS feature, if properly calibrated, would promote opportunistic formation of collaborative relationships.

Addressing the following questions could be useful in determining how RNSs can best facilitate serendipitous collaborations.

—What algorithms are most useful for identifying potential collaboration partners? What variables should they take into account?
—Should users be able to customize the recommendation and matching algorithms used in RNSs? What features/aspects of the matching process should be user-modifiable?
—How can RNSs obtain and incorporate feedback about the usefulness of suggested matches [Melin 2000]?
—Can RNSs help identify the "gaps" in science which present significant research opportunities? How can results of conceptual gap analyses be combined with social network data, researcher profiles, and user characteristics to recommend meaningful novel collaboration opportunities?

The emergence of RNSs creates an opportunity for HCI researchers and developers to apply their understanding of presentation and user experience design to a problem domain that has previously only marginally been supported by technology. Novel aspects of research networking, such as presenting multidimensional researcher profiles, supporting boundary-crossing discovery, and balancing the often conflicting needs of searchers and subjects, present important design challenges. Addressing these challenges will advance our understanding of how to develop complex but usable interfaces can facilitate research networking.

3.3. Architecture

Although individual researchers have significant autonomy in determining the direction and nature of their collaborative efforts, research collaborations and the relationships that support them are solidly embedded in a web of social and institutional systems. Resources and individuals are associated with departments, labs, centers, and universities. Journals, conferences, and associations provide networking opportunities and outlets for work within specific disciplines. Corporate sponsors, government agencies, and private foundations provide resources and collect data about research activities. These overlapping institutions each have their own practices, procedures, formats, and systems for managing data, all of which place demands on researchers and affect efforts to form collaborations. To be effective, RNSs must account not only for the needs of the individual users, but also for the nature of the larger social and institutional contexts in which researchers work and live.

Claim 9. RNSs must integrate information from multiple systems, make use of meta-information such as indexing terms to synthesize the information, and present results in a cohesive manner.

Researchers produce many artifacts, including papers, abstracts, presentations, grant applications, Web pages, Internet postings, tools, methods, and datasets. These artifacts are stored in a variety of personal, local, regional, national, or global systems. Representing a researcher's work comprehensively requires information from many different sources. For instance, information about a paper may reside on the author's computer, an electronic journal Web site, and in MEDLINE, CiteSeer, and the Web of Science. Integrating data from heterogeneous sources is a significant challenge because few systems are designed to support machine-based information access or exchange.

RNSs must merge data about a person from several sources in the absence of a common identifier. One common, if mundane, example is retrieving an author's publications unambiguously from MEDLINE [Bedrick and Sittig 2008; McKibbon et al. 2002]. Queries for authors with common names result in many false positives which require additional processing or manual review. Similar problems on the Web have led to the emergence of semantic Web standards for data interchange and interoperation such as SIOC (Semantically Interlinked Online Communities) and FOAF (Friend-of-a-Friend) [Bojars et al. 2008].

Once documents about a person have been retrieved, their content must be meaningfully integrated. Many domains lack the strong tradition of indexing information using controlled vocabularies that the National Library of Medicine has established in biomedicine [Coletti and Bleich 2001]. Therefore, documents may be indexed using different controlled terminologies/ontologies or not at all. Various approaches have been proposed to solve this problem. Liu et al. [2005] proposed the Resource Description Framework (RDF) that combines semantically rich information with a domain ontology to facilitate integration. Cameron et al. [2007] showed how semantic annotation and FOAF can be used to determine the expertise of researchers across various areas of computer science. Jung et al.'s research [2007] discussed a method for finding topic-centric experts from open-access metadata and full text documents using OntoFrame, a semantic Web-based academic research information service. Other approaches to integrating information from multiple sources include ontology-based integration methods [Wache et al. 2001], Digital Object Identifiers (http://www.doi.org), and persistent URL mechanisms (http://purl.org), MOMIS (Mediator envirOnment for Multiple Information Sources), a model of information integration based on the conceptual schema or metadata of the information sources [Bergamaschi et al. 1999], and automated approaches to unifying heterogeneous information based on machine-processable metadata specifications [Singh 1998]. While these methods may be useful in particular contexts, the integration of large-scale classification systems and ontologies and, therefore, the information indexed by them, remains a fundamentally difficult problem [Prévot et al. 2005].

Being able to aggregate a scientist's information artifacts does not mean that they can be easily synthesized into a comprehensive and coherent whole. The process is hindered because documents differ with respect to currency, validity, representation scheme, level of abstraction, audience, and focus. For instance, a list of recently published abstracts may be relatively current in representing a researcher's interests. Nonetheless, it might not be valid if the researcher has abandoned some of the projects. Similarly, recent grants, abstracts, and papers drawn from a departmental Web site will only be useful as a source of current research interests if they can be correlated with the keyword terms that the individual has provided to describe his interests in other systems.

In addition to the technical problems of integration, RNS developers must also consider and address the social and organizational consequences of integration. Researchers are very conscious of the role that their work plays in the formation of their professional identity and reputation (see *Claim 7*). As a result, a composite profile drawing on data from multiple sources that is not under the control of the individual being profiled may create concern. This is further complicated if the technology incorporates information from systems that focus on informal or personal networking, such as Facebook and YouTube [Bateman et al. 2011]. Balancing different perspectives of various information sources is critical if an RNSs are to be effective catalysts for collaborative partnerships.

This discussion suggests the following research questions about integration challenges faced by RNSs.

