Successful Collaborations:

Helping biomedicine and computation play well together

BY KATHARINE MILLER

Social scientists who study science have noticed a trend: More and more researchers are collaborating. Over the last twenty years, the number of co-authored papers has increased in every scientific discipline and across diverse geographic areas. Co-authored papers are also cited more frequently than single-authored papers, according to what are called "bibliometric" studies.

And many collaborations bridge disciplines. Biomedical computation—interdisciplinary by nature—is no exception. Many of its goals require the involvement of people with different expertise.

So if collaborations will be a fact of life for many involved in biomedical computing, what can be done to make them productive? Can social scientists provide any insights?

Skepticism abounds about whether social scientists' observations of scientists are more informative than scientists' own experience. The ingredients of a successful collaboration seem obvious: good leadership, trust among the participants, face-to-face meetings and strong communication skills.

But then why do many initial collaborations fail? Studies show that even when collaborators are in the same location (a best case scenario), fewer than a third of collaborations succeed, says **Gary Olson, PhD**, a professor of human computer interaction at the University of Michigan who has studied collaborative science. So if common sense can only take us so far, perhaps rigorous research is needed to fill in the gaps. Sociological research can produce counterintuitive findings; answer debates about contrasting models of collaboration; and provide specifics about what works and what doesn't work. "We have an idea about what factors matter, how they matter, and how to intervene to make a collaboration work," Olson says.

Judy Olson, PhD, also a professor of human computer interaction at the University of Michigan, has developed a "Theory of Remote Scientific Collaboration." The "TORSC," as it's known in social science circles, describes a number of factors that can affect the success of collaborations. As suggested by the word "remote" in the theory's name, the most important factor is distance itself. In addition, collaboration readiness, technical readiness, modularity of tasks, and a management plan can make a huge difference—as the leaders of various collaborations attest below.

Some believe that changes in the next generation's social world—which is so reliant on computer interaction—will alter the collaborative landscape. If so, it will provide plenty of fodder for further sociological study. But for now, using guidelines based on the TORSC factors seems wise.

PROXIMITY MATTERS

Thirty meters. That's the rule of thumb. When coworkers are located more than 30 meters from one another, a collaboration's effectiveness declines precipitously, according to the "Allen Curve" discovered by Massachusetts Institute of Technology researcher **Thomas Allen** in the 1970s. Accidental meetings in the hall, water-cooler conversations, lunchroom chats—all of these unplanned encounters between collaborators drop off beyond that distance. And that matters a great deal, says **Jonathon Cummings, PhD**, associate professor of management at Duke University's Fuqua School of Business. It becomes harder to foster a collegial social environment, build common ground, maintain awareness of what others are doing, attend to the project, and adjust to surprises, research shows.

Collaborators located in different buildings within the same university or institution can compensate for a lack of informal interaction with regular in-person meetings. "But as soon as it's not built into their plans, then the 30-meter rule really does operate," Cummings says.

Furthermore, many collaborations—including biomedical computation collaborations—happen at distances measured in miles rather than meters. They range across differ-

ent states, regions, or countries and connect multiple institutions. And these distances really do get in the way of success, according to many sociological studies.

Research by Cummings and his colleague **Sara Kiesler, PhD**, Hillman Professor of Computer Science and Human-Computer Interaction at Carnegie Mellon University, specifically shows that multi-institutional collaborations are less successful than collaborations within a single institution. They studied nearly 500 collaborations funded under the National Science Foundation's Information Technology Research (ITR) program. The collaborations were relatively small—five to ten principal investigators each. But even at that scale, they found that as the number of universities involved increased, researchers spent less of their funding on practices that foster collaboration. They held fewer meetings, made less effort to divide up responsibilities effectively, and transferred less knowledge (such as the best way to do things) from one part of the organization to another. The result: collaborations involving more institutions actually generated fewer positive outcomes (such as papers, new models, new ideas, new software, spin-off projects, or PhD dissertations for graduate students).

"So in a sense, these distributed projects are shooting themselves in the foot by not investing in the very things that would help them succeed," Cummings says. Cummings believes this happens because of budgetary selfishness. If you're spread across multiple institutions, he says, "you're more likely to spend the money on your own institution's needs—salaries and graduate students—than to spend it on a shared workshop or conference."

