SARAH MORRISON-SMITH | RESEARCH STATEMENT

I design **systems to support multidisciplinary research teams** that rely on computers, including life science research and remote work. My main research interests lie in computer-supported cooperative work, a sub-field of human-computer interaction (HCI). Research that addresses problems with significant societal impact is increasingly conducted in large, interdisciplinary collaborations. The goal of my work is to develop systems that facilitate research and address the challenges faced by researchers using technology. To do this, I use two approaches; 1) I conduct qualitative research to understand the problems that people have when doing scientific work to inform the design of new technology, and 2) I design, implement, and evaluate systems that solve collaboration challenges. This statement will focus on three of the key challenges my research has centered on.

Challenge 1: Collaboration Activities Are Insufficiently Supported

My preliminary work [1] revealed that the collaboration activities in multidisciplinary, multi-institutional life science research is currently insufficiently supported by existing technology. Collaboration is the keystone of modern research. Without collaboration, problems with high societal impact are difficult if not impossible to solve [7]. To address this issue, I performed a qualitative study to better understand the use of technology in multidisciplinary life science research and the design needs of tools that facilitate large-scale research collaborations. Results from this work revealed that: (1) large, interdisciplinary, and multi-institutional collaborations are unavoidable in life science research, (2) the need for expertise outweighs the difficulties associated with remote collaboration, and lastly, (3) current technology is ineffective for scientific communication. From these findings, I uncovered implications for life science training programs. These implications could inform the design of new systems that support collaborative work by highlighting the need for future technology to support the discussion and negotiation process as methodologies and data sharing standards evolve; provide mechanisms for facilitating, simplifying, and documenting conversations surrounding scientific knowledge in a manner that allows users to search for abstract representations of information; and facilitate explicit management in large projects. This work is under review [4].

Challenge 2: Researchers Have Difficulty Perceiving Remote Teammates' Priorities

I have developed a system addressing one of the challenges identified by the work described above: researchers' perceptions of their collaborators' project involvement. This challenge has since been magnified by the prevalence of remote collaboration due to the COVID-19 pandemic. My first step was to conduct an extensive literature review of the challenges faced by remote workers [3]. Next, I developed *AmbiTeam* (Fig 1), which monitors and provides an ambient visualization of project-related activity for each collaborator in a research project with the goal of reducing uncertainty regarding collaborators' priorities, therefore bolstering collaboration. Ambient displays communicate contextual or background information in the periphery of the user's awareness and only require the user's attention when it is appropriate or desired [6]. By providing information in an ambient form, I minimized disruptions to the user's existing workflow. AmbiTeam was evaluated with 10 scientists from four existing collaborations across four institutions in the United States. I found that using AmbiTeam had a quantifiable effect on researchers' perceptions of their collaborators' project prioritization and that use of the system motivated researchers to work on their collaborative projects. This work has implications for creating effective awareness-based technology for supporting collaborative work. I recommend that future awareness systems consider (a) using file activity to measure effort and (b) implementing ambient displays that do not interrupt the user's workflow. This work appears in Graphics Interface 2021 [2].

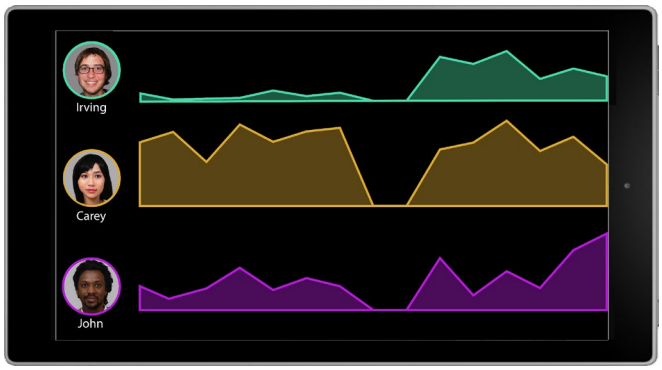

Fig. 1: This example shows activity from three teammates using random data. Like our participants, at the end of the first week, Irving notices that Carey has been working hard on their project. This makes him feel more motivated to work. [5].