—RNS that generate comprehensive profiles must acquire and integrate information from heterogeneous sources, such as CVs, MEDLINE, the NIH's Reporter database, conference proceeding sites, online communities, and Web pages. How should RNSs interface with these sources and aggregate data about researchers?
—How should different information artifacts about a researcher be synthesized? What attributes, such as currency, validity, representation scheme, level of abstraction, audience, and focus should be taken into account when creating comprehensive profiles?
—Should data about researchers be managed in a central repository or using a federated approach, in which data are retrieved and synthesized on-the-fly? What issues and problems arise in managing data using either approach?
—How should information content annotated with different types of meta-information, such as controlled vocabularies and social tags, be synthesized? How should information artifacts without meta-information be handled?
—How does combining information from different spheres (e.g., personal and professional) affect the impressions that people form of one another?

Claim 10. RNS must integrate seamlessly with an individual's workflow and the software applications that are part of it.

The scientific workflow in biomedical research and the software applications associated with it are a complex and challenging environment with which RNSs must be integrated. Researchers use a variety of tools, such as data management applications; general office applications, such as Microsoft Word and PowerPoint; reference databases, such as EndNote and CiteULike; conference and journal submission sites; and computer-supported cooperative work applications. Introducing RNSs that duplicate data entry, management, and reporting functions places unnecessarily burdens users and is likely to be met with resistance. Therefore, close integration with researchers' existing workflows and practices is a key factor in facilitating the adoption of RNSs [Schleyer et al. 2008a].

In addition, RNSs must operate across organizational and disciplinary boundaries to be effective. Given the increasingly inter- and multidisciplinary nature of research, a researcher with several research interests is likely to join different communities that are independent, isolated, and supported by incompatible systems. The ability to easily bridge these systems is an essential part of facilitating cross-boundary collaborations. One attempt to solve this problem was introduced by Mitchell-Wong et al. [2007] in the OpenSocial framework. The DIRECT[7] project has also begun to interlink several major current research networking systems.

Research networking is simultaneously critical and secondary. Failure to collaborate undermines a researcher's ability to complete many of the activities critical to successful scientific work. Hence, research networking activities are pervasive and important. At the same time, researchers do not develop collaborations for their own sake. In this sense, research networking is a secondary support activity. Successful RNS must balance these two concerns by supporting lightweight, low-impact integration between the networking system and the systems that are the primary tools of research. This suggests the following research questions regarding integration of RNS, other networking systems, and research workflow systems.

- How should RNSs interface with each other and related systems, such as general social networking platforms? What standards for information exchange should be developed?
- Researchers' activities continuously produce artifacts and information that may be useful in RNS profiles. How can workflows for activities such as conducting experiments or writing a paper be leveraged to facilitate RNS profile maintenance?
- How should RNSs integrate with other systems that researchers use in their work, both from a back-end and user interface perspective? For instance, RNSs could automatically populate an individual citation library in CiteULike or feed an expertise database for paper reviews.
- How can RNSs help address the problem of duplicate information management requirements? For instance, academic and funding institutions require a variety of documents, performance reviews, and progress reports. How should RNS data be structured to facilitate sharing and reuse in other systems?

Research networking is an activity inherently tied to the institutional and social context. Researchers' efforts to form and maintain collaborations are directly affected by the practices and systems around them. Successful RNSs must work with these existing systems, interconnected where the integration provides value and deliberately separate where they are able to improve on the existing capabilities. Hence, designing RNS architectures to allow for various forms of integration is essential to their ability to facilitate the formation of collaborations.

3.4. Evaluation

RNSs require buy-in from a range of stakeholders. Researchers must use the system, both maintaining their profile and searching for others. Administrators must provide

[7] http://www.direct2experts.org

the resources needed to implement RNSs, and support their integration with the systems and procedures of the local institutions. Each of these groups has different needs which may only be partially addressed by RNSs. Making the case for an RNS requires answering a range of fundamental questions about how it provides value for individuals, relationships, and organizations.

Claim 11. Evaluating RNS search results requires metrics which combine traditional information retrieval measures with those specific to collaboration.

Supporting collaboration seeking with an RNS requires that designers define criteria used to select candidates from the pool of available individuals. Although researchers often feel that selecting collaborators is idiosyncratic, context-specific, or even random, the capability to systematically evaluate individual profiles is critical in RNSs.

Evaluating RNSs for collaborator discovery in some ways parallels evaluating Information-Seeking Support Systems (ISSS) for Information Retrieval (IR). Models of information-seeking which can inform RNS design and evaluation include the five-stage information seeking process model [Cole 1997], the Information Seeking Process [Kuhlthau 1991], and the model of general information behavior [Wilson 1999]. To evaluate ISSSs, Kelly et al. [2009] advocate the development of alternative user and task models, methods for assessing support of complex, evolving tasks, and longitudinal designs. As systems providing essential information to researchers to help them make decisions on potential collaborators, RNSs can be considered a type of ISSS. This suggests a need for RNS research which extends IR models to integrate models of the collaboration seeking processes, adds new evaluation methods and measures, and develops longitudinal designs with process-specific measures of learning, cognition, and engagement.

While traditional IR approaches provide a starting point for the social, relational, and instrumental aspects of collaborator discovery, critical differences between person discovery and document retrieval suggest that effective evaluation of RNSs will require fundamentally different approaches. One approach is to consider various frameworks for describing collaboration. For example, Larson [2003] identified three key components of collaboration: structure, process, and outcomes. Structure includes characteristics such as standardized methods of communicating, decision-making, and formal agreements for sharing data and other collaborative activities. Process is characterized by clear and explicit shared research goals and objectives, experience with the change process, strong and clear leadership, and efficient work procedures. Outcomes include measurable work products such as publications, dissertations, and presentations. Another framework more directly related to RNSs is the work of Kraut et al. [1987].