This finding raises an important question: as interdisciplinary research collaborations become more and more common, should they be promoted more within universities or between them? In recent years, top universities have invested heavily in bringing multiple disciplines physically together under one roof. The Clark Center at Stanford, the Lewis-Sigler Institute for Integrative Genomics at Princeton, and the Broad Institute of MIT and Harvard, represent a few prominent examples. In contrast, funding agencies have been "looking to get the most bang for the buck" by supporting between-university collaborations, Cummings says. Cummings' research supports the former strategy. If given $100 million to invest in either a Clark Center or a collaboration among multiple institutions, he says he would "without a doubt" build a Clark Center.

But **Mark Ellisman, PhD**, director of the Biomedical Informatics Research Network (BIRN), thinks otherwise. BIRN, launched in 2001, was one of the first large-scale biomedical "collaboratories"—a term that refers to large distributed collaborations that rely heavily on tools of the digital age. BIRN, a National Institutes of Health initiative, consists of 31 research groups at 23 universities around the United States and in England. All are working together on infrastructure development and three projects centered around the imaging of human or mouse brains.

As a result of his experience with BIRN, Ellisman thinks we're coming to the end of the era when universities need to attract the best and brightest to their own faculty. "I can gather the best scientists in the world in a virtual collaboration more quickly, and act to conquer big challenges," he says. "I don't need to find a way to move everybody here."

"I can gather the best scientists in the world in a virtual collaboration more quickly, and act to conquer big challenges," says Mark Ellisman. "I don't need to find a way to move everybody here."

Cummings concedes that it may not be possible to address some complex problems without involving multiple institutions—particularly where there's scarce expertise or scarce equipment. "But I would say those are rare or far less likely than other types of projects."

And Cummings admits that distance isn't everything. As described below, readiness and modularity of tasks can, to some extent, help to overcome the problem of distance. As can taking the lessons of Cummings' research to heart: if you're planning a multi-institutional collaboration, set aside money and time to make it all come together—plan on holding symposia; dividing up tasks effectively, and meeting face-to-face on a regular basis. Don't leave it to chance, or it won't happen—unless your institutions are less than 30 meters apart.

READINESS MATTERS

COLLABORATION READINESS

Collaborations collapse when people don't have the right motivations and experience before they launch—what Gary Olson's group calls "collaboration readiness." "All kinds of projects fail when people try to collaborate because the funding agency said they had to or they think they'll get more money if they collaborate—exogenous reasons that don't really make the collaboration work," Olson says.

For example, in the case of the ITR projects that Cummings studied, few of the researchers involved had worked together before. "The funding agencies were holding a big carrot that says we'll only fund you if you put together a distributed interdisciplinary project," Cummings says. "So you had all these people clamoring to find partners to collaborate with. In my view, that's kind of going about it all wrong. . . . Taxpayer money is being spent on sort of trying out a relationship." In a separate study, Cummings showed that people who have never worked together before are less likely to overcome the barriers of working across institutions.

Ellisman admits that in the early days, BIRN researchers had to overcome a lack of collaboration readiness. "Many collaborators were happy to see extra research dollars but actually doing something beyond the office next door, or having to sit in video teleconferences was a bit more painful. They had to work up to it."

And participants have sometimes been unwilling to share data too soon—a factor in collaboration readiness. But, as Ellisman puts it, "With NIH saying 'thou shalt do things differently,' we got everyone to sign off that if they were going to take money under BIRN, there would be open access."

A lack of common ground can also raise readiness concerns, as BIRN researchers discovered in the Mouse BIRN project: Researchers from different subdisciplines referred to the same location in the brain using different terminology. BIRN solved this problem by building ontologies to establish a common vocabulary and by creating a "Smart Atlas," which allowed data to be placed within a common coordinate system.

TECHNICAL READINESS

The TORSC holds that remote collaborations must also exhibit "technical readiness," meaning the participants must be comfortable with the use of communication tools that make long-distance collaboration easier. "If physicians are working together and the video conferencing tools don't work, they'll just walk away," says Gary Olson. "In the early stages, with new or experimental software, there's a breaking-in period and you could lose an entire collaboration if you don't provide good support."

When BIRN first set about creating tools to standardize and calibrate magnetic resonance imaging (MRI) machines at multiple institutions, they solved the local support problem with "BIRN-in-a-box." They shipped an entire integrated hardware system, preloaded and preconfigured with BIRN software, to each of the institutions involved. This minimized the amount of local technical development required.