Challenge 3: Balancing Data Control and Collaboration

I am currently exploring the use of data-sharing technology in multidisciplinary life science research and how to better support this facet of research. One of my recent studies has revealed that the data sharing needs of life

scientists are not adequately supported by existing technology [1]. A current collaboration with Dr. Jaime Ruiz, Dr. Christina Boucher, and Dr. Aleksandra Sarcevic, as well as a team of Barnard and Columbia undergraduate students that I am mentoring, has begun exploring the challenges and barriers to data sharing within life science collaborations. The initial results have allowed us to begin to identify fundamental issues including that researchers struggle to balance data access control and collaboration. Current technology, such as Google Drive, tends to lack discoverable fine-grained control. Although Google Drive integrates with Google Groups to provide group access to shared documents, access to files is universal across the group without a way to account for situations where some group members (e.g., students) need full access while others (e.g., outside collaborators) should only have view access. While frameworks for fine-grained access control methods have been developed, they lack widespread implementation and are not integrated into Google Drive, which operates solely on an identity-based model [5]. Findings from my recent work shows that researchers frequently find using existing settings in Google Drive to be painful and inefficient, particularly when controlling access for multiple files and folders, due to inappropriate default settings. Furthermore, researchers are often unaware of the access control features that are supported by the technology they use every day. My goal is to develop, implement, and evaluate a system called *DriveGroups* that provides fine-grained control for Google Drive allow users to set workspace or project wide defaults in a transparent, accessible, and discoverable way. DriveGroups will accomplish this by providing an optional UI that clearly indicates individual and group file access levels, allows users to easily set file access settings, and collates information about who has access to what in a centralized location, as opposed to being buried deep in settings. The first step is to utilize ethnographic methods to investigate access control challenges in situ. Next, I will design and implement DriveGroups using Google's AppScript workspace. After implementation, DriveGroups will be evaluated by laboratory-based and in-situ user evaluations with local life science researchers. This research is in the preliminary stage and will be the topic of an upcoming NSF (CISE) Research Initiation Initiative (CRII) grant proposal.

Future Research

My future work will continue to study user interaction for multidisciplinary research to better support the scientific efforts that have important social impact. Two important areas I will address are interruptability when working remotely and data organization when sharing data within collaborations. My current research addressing Challenge 3 has revealed that data organization can hinder sharing efficiency, particularly with regards to the metadata required to make sense of biological data. Together with my findings regarding balancing data control and collaboration, these observations allow me to start developing a foundation of guidelines and principles that inform the design of technology for sharing data.

Future Challenge 1: Data Organization Can Help or Hinder Sharing Efficiency

In the context of life science research, metadata (e.g., sampling methods, analysis method etc.) is used to make decisions regarding reuse of existing content and resources and simplify retrieval and sharing of data [9]. While prior work has focused on developing frameworks for sharing metadata alongside public datasets [8], the organization and preservation of metadata within collaborations, before the data is made public, is still an open challenge. In particular, the data sharing technologies that our current work uncovered as being the primary means of sharing data within a collaboration (e.g., Google Drive, Dropbox) currently lack this functionality. My goal is therefore to develop, implement, and evaluate a system called *MetaDrive* that incorporates metadata into Google Drive's organizational schema. With a team of undergraduate researchers, I will design and implement MetaDrive using Google's AppScript workspace. Like DriveGroups, after implementation, MetaDrive will be evaluated by a series of laboratory-based and in-situ user evaluations with local life science researchers.