Another critical aspect of RNS functionality is candidate ranking. In general, expertise location systems do not distinguish levels of expertise. Zhang et al. [2007] nonetheless have proposed an expertise-finding mechanism that can automatically infer expertise level from characteristics of postings in an online community. As a result, potential collaborators might be personalized to a candidate's expertise level as well as to keyword similarity.

An overarching issue regarding searching in RNSs is what metrics should be used to assess the quality of the search. Measures typically used in information retrieval include recall and precision, but they require a gold standard against which they can be calculated. While it may be possible to identify a gold standard for RNS searches under narrowly scoped circumstances, such scenarios are not likely to fully reflect the range of concerns involved in forming collaborations.

RNSs depend on criteria for systematically evaluating and ranking potential collaborators to a user. As a result, the following research questions regarding candidate evaluation are central to the development of effective RNSs.

—What model(s) of collaboration and information seeking are most appropriate and relevant to the evaluation of RNS results?
—How should similarity and complementarity be incorporated into the metrics used by RNSs to evaluate potential collaborators? When should the similarity of two people be highly weighted? When should complementarity be emphasized?
—What metrics are appropriate for assessing the outcome of a search for a collaborator using an RNS? Under what circumstances can IR metrics such as recall and precision be used?
—How can process model(s) of collaboration formation inform the design of RNS evaluation metrics? For example, if we use Kraut et al.'s [1987] framework, potential questions include: What specific tasks are involved in forming a collaborative relationship? What strategies and tools do researchers use to complete each task? How does an RNS support the completion of these tasks?

Claim 12. Evaluation of RNSs must assess actual and perceived effects on individual users' collaboration practices and outcomes.

In addition to evaluating the quality of potential collaborators identified by RNSs, it is necessary to assess the general effects of RNSs use on individual users. Such effects could include how individuals' perceptions of RNS functionality and performance develop, and how these perceptions affect users' decisions to participate as collaboration seekers, potential collaborators, or both.

Unlike traditional CSCW applications which focus on performance of tasks by members of well-defined teams, RNSs focus on facilitating a general class of social practices within a diverse, poorly defined community [Neale et al. 2004]. While the general goal of RNSs is relatively clear, the particulars of how the goal is achieved, who is involved, when it is successfully achieved, and what constitutes successful use of the system are difficult to articulate. As a result, assessing RNS performance is highly complex, having more in common with evaluating medical decision support systems [Friedman et al. 2006] than with evaluating traditional process-oriented applications. As with decision support systems, the evaluation of RNS faces challenges arising from crossing multiple research disciplines. As a result, to be useful for design improvement, assessment of RNSs must take into account a plethora of factors. Functional usability and perceived ease-of-use are likely important, but so too are questions of whether the system significantly impacts a researcher at various stages of a collaboration process, as well as long-term career advancement, research directions, and scientific impact.

While the primary goal of RNSs is to facilitate the formation of productive collaboration relationships, the outcome of these relationships is dependent on many other factors, including standardized communication modes, a highly efficient work process, and strong and clear leadership [Larson 2003]. Given the difficulty of delineating the functional boundary between forming collaborations, maintaining the resulting relationships, and executing collaborative work tasks, it is impossible to evaluate the impact of RNSs in isolation. Therefore, it is important to define and assess variables at the various stages of collaboration that RNS may significantly impact.

Another consequence of the complexity of the collaboration formation process is that individual users will rarely have extensive, objective measures of systems performance on which to base their adoption and participation decisions. The presence of potentially conflicting user roles, that is, collaboration seeker and potential collaborator, means that past experience with the system may not be a clear indicator of future effort or outcomes. The extended timeframe of collaborative relationships and the presence of confounding factors also significantly limit an individual's ability to accurately assess the correlation between use of a particular RNS and successful formation of a collaborative relationship. Consequently, user perceptions of system characteristics and impacts are likely to play a significant role in adoption decisions regardless of whether they are based on objective data or not. This suggests that the following questions regarding user perceptions and system assessment will be central to efforts to develop meaningful evaluations of RNSs.

—What is a good collaboration decision? What are near-, medium- and long-term outcomes variables? Are individuals' perceptions of desirable collaborative relationships consistent with those found in empirical studies [Cummings and Kiesler 2008]?
—How do individual users determine if an RNS is useful? What forms of evidence do they use to assess whether a networking system has significantly contributed to their efforts to form and maintain a collaborative relationship?
—What indicators do users rely on to assess whether an RNS has enough participants to be worthwhile as a source of potential collaborators (i.e., critical mass)? How do users determine whether it is beneficial for them to maintain their profile in an RNS?
—How do individuals assess the costs and benefits of using an RNS? What prior experiences provide the basis for expected costs and benefits? What features and outcomes are most salient in development of users' overall assessment of the system?

Claim 13. Evaluation of RNSs must assess their impact on organizational and societal outcomes.

RNSs are infrastructure systems that can only prove their value through the effects they have on their users, the community and/or organization, and the scientific field(s) in which they are used. This raises question of who should invest in these systems and who will derive value from this investment.

The decision makers with the authority to allocate resources for development and maintenance of an RNS are typically not its target users. As a result, their view of the value and cost of an RNS is rarely the same as, or even consistent with, that of the individual users of the system. Where each user may consider the time and effort to maintain their profile a significant cost, an administrator may only see the cost of additional personnel needed to gather the information from external systems (treating researchers' time as "free"). While a researcher might consider the system useful if it allows her to maintain her general awareness of activities taking place in her social network, a funder may seek more quantifiable outcomes such as cost reduction or increased volume of publications. Therefore, although user perceptions of RNSs are critical for its success, evaluation of the organizational- and societal-level impacts is also necessary for their success as sustainable infrastructure systems.