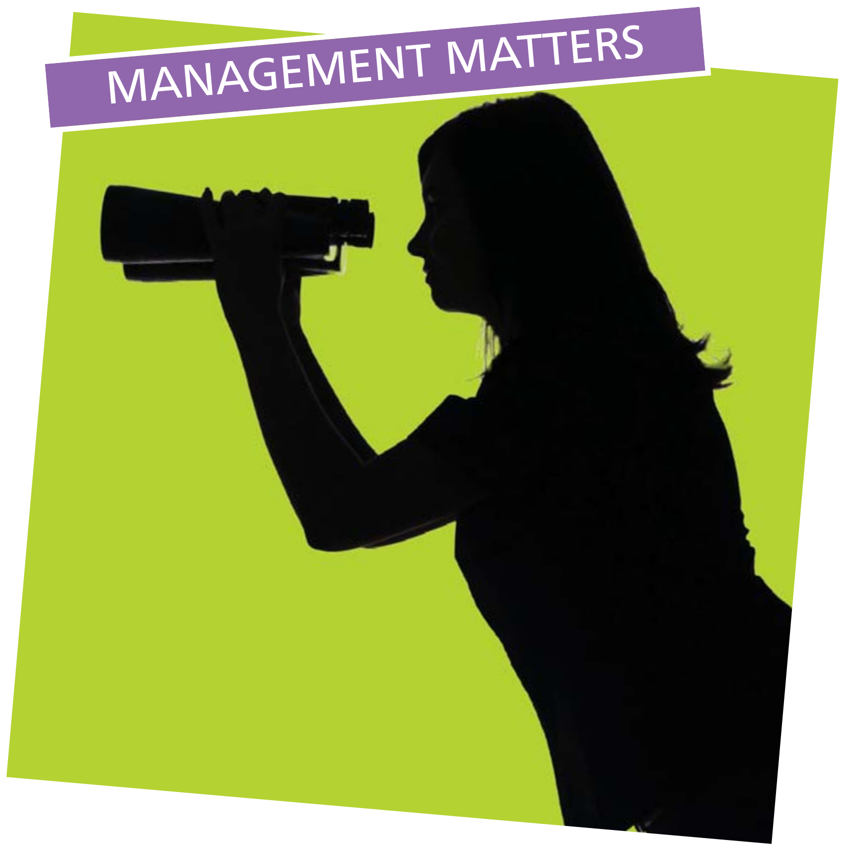
These days, it's BIRN-on-a-disk, which uses virtualization software—with the same goal. That shift to a much simpler technological model is not insignificant. Indeed, Cummings suggests, in the United States these days, the technology needed for collaboration is already robust. "Most scientists are pretty savvy technology-wise. They can use wikis, email, video and tele-conferencing, instant messaging," he says. "In this day and age, it's hard to see that as an issue." But, he notes, international collaborations might present a different picture if issues around broadband and computer power lead to a tilted playing field.

> "In the early stages, with new or experimental software, there's a breaking-in period and you could lose an entire collaboration if you don't provide good support," says Olson.

However, even in this country there are signs that people are only partly technically ready—and that the technology is only partly there. When researchers at the University of Illinois, Urbana-Champaign, created a collaborative, online research environment called BioCoRE, they thought others would be drawn to the interface for securely managing a number of research projects, sharing files, and scheduling supercomputer time. But, says **Kirby Vandivort**, senior research programmer on BioCoRE, "A lot of people saw BioCoRE as only being a Web interface to their normal tools." It was very difficult to convince people of the value added over email and SSH, a program that lets you securely access and submit jobs on remote computers. So although 2500 people have registered for BioCoRE, only about 50 to 100 folks use it regularly—and approximately half of those are located at the University of Illinois, Urbana-Champaign.

Perhaps the most interesting thing about BioCoRE is how it's evolving. The BioCoRE interface is now built right into the research group's most popular software—VMD, which is used for molecular dynamics visualization. So if researchers want to chat or share a molecule with a collaborator, or schedule supercomputer time, they can now do it entirely within VMD. "This is the real jewel that we didn't actually anticipate when we started," says Vandivort. And it suggests that technology is still finding new ways to make collaboration easier.

MANAGEMENT MATTERS



In larger collaborations, especially at a distance, having a management plan is key, according to the TORSC. "The more seriously the scientists take that plan, working out exactly who will do what. . . the more likely the success," writes Judy Olson.

The tricky thing, Gary Olson says, is that scientists and funding agencies are not keen on spending money on management. They'd rather see the money go to the science itself. But some organizations see the need to take management seriously. BIRN has had a management plan from the

"It's easy for the home institution to gobble up the lion's share of the resources with others getting dribs and drabs. But that's going to kill the collaboration if people feel they're not being treated fairly," Olson says.