Future Challenge 2: Interruptability Is Difficult to Determine in Remote Teams

My prior work from Challenge 2 has shown that in addition to providing opportunities to answer quick questions related to the project, researchers find that spontaneous, informal project inquiries are beneficial when attempting to keep collaborators up-to-date on the project status. This type of informal project inquiry helps researchers show interest in the project—which can alleviate concerns regarding a collaborator's priorities. Unfortunately, this type of communication is only possible if the collaborators are co-located, putting remote collaborations at a disadvantage. To address this problem, I intend to develop, implement, and evaluate a system called *Knock-Knock*

that uses an ambient display to indicate the availability of collaborators as if they were as close as the next office over, facilitating spontaneous communication. Knock-Knock will also feature a microphone and speaker which that will allow users to initiate a conversation by selecting a collaborator on the ambient display and then knocking on their table as if they were knocking on their collaborator's door. The knock is played in the recipient's office. Knock-Knock will be connected to videoconferencing software, such as Zoom, which will allow the recipient to choose to answer by initiating a virtual face-to-face meeting or ignore the knock. Furthermore, since knocking on the table is similar to knocking on a collaborator's door, Knock-Knock affords the same feeling of informality. After implementation, Knock-Knock will be evaluated by a series of laboratory-based and in-situ user evaluations with researcher in existing remote collaborations.

Long Term Research Agenda

The current state of collaboration is plagued by many challenges. In this statement, I have discussed navigating the push/pull dynamic of working with collaborators who are less invested than their team, adhering to institutional data control standards while maintaining the spirit of collaboration, and effectively sharing knowledge with collaborators outside your field of expertise. Collaboration should feel natural. Teammates should be able to seamlessly interact with each other while equitably contributing to the project. More generally, my future research will focus on understanding how to employ new technology to help fundamental collaboration problems and ensure that the systems we design will be used in practice. In the future, I will continue to harness the understanding gained through qualitative research to design, implement, and evaluate systems to support multidisciplinary research teams that rely on computation Furthermore, to support my research activities, I plan to submit grant proposals to various programs in the following agencies including NSF IIS (CHS), NSF I-USE, and the NSF Faculty Early Career Development (CAREER) Program, among others.

Publications

1. Sarah Morrison-Smith, Christina Boucher, Andrea Bunt, and Jaime Ruiz. 2015. Elucidating the role and use of bioinformatics software in life science research. In *Proc. of British HCI '15*, 230–238.
2. Sarah Morrison-Smith, Lydia B Chilton, and Jaime Ruiz. 2021. AmbiTeam: Providing Team Awareness Through Ambient Displays. In *Graphics Interface*.
3. Sarah Morrison-Smith and Jaime Ruiz. 2020. Challenges and barriers in virtual teams: a literature review. *SN Applied Sciences* 2: 1–33.
4. Sarah Morrison-Smith, Aleksandra Sarcevic, Noelle R Noyes, Christina Boucher, and Jaime Ruiz. 2021. Challenges in Large-Scale Bioinformatics Projects. *In submission*.

Other References

5. Li Gong. 1989. A Secure Identity-Based Capability System. In IEEE symposium on security and privacy, 56–63.
6. Jeremy M Heiner, Scott E Hudson, and Kenichiro Tanaka. 1999. The information percolator: ambient information display in a decorative object. In *Proceedings of the 12th annual ACM symposium on User interface software and technology*, 141–148.
7. Gary M Olson and Airong Luo. 2007. Intra-and inter-cultural collaboration in science and engineering. In *International Workshop on Intercultural Collaboration*, 249–259.
8. Anushka Rajesh, Yutong Chang, Malak S Abedalthagafi, Annie Wong-Beringer, Michael I Love, and Serghei Mangul. 2021. *Improving the completeness of public metadata accompanying omics studies*. BioMed Central.
9. Katy Wolstencroft, Stuart Owen, Franco du Preez, Olga Krebs, Wolfgang Mueller, Carole Goble, and Jacky L. Snoep. 2011. Chapter twenty-nine - The SEEK: A Platform for Sharing Data and Models in Systems Biology. In *Methods in Enzymology*, Daniel Jameson, Malkhey Verma and Hans V. Westerhoff (eds.). Academic Press, 629–655.