While RNSs and the associated collaborative relationships can be beneficial for researchers and institutions, they can also be costly. Katz and Martin [1997] describe the money, time, and increased administrative effort required to support cross-institutional collaborations. These costs must also be considered when assessing RNS impacts. Together these issues suggest the following questions regarding larger-scale outcomes of implementing RNSs.

—How can the benefits of RNS deployment be quantified? Will there be significant cost reductions for organizations that implement RNSs or do they just shift work from one part of the organization to another? How can the outcomes of supporting collaboration formation be measured?
—What is the appropriate timeframe for evaluation of RNSs? Is it reasonable to expect impacts of RNS use to be visible in months, years, or decades?
—What is the relationship between RNS use and organizationally significant impact measures? Which outcomes supported by RNSs, such as increased research productivity and innovative projects, are most likely to result in significant cost reductions?

—Under what conditions will introduction of RNSs have the greatest impact? What disciplines, areas, and populations will be most affected by the availability of RNSs?

4. CONCLUSION

Choosing appropriate collaborators in science is important and likely to become more so. As this review has shown, the HCI and CSCW literatures provide important background knowledge and foundational concepts for research on RNSs. Beyond core areas such as expertise location systems and virtual communities, advancing our knowledge of research networking must also draw on knowledge representation, ontologies/controlled terminologies, human-computer interaction, social network formation, social matching, and the semantic Web. Moving RNSs forward requires a broad but integrated research program.

Given the current state of RNS development, a rapid, iterative cycle between foundational research, design, implementation, and evaluation seems desirable. The major funding agencies for biomedical (NIH) and basic science (NSF) research in the U.S. are keenly interested in a rapid reengineering of the research enterprise towards a more collaborative approach [Cummings et al. 2008]. CSCW and HCI are disciplines that can add tremendous value to this transformation.

A primary goal of this article is to stimulate the HCI, CSCW, and related communities to consider studying research networking systems. As such, we view our work as a starting point to motivate a much more expansive discussion of research networking systems, and the pursuit of a broad and comprehensive research agenda.

ACKNOWLEDGMENTS

We appreciate all reviewers' thorough and thoughtful comments and suggestions, Ellen Detlefsen's input, Janine Carlock's copy edits, and Michael Dziabiak's help with formatting and submission.

REFERENCES

Ackerman, M. S. and Palen, L. 1996. The Zephyr help instance: Promoting ongoing activity in a CSCW system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Common Ground (CHI '96)*. Tauber, M. J., Ed., ACM, New York, 268–275.

Ackerman, M. S., Pipek, V., and Wulf, V. Eds. 2003. *Sharing Expertise: Beyond Knowledge Management*. MIT Press, Cambridge, MA.

Adamic, L. and Adar, E. 2005. How to search a social network. *Soc. Netw. 27,* 3, 187–203.

Adams, J. D., Black, G. C., Clemmons, R., and Stephan, P. E. 2002. Patterns of research collaborations in U.S. universities, 1981–1999.

Adler, P. S. and Kwon, S. W. 2002. Social capital: Prospects for a new concept. *Acad. Manage. Rev. 27,* 1, 17–40.

Arzberger, P. and Finholt, T. A. 2002. Data and collaboratories in the biomedical community. Tech. rep. CREW-02-01, Collaboratory for Research on Electronic Work, School of Information, University of Michigan, Ann Arbor, MI.

Axelrod, R. 1984. *The Evolution of Cooperation*. Basic Books, New York.

Bateman, P. J., Pike, J. C., and Butler, B. S. 2011. To disclose or not: Publicness in social networking sites. *Inf. Technol. People 24,* 1, 78–100.

Beaver, D. D. 2001. Reflection on scientific collaboration (and its study): Past, present, and future. *Scientometrics 52,* 3, 365–377.

Becerra-Fernandez, I. 2006. Searching for experts on the web: A review of contemporary expertise locator systems. *ACM Trans. Internet. Technol. 6,* 4, 333–355.

Bedrick, S. D. and Sittig, D. F. 2008. A scientific collaboration tool built on the facebook platform. In *Proceedings of the AMIA Annual Symposium.* 41–45.

Belkhodja, O. and Landry, R. 2007. The Triple-Helix collaboration: Why do researchers collaborate with industry and the government? What are the factors that influence the perceived barriers? *Scientometrics 70,* 2, 301–332.

BELLOTTI, V. 1996. What you don't know can hurt you: Privacy in collaborative computing. In *Proceedings of HCI on People and Computers XI (HCI '96)*. Sasse, M. A, Cunningham, J., and Winder, R. L., Eds. Springer, 241–261.

BERGAMASCHI, S., CASTANO, S., AND VINCINI, M. 1999. Semantic integration of semistructured and structured data sources. *SIGMOD Rec. 28,* 1, 54–59.

BIRNHOLTZ, J. P. 2007. When do researchers collaborate? Toward a model of collaboration propensity. *J. Amer. Soc. Inf. Sci. Tec. 58,* 14, 2226–2239.

BOJARS, U., BRESLIN, J. G., PERISTERAS, V., TUMMARELLO, G., AND DECKER, S. 2008. Interlinking the social web with semantics. *IEEE Intell. Syst. 23,* 3, 29–40.

BÖRNER, K., CONTRACTOR, N., FALK-KRZESINSKI, H. J., FIORE, S. M., HALL, K. L., KEYTON, J., SPRING, B., STOKOLS, D., TROCHIM, W., AND UZZI, B. 2010. A multi-level systems perspective for the science of team science. *Sci Transl. Med 2,* 49, 49cm24.