STARTING OFF RIGHT

A failed collaboration can sometimes lead to legal disputes and misery. To avoid disaster, start your collaboration off right. You might look for Gary and Judy Olson's soon-to-be-released book, *Science on the Internet*. Or, if you want more of a short course, consult "Guidelines for Negotiating Scientific Collaboration," published in *PLoS Biology* in June 2005, or "With All Good Intentions," published in Nature in April 2008, which includes a "collaborators' prenup."

COUNSELLING FOR COLLABORATORS

If things start to go sour, get help. The Olsons and their colleagues at the University of Michigan often provide advice to collaborations that are struggling. In addition, Gary Olson now has a grant to create an online tool to help researchers evaluate their own collaborations using the TORSC. He calls it a "Wizard"—a "Collab-o-matic," if you will. Answer a series of interview questions about your project, and it will give you automatic feedback. The wizard is currently being developed and tested. To Olson, it's a way to share what his group has learned and also to collect enormous amounts of additional data. So keep your eye out for it. You might be surprised what it will tell you. And what you can tell the Olsons.

get go. They created an oversight committee that, in turn, commissions a variety of standing and ad hoc committees. And BIRN also has as one of its cores the BIRN Coordinating Center.

Good management helps ensure fairness, Gary Olson says. "It's easy for the home institution to gobble up the lion's share of the resources with others getting dribs and drabs. But that's going to kill the collaboration if people feel they're not being treated fairly. We've seen lots of examples of that."

Another challenge is getting everyone in a collaboration to treat each other as equals. For example, Ellisman points to various NSF-funded cyberinfrastructure projects, in which tension between computer scientists and biomedical scientists impeded collaboration. The result: The computer scientists built something that was underutilized because biologists hadn't been fully engaged in development.

Ellisman says avoiding this pitfall was his hardest challenge as director of BIRN. "People want to feel recognized for their contribution," he says. For example, because BIRN needed to involve computer scientists working at the cutting-edge of their own field, it was critical that the biomedical researchers realize the computer scientists' contributions were equally important—"so they aren't computer scientists in the service of biomedicine."

"We started out talking about 'if you're a biologist, hug your computer scientist,'" Ellisman says.

Besides fairness, good management also helps build and maintain a collaborative infrastructure. For example, BIRN set out to create a geographically distributed repository of medical images and make them available for large-scale cooperative studies more or less in real time. This involved inventing new software tools to normalize MRIs across multiple institutions; de-identifying health records; and getting approval from the numerous institutional human subject review boards.

"It's hard work and you take a lot of lumps," he says. And the coordinating center for BIRN ends up being viewed as a service provider. "Like if the telephone doesn't work. It's not how great it is, but what's not working today," he says. "It's part of what happens when you build something that becomes like a utility."

MODULARITY MATTERS

According to the TORSC, successful collaborations divide the work so that it can be done without the need for a lot of chit-chat. "We have seen a number of projects fail because tightly coupled work spanned people in different locations," writes Judy Olson and her TORSC colleagues. "The more modularized the work at different locations, the more likely is success."

Indeed, distance may not matter so much when a collaborative task is easily divided into distinct tasks. And Cummings says this should provide a glimmer of hope to biomedical computing collaborations located at multiple universities. According to his research with the ITRs, projects that produced computer hardware, software and datasets seemed to suffer less from being located at multiple universities.

"I believe there's something fundamentally different about tools projects," Cummings says. "Tool development often can be decomposed or broken down, and programmers and other developers seem to have a shared understanding of how to visualize that process." By contrast, Cummings says, other scientists might not have a clear and common understanding of what a successful outcome would look like or what the steps would be to get there. "The nature of the work—the ability to modularize it and divide it up," says Cummings—that's where biomedical computing researchers might have an edge.

The Protein Structure Initiative (PSI) is one prominent collaborative project that, on the surface at least, seemed to divvy up tasks effectively. For example, the PSI funded four large-scale centers in various locations around the country. Each of them then created its own multi-institution pipeline for crystallizing proteins, analyzing their structure, and maintaining and disseminating the data generated. Each institution had a discrete job.

"You have to have independent tasks," says **Ian Wilson, PhD**, professor of molecular biology at the Scripps Research Institute and director of the Joint Center for Structural Genomics, one of the four PSI centers. But structuring the pipeline in a modular fashion isn't quite enough, Wilson says. "It has to be seamless." Making the work flow smoothly took a lot of hard work over the first five years of the project. Some of that involved building relationships among

> "Tool development often can be decomposed or broken down, and programmers and other developers seem to have a shared understanding of how to visualize that process," Cummings says.

people who didn't know each other before—establishing "collaboration readiness" after the fact. "We had meetings every week among the cores to discuss how communication could be better and to determine how to move materials through the pipeline more efficiently." At this point, Wilson says, "We've got a fantastically well-integrated team working on this seamless pipeline."