BOS, N., ZIMMERMAN, A., OLSON, J., YEW, J., YERKIE, J., DAHL, E., AND OLSON, G. 2007. From shared databases to communities of practice: A taxonomy of collaboratories. *J. Comput.-Mediat. Comm. 12,* 2.

BRAUN, T. AND SCHUBERT, A. 2003. A quantitative view on the coming of age of interdisciplinarity in the sciences 1980–1999. *Scientometrics 58,* 1, 183–189.

BUDZIK, J., BRADSHAW, S., FU, X., AND HAMMOND, K. J. 2002. Clustering for opportunistic communication. In *Proceedings of the 11th International Conference on World Wide Web (WWW '02)*. Lassner, D., De Roure, D., and Iyengar, A., eds., ACM, New York, 726–735.

CAMERON, D., ALEMAN-MEZA, B., AND ARPINAR, I. B. 2007. Collecting expertise of researchers for finding relevant experts in a peer-review setting. In *Proceedings of the 1st International ExpertFinder Workshop (EFW)*.

CASCIARO, T. AND LOBO, M. S. 2005. Competent jerks lovable fools, and the formation of social networks. *Harvard Bus. Rev. 83,* 6, 92–99.

CASCIARO, T. AND LOBO, M. S. 2008. When competence is irrelevant: The role of interpersonal affect in task-related ties. *Admin. Sci. Quart. 53,* 4, 655–684.

COLE, C. 1997. Information as process: The difference between corroborating evidence and "information" in humanistic research domains. *Inf. Proces. Manag. 33,* 1, 55–67.

COLETTI, M. H. AND BLEICH, H. L. 2001. Medical subject headings used to search the biomedical literature. *J. Amer. Med. Inf. Assoc. 8,* 4, 317–323.

CUMMINGS, J., FINHOLT, T., FOSTER, I., KESSELMAN, C., AND LAWRENCE, K. A. 2008. *Beyond Being There: A Blueprint for Advancing the Design, Development, and Evaluation of Virtual Organizations*. National Science Foundation, Arlington, VA.

CUMMINGS, J. N. AND KIESLER, S. 2005. Collaborative research across disciplinary and organizational boundaries. *Soc. Stud. Sci. 35,* 5, 703–722.

CUMMINGS, J. N. AND KIESLER, S. 2007. Coordination costs and project outcomes in multi-university collaborations. *Res. Policy 36,* 10, 1620–1634.

CUMMINGS, J. N. AND KIESLER, S. 2008. Who collaborates successfully?: Prior experience reduces collaboration barriers in distributed interdisciplinary research. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW '08)*. Begole, B. and McDonald, D. W., Eds., ACM, New York, 437–446.

DE VRIES, S. AND KOMMERS, P. 2004. Online knowledge communities: Future trends and research issues. *Int. J. Web Based Communities 1,* 1, 115–123.

DEMARTINI, G. 2007. Finding experts using Wikipedia. In *Proceedings of the 2nd International Workshop on Finding Experts on the Web with Semantics (FEWS'07)*. Zhdanova, A. V., Nixon, L. J. B., Mochol, M., and Breslin, J. G., Eds., 33–41.

DIMICCO, J. M. AND MILLEN, D. R. 2007. Identity management: Multiple presentations of self in facebook. In *Proceedings of the International ACM Conference on Supporting Group Work (GROUP '07)*. Gross, T. and Inkpen, K., Eds., ACM, New York, 383–386.

EHRLICH, K., LIN, C., AND GRIFFITHS-FISHER, V. 2007. Searching for experts in the enterprise: Combining text and social network analysis. In *Proceedings of the International ACM Conference on Supporting Group Work (GROUP '07)*. ACM, New York, 117–126.

EHRLICH, K. 2003. Locating expertise: design issues for an expertise locator system. In *Sharing Expertise: Beyond Knowledge Management*. M. S. Ackerman, V. Pipek, and V. Wulf, Eds., MIT Press, Cambridge, MA, 137–158.

ERICKSON, T. AND KELLOGG, W. A. 2003. Knowledge communities: Online Environments for supporting knowledge management and its social context. In *Sharing Expertise: Beyond Knowledge Management*, M. S. Ackerman, V. Pipek, and V. Wulf, Eds., MIT Press, Cambridge, MA, 299–325.

FINHOLT, T. A. AND OLSON, G. M. 1997. From laboratories to collaboratories: A new organizational form for scientific collaboration. *Psychol. Sci. 8,* 1, 28–36.

FLYNN, D. A. 2005. Seeking peer assistance: Use of e-mail to consult weak and latent ties. *Libr. Inf. Sci. Res. 27,* 1, 73–96.

FOGEL, J. AND NEHMAD, E. 2009. Internet social network communities: Risk taking, trust, and privacy concerns. *Comput. Hum. Behav. 25,* 1, 153–160.

FONER, L. 1996. A multi-agent referral system for matchmaking. In *Proceedings of the 1st International Conference on the Practical Application of Intelligent Agents and Multi-Agent Technology (PAAM '96)*. 245–262.

FOX, M. F. AND FAVER, C. A. 1984. Independence and cooperation in research: The motivations and costs of collaboration. *J. High. Educ. 55,* 3, 347–359.

FRIEDMAN, C. P. AND WYATT, J. C. 2006. *Evaluation Methods in Medical Informatics* 2nd Ed. Springer, New York.

FRIEDMAN, P. W., WINNICK, B. L., FRIEDMAN, C. P., AND MICKELSON, P. C. 2000. Development of a MeSH-based index of faculty research interests. In *Proceedings of the AMIA Annual Symposium*. 265–269.

GAGLIO, C. AND WINTER, S. 2009. Entrepreneurial alertness and opportunity identification: Where are we now? In *Understanding the Entrepreneurial Mind: Opening the Black Box*, A. L. Carsrud and M. Brännback, Eds., Springer, 305–325.

GEWIN, V. 2010. Collaboration: Social networking seeks critical mass. *Nature 468,* 993–994.

GITLIN, L. N., LYONS, K. J., AND KOLODNER, E. 1994. A model to build collaborative research or educational teams of health professionals in gerontology. *Educ. Gerontol. 20,* 1, 15–34.

GOFFMAN, E. 1959. *The Presentation of Self in Everyday Life*. Doubleday.

GRANOVETTER, M. S. 1973. The strength of weak ties. *Amer. J. Sci. 78,* 6, 1360–1380.

GROSS, R., ACQUISTI, A., AND HEINZ, H. J. 2005. Information revelation and privacy in online social networks. In *Proceedings of the ACM Workshop on Privacy in the Electronic Society (WPES '05)*. ACM Press, New York, 71–80.

GUY, M. AND TONKIN, E. 2006. Folksonomies: Tidying up tags? *D-Lib Mag. 12,* 1.

HEWITT, A. AND FORTE, A. 2006. Crossing boundaries: identity management and student/faculty relationships on the Facebook. In *Proceedings of the 20th Anniversary Conference on Computer Supported Cooperative Work (CSCW'06)*.

HINDS, P. J. AND PFEFFER, J. 2003. Why organizations don't "know what they know:" Cognitive and motivational factors affecting the transfer of expertise. In *Sharing Expertise: Beyond Knowledge Management*. M. S. Ackerman, V. Pipek, and V. Wulf, Eds., MIT Press, Cambridge, MA, 3–26.

JACOVI, M., SOROKA, V., AND UR, S. 2003. Why do we ReachOut?: Functions of a semi-persistent peer support tool. In *Proceedings of the International ACM SIGGROUP Conference on Supporting Group Work (GROUP '03)*. ACM, New York, 161–169.

JENERETTE, C. M., FUNK, M., RUFF, C., GREY, M., ADDERLEY-KELLY, B., AND MCCORKLE, R. 2008. Models of inter-institutional collaboration to build research capacity for reducing health disparities. *Nurs. Outlook 56,* 1, 16–24.

JOHNSON, C. M. 2001. A survey of current research on online communities of practice. *Internet Higher Educ. 4,* 1, 45–60.

JUNG, H., LEE, M., KANG, I. S., LEE, S. W., AND SUNG, W. K. 2007. Finding topic-centric identified experts based on full text analysis. In *Proceedings of the 2nd International Workshop on Finding Experts on the Web with Semantics (FEWS'07)*. Zhdanova, A. V., Nixon, L. J. B., Mochol, M., and Breslin, J. G., Eds., 56–63.

KATZ, J. S. AND MARTIN, B. R. 1997. What is research collaboration? *Res. Policy 26,* 1, 1–18.

KAUTZ, H., SELMAN, B., AND SHAH, M. 1997a. Referral web: Combining social networks and collaborative filtering. *Comm. ACM 40,* 3, 63–65.

KAUTZ, H., SELMAN, B., AND SHAH, M. 1997b. The hidden web. *AI Mag. 18,* 2, 27–36.

KELLY, D., DUMAIS, S., AND PEDERSEN, J. O. 2009. Evaluation challenges and directions for information-seeking support systems. *Computer 42,* 3, 60–66.

KOUZES, R. T., MEYERS, J. D., AND WULF, W. A. 1996. Collaboratories: Doing science on the Internet. *Computer 29,* 8, 40–46.

KRAUT, R. E., GALEGHER, J., AND EGIDO, C. 1987. Relationships and tasks in scientific research collaboration. *Hum.-Comput. Interact. 3,* 1, 31–58.

KUHLTHAU, C. 1991. Inside the information search process: Information seeking from the user's perspective. *J. Amer. Soc. Inf. Sci. 42,* 5, 361–371.

Lakhani, K. R. and von Hippel, E. 2003. How open source software works: "Free" user-to-user assistance. *Res. Policy 32,* 6, 923–943.

Larson, E. L. 2003. Minimizing disincentives for collaborative research. *Nurs. Outlook. 51,* 6, 267–271.

Laudel, G. 2002. What do we measure by co-authorship? *Res. Evaluat. 11,* 1, 3–15.

Leary, M. R. 1996. *Self Presentation: Impression Management and Interpersonal Behavior*. Westview Press, Boulder, CO.

Lee, D. H. and Schleyer, T. K. 2012. Social tagging is no substitute for controlled indexing: A comparison of medical subject headings and Citeulike tags assigned to 231,388 papers. *J. Amer. Soc. Infor. Sci. Technol.* To appear.

LeGris, J., Weir, R., Browne, G., Gafni, A., Stewart, L., and Easton, S. 2000. Developing a model of collaborative research: The complexities and challenges of implementation. *Int. J. Nurs. Stud. 37,* 1, 65–79.

Li, J., Tang, J., Zhang, J., Luo, Q., Liu, Y., and Hong, M. 2007. EOS: Expertise oriented search using social networks. In *Proceedings of the 16th International Conference on World Wide Web (WWW '07)*. ACM, New York, 1271–1272.

Liu, P., Curson, J., and Dew, P. 2005. Use of RDF for expertise matching within academia. *Knowl. Inf. Syst. 8,* 1, 103–130.