But just as that happened, a new coordination task emerged: in PSI-2 (the initiative's second five years), the four centers were tasked with working together on communal goals, Wilson says. And the bioinformaticians across the four centers have to work together as well. "We have to decide as a group what to do for 70 percent of the project," he says. So modularity only got them so far. Ultimately, all of the centers have to agree on how to achieve the overall mission of the PSI.

THE FUTURE OF COLLABORATION

WILL WEB 2.0 CHANGE HOW WE COLLABORATE?

Some social scientists suggest that the shift toward collaborative science mirrors changes in the social order: As global networking becomes the modus operandi in all realms, including business and social life, so too will it become a natural part of science. If this is true, then the culture vultures who see Web 2.0 as the wave of the future—with its MySpaces, Facebooks, and Wikipedias—would also predict that this social climate will affect the scientific endeavor. An impact on collaborative science seems almost inevitable.

According to Cummings, this idea may take some time to be realized. "I think the social structure of science is very resistant to change," he says. Web 2.0 tools promoting collaboration won't overcome a culture in which junior faculty are less rewarded for playing a role in a collaboration than they are for producing strong individual research. "You're looking at hiring and promotion of junior faculty, tenure committees, how departments are structured in universities and labs, norms for sharing and being open about information, journals and their restrictions on intellectual property," Cummings says. "All of these factors play a much larger role than people who are optimistic about the technology may realize."

But Ellisman disagrees: "The world people sit at everyday is a collaborative one," he says. In fact, the BIRN portal right now is a customizable, chattable, bloggable, build-your-own environment type of space. And the next generation will move toward standardization. Just as browsers such as Firefox, Explorer and Netscape have all started to look alike, with similar menus, he says, "for these collaboration spaces we're going to see a kind of refinement in what is practical and familiar."

And he sees a gradual shift toward openness as well—a fundamental change in scientific culture. Colleagues Ellisman's age (60), may still cling to their data, he says. "But for my youngest children, it's obvious that in the open, available, electronic age where you can exchange music legally or illegally, everything should be open." Indeed, he says, "We need to err on the side of open access all the way down to the laboratory notebook."

Some scientists are already there. People are blogging about the potential for open lab notebooks at

OneBigLab.Blogspot.com. The Web site myExperiment.org is giving it a go by allowing researchers to upload workflows so that other people can benefit from them. And in April 2008, Scientific American reported on the phenomenon—pointing to the early success of OpenWetWare.org, a wiki created by graduate students at the Massachusetts Institute of Technology in 2005. Protocols posted on that site have become useful to many other scientists around the world.

As for scientists using Web 2.0 to find collaborators, Nature ran a story in February 2008 titled "The New Networking Nexus." It describes the spawning of Web sites geared toward bringing researchers together online to discuss their common interests. Examples include Nature's own site called Nature Network (network.nature.com), which apparently draws interdisciplinary scientists, and Community of Science (COS.com). But, some say, getting a critical mass of participants limits the usefulness of these and other sites.

Right now, so many are attempting to develop Web 2.0 tools for science that it's tough to see where they'll lead. Tools such as the BioCoRE Web site are created with a specific purpose in mind, but they may end up being used in a different way, or perhaps not being used at all.

Ellisman predicts that Wikipedia is the model of the future. "This is what we expect the world to look like—where all mankind's knowledge is available to all from anywhere, anytime. That's what the electronic age we're poking through right now is going to make possible."

If he's right, as data becomes available virtually and search engines become more intelligent, perhaps collaboration among colleagues will become so interwoven with the everyday lives of scientists that it won't even be called collaboration anymore. But don't count on that happening anytime soon.

"The world people sit at everyday is a collaborative one," Ellisman says.

For now, say the social scientists, if you want to succeed, your best bet lies in collaborating with people who work less than thirty meters away and with whom you've worked before. Oh, and don't forget to divvy up the tasks effectively and create a management plan. Though research shows that following these guidelines only gives you a 30 percent chance for a successful collaboration, it is a possibility that can become a reality, even for remote collaborations.

"The biggest reward from running BIRN," Ellisman says, "has been seeing the acceptance by the biomedical research community of the notion of working across the boundaries of institutions and different domains, and of cooperating to take on larger challenges for the benefit of society and for our understanding of biomedicine and the human predicament." □