Madanmohan, T. R. and Navelkar, S. 2004. Roles and knowledge management in online technology communities: An ethnography study. *Int. J. Web Based Communities 1,* 1, 71–89.

Mann, M. D. 2007, May 7. Some job hunters are what they post. *Nat. Law J*.

Maron, M., Curry, S., and Thompson, P. 1986. An inductive search system: Theory, design, and implementation. *IEEE Trans. Syst. Man. Cyb. 16,* 1, 21–28.

Mattessich, P. W. and Monsey, B. R. 1992. *Collaboration: What Makes it Work. A Review of Research Literature on Factors Influencing Successful Collaboration*. Amherst H. Wilder Foundation, St. Paul, MN.

Mattox, D., Maybury, M. T., and Morey, D. 1999. Enterprise expert and knowledge discovery. In *Proceedings of the 8th International Conference on Human-Computer Interaction*. Bullinger, H. and Ziegler, J., Eds., Lawrence Erlbaum Associates, Mahwah, NJ, 303–307.

McDonald, D. W. 2003. Recommending collaboration with social networks: A comparative evaluation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, 593–600.

McDonald, D. W. and Ackerman, M. S. 1998. Just talk to me: A field study of expertise location. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW '98)*. Poltrock, S. and Grudin, J., Eds. ACM, New York, 315–324.

McDonald, D. W. and Ackerman, M. S. 2000. Expertise recommender: A flexible recommendation system and architecture. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW '00)*. ACM, New York, 231–240.

McKibbon, K. A., Friedman, P. W., and Friedman, C. P. 2002. Use of a MeSH-based index of faculty research interests to identify faculty publications: An IAIMSian study of precision, recall, and data reusability. In *Proceedings of the AMIA Annual Symposium* 514–518.

Melin, G. 2000. Pragmatism and self-organization: Research collaboration on the individual level. *Res. Policy 29,* 1, 31–40.

Millen, D. R., Fontaine, M. A., and Muller, M. J. 2002. Understanding the benefit and costs of communities of practice. *Comm. ACM 45,* 4, 69–73.

Mitchell-Wong, J., Kowalczyk, R., Roshelova, A., Joy, B., and Tsai, H. 2007. OpenSocial: From social networks to social ecosystem. In *Proceedings of the Inaugural IEEE International Conference on Digital Ecosystems and Technologies (IEEE DEST '07)*. Chang, E. and Hussain, F. K., Eds., IEEE Computer Society, 361–366.

Mockus, A. and Herbsleb, J. D. 2002. Expertise browser: A quantitative approach to identifying expertise. In *Proceedings of the 24th International Conference on Software Engineering (ICSE '02)*. ACM, New York, 503–512.

Moore, D. A., Kurtzberg, T. R., Thompson, L. L., and Morris, M. W. 1999. Long and short routes to success in electronically mediated negotiations: Group affiliations and good vibrations. *Organ. Behav. Hum. Dec. 77,* 1, 22–43.

National Center for Research Resources, National Institutes of Health, and Department of Health and Human Services. 2009. Recovery Act 2009 limited competition: Enabling national networking of scientists and resource discovery (U24). http://grants.nih.gov/grants/guide/rfa-files/RFA-RR-09-009.html

Neale, D. C., Carroll, J. M., and Rosson, M. B. 2004. Evaluating computer-supported cooperative work: Models and frameworks. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW '04)*. ACM, New York, 112–121.

NUMPRASERTCHAI, S. AND IGEL, B. 2005. Managing knowledge through collaboration: Multiple case studies of managing research in university laboratories in Thailand. *Technovation 25,* 10, 1173–1182.

OGATA, H., YANO, Y., FURUGORI, N., AND JIN, Q. 2001. Computer supported social networking for augmenting cooperation. *Comp. Support. Comp. W. 10,* 2, 189–209.

OLSON, G. M., ZIMMERMAN, A., AND BOS, N. 2008a. Introduction. In *Scientific Collaboration on the Internet*, G. M. Olson, N. Bos, and A. Zimmerman, Eds., MIT Press, Cambridge, MA, 1–12.

OLSON, G. M., ZIMMERMAN, A., AND BOS, N. 2008b. *Scientific Collaboration on the Internet*. MIT Press, Cambridge, MA.

PAVLOV, M. AND ICHISE, R. 2007. Finding experts by link prediction in co-authorship networks. In *Proceedings of the 2nd International Workshop on Finding Experts on the Web with Semantics (FEWS'07)*. Zhdanova, A. V., Nixon, L. J. B., Mochol, M., and Breslin, J. G., Eds., 42–55.

PETERS, I. AND STOCK, W. G. 2007. Folksonomy and information retrieval. *Proc. Amer. Soc. Inf. Sci. Technol. 44,* 1, 1–28.

PRÉVOT, L., BORGO, S., AND OTLRAMARI, A. 2005. Interfacing ontologies and lexical resources. In *Proceedings of Ontologies and Lexical Resources*. Asian Federation of Natural Language Processing 91–102.

RHOTEN, D. 2007, September 7. The dawn of networked science. *The Chronicle,* B12.

ROSENBLUM, D. 2007. What anyone can know: The privacy risks of social networking. *IEEE Secur. Priv. 5,* 3, 40–49.

SCHLENKER, B. R. 2003. Self-Presentation. In *Handbook of Self and Identity*, M. R. Leary and J. P. Tangney, Eds., The Guilford Press, New York, 492–519.

SCHLENKER, B. R. AND LEARY, M. R. 1982. Audiences' reactions to self-enhancing, self-denigrating, and accurate self-presentations. *J. Exp. Soc. Psychol. 18,* 1, 89–104.

SCHLEYER, T., SPALLEK, H., BUTLER, B. S., SUBRAMANIAN, S., WEISS, D., POYTHRESS, M. L., RATTANATHIKUN, P., AND MUELLER, G. 2008a. Facebook for scientists: Requirements and services for optimizing how scientific collaborations are established. *J. Med. Internet. Res. 10,* 3, e24.

SCHLEYER, T., SPALLEK, H., BUTLER, B. S., SUBRAMANIAN, S., WEISS, D., POYTHRESS, M. L., RATTANATHIKUN, P., AND MUELLER, G. 2008b. Requirements for expertise location systems in biomedical science and the semantic Web. In *Proceedings of the 3rd Expert Finder Workshop on Personal Identification and Collaboration: Knowledge Mediation and Extraction (PICKME'08)*. Mochol, M., Zhdanova, A. V., Nixon, L., Breslin, J., and Polleres, A., Eds., 31–41.

SINGH, N. 1998. Unifying heterogeneous information models. *Comm. ACM. 41,* 5, 37–44.

SPALLEK, H., SCHLEYER, T., AND BUTLER, B. S. 2008. Good partners are hard to find: The search for and selection of collaborators in the health sciences. In *Proceedings of the 4th IEEE International Conference on eSciene (eScience'08)*. IEEE Computer Society, 462–467.

STOCK, W. G. 2007. *Information Retrieval. Informationen suchen und finden [Information Retrieval. Searching and Finding Information].* Oldenbourg, Munich, German.

STREETER, L. A. AND LOCHBAUM, K. E. 1988. Who knows: A system based on automatic representation of semantic structure. In *Proceedings of RIAO'88*. 380–388.

SUCHMAN, L. A. AND TRIGG, R. H. 1986. A framework for studying research collaboration. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work (CSCW '86)*. ACM, New York, 221–228.

TERRY, M., MYNATT, E. D., RYALL, K., AND LEIGH, D. 2002. Social net: Using patterns of physical proximity over time to infer shared interests. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems*. Terveen, L. and Wixon, D., Eds., ACM, New York, 816–817.

TERVEEN, L., HILL, W., AMENTO, B., MCDONALD, D., AND CRETE, J. 1997. PHOAKS: A system for sharing recommendations. *Comm. ACM 40,* 3, 59–62.

TERVEEN, L. AND MCDONALD, D. W. 2005. Social matching: A framework and research agenda. *ACM Trans. Comput.-Hum. Int. 12,* 3, 404–434.

WACHE, H., VOGELE, T., VISSER, U., STUCKENSCHMIDT, H., SCHUSTER, G., NEUMANN, H., AND HÜBNER, S. 2001. Ontology-based integration of information: A survey of existing approaches. In *Proceedings of the International Workshop on Ontologies and Information Sharing.* 108–117.

WALSH, J. P. AND MALONEY, N. G. 2007. Collaboration structure, communication media, and problems in scientific work teams. *J. Comput.-Mediat. Comm. 12,* 2.

WENG, C., GALLAGHER, D., BALES, M. E., BAKKEN, S., AND GINSBERG, H. 2008. Understanding interdisciplinary health sciences collaborations: A campus-wide survey of obesity experts. In *Proceedings of the AMIA Annual Symposium*. 798–802.

WILSON, T. D. 1999. Models in information behavior research. *J. Doc. 55,* 3, 249–270.

WOOLWINE, D., FERGUSON, M., JOLY, E., PICKUP, D., AND UDMA, C. M. 2011. Folksonomies, social tagging and scholarly articles. *Can. J. Inf. Lib. Sci. 35,* 1, 77–92.

YANG, B. AND GARCIA-MOLINA, H. 2002. Improving search in peer-to-peer networks. In *Proceedings of the 22nd International Conference on Distributed Computing Systems (ICDCS'02)*. Rodrigues, L. E. T., Raynal, M., and Chen, W. S. E., Eds., IEEE Computer Society, 5–14.

YANG, S. J. H. AND CHEN, I. Y. L. 2008. A social network-based system for supporting interactive collaboration in knowledge sharing over peer-to-peer network. *Int. J. Hum.-Comput. St. 66,* 1, 36–50.

YU, B. AND SINGH, M. P. 2003. Searching social networks. In *Proceedings of the 2nd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS '03)*. Rosenschein, J. S., Wooldridge, M., Sandholm, T., and Yokoo, M., Eds., ACM, New York, 65–72.

ZERHOUNI, E. 2003. Medicine. The NIH roadmap. *Sci. 302,* 5642, 63–72.

ZHANG, J. AND ACKERMAN, M. S. 2005. Searching for expertise in social networks: A simulation of potential strategies. In *Proceedings of the International ACM SIGGROUP Conference on Supporting Group Work (GROUP '05)*. ACM, New York, 71–80.

ZHANG, J., ACKERMAN, M. S., ADAMIC, L., AND NAM, K. K. 2007. QuME: A mechanism to support expertise finding in online help-seeking communities. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology (UIST '07)*. ACM, New York, 111–114.

ZHENG, J., VEINOTT, E., BOS, N., OLSON, J. S., AND OLSON, G. M. 2002. Trust without touch: Jumpstarting long-distance trust with initial social activities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Changing Our World, Changing Ourselves (CHI '02)*. Wixon, D., ed. ACM, New York, 141–146.

ZIMAN, J. M. 1994. *Prometheus Bound: Science in a Dynamic Steady State*. Cambridge University Press, Cambridge, UK.

Received July 2009; revised April 2011; accepted June 2011