My research advances the design of integrative human-machine systems that solve complex human problems and advance desired values at scale. Specifically, I design, build, and study *Computational Ecosystems* that interweave community process, social structures, and intelligent systems to unite people and machines to overcome large challenges. My research sits at the intersection of social and crowd computing, human-computer interaction, and artificial intelligence, and applies principles and methods from decision science and learning science to enable specific applications.

Despite the continued development of individual technologies for supporting human endeavors, major leaps in solving complex human problems will require a radical re-thinking of how to combine wedges of human and machine competencies into the design of *integrative* systems [9, 18, 4, 8]. In practice, most socio-technical solutions are built component by component, and layer by layer. Researchers and practitioners focus primarily on advancing individual technologies (i.e., components), and on designing technologies that complement existing ways of working and organizing (i.e., as a layer). While this approach can solve focused problems that are within each technology's scope, it can fail to address the underlying challenges inherent in creating comprehensive solutions to complex, real-world problems that require new, integrative solutions. **Taking an alternative approach, my work on computational ecosystems integrates multiple components, and designs entirely new compositions of processes, social structures, and intelligent systems that work well together to enable unique, scalable solutions to complex, real-world problems.**

Using this approach, I devised scalable solutions to advance community-based problem solving and the learning of complex skills. Specifically, I designed computational ecosystems that (1) facilitate complex planning with many knowledgeable participants; (2) use crowds and mobility as potential resources for solving local problems in ways that balance the costs and benefits; (3) transform professionally-made artifacts into an effective resource for learning complex skills; and (4) support training a large number of independent researchers to cultivate self-directed learners. To overcome the inherent challenges that limit the efficacy and scale of existing approaches, I introduced new ways to integrate multiple sources of human and machine intelligence; foster sustainable communities and promote synergistic interactions among participants; and shift (but not replace) the roles of human participants. In doing so, my contributions extend beyond the specific application areas to broadly advance general principles and methods for designing integrative human-machine systems. In presenting my work, I will highlight conceptual and technical advances in (1) *integrative computing* to decompose, distribute, and connect problem solving across diverse people or machines most able to address them; and (2) *ecological thinking* to create sustainable processes and interactions that support jointly the goals of ecosystem members and proper ecosystem function.

1 Facilitating Community-Based Planning

The first track of my research transforms the way we plan large events. I contribute a *community-informed planning (CIP)* approach that produce outcomes that better serve the community as a whole, significantly ease the load on organizers, and engage many participants to communicate their needs and wants.

Planning community-based events involving thousands of people is a challenging organizational and computational problem. Organizers often lack effective processes and tools for gathering and acting on the diverse preferences, constraints, and domain knowledge held by community members that are critical for forming effective plans. Automated systems can resolve easy-to-encode hard constraints, but the plans they output are often incomplete and undesirable because they cannot account for important community goals that remain latent and tacit. As a consequence, the planning task remains arduous, time-consuming, and short-sighted. As an illustrative example, consider the task of scheduling talks for a large academic conference. Typically, a few dedicated organizers produce the schedule with minimal tool support and little input from community members. Despite significant efforts from organizers, final schedules often contain incoherent sessions, parallel sessions that are similarly-themed, and attendee-specific conflicts.

To address this challenge, I created a computational ecosystem for community-informed planning called Cobi that provides a general model for engaging various stakeholders in the planning process and mixed-initiative tools for planning with community input [13, 2, 7, 3]. As overall project lead, I worked with Rob Miller, Steven Dow, Juho Kim, Lydia Chilton, and others to apply this approach to reinvent how we plan large conferences. Our deployments at ACM CHI and CSCW produced three key outcomes: **(1) organizers produced better schedules by resolving hundreds of previously hidden conflicts; (2) our tools reduced organizers' planning time from 100 hours to 5 hours; and (3) our process engaged 1.5k+ community members across phases of planning.** This research **received a best paper honorable mention at ACM CHI 2014 [7] and a notable paper award at HCOMP 2013 [2]**. Given the success of our deployments, Cobi is now part of ACM CHI's standard process for planning the schedule for its 3000+ attendees.

Advancing ecological thinking, CIP contributes new approaches for engaging diverse participants in ways that benefits them and the planning process as a whole. Instead of treating data collection tasks as isolated problems, I designed a socio-technical process called *incentive chaining* to support participants across multiple phases of planning by using the data collected in earlier interactions to promote further data collection in subsequent interactions. For example, the affinities generated by program committee members produces a short list of similar papers for authors; this list helps authors focus their attention rather than scan through hundreds of accepted papers [2]; their data provide the seeds for our recommendation engine for attendees, avoiding the "cold-start problem" faced by other collaborative filtering systems [3]. In other words, by building systems that considers all aspects of the data collection and the different needs and interests of varied stakeholders across various phases of planning, we can develop solutions that make the process of gathering data more useful for participants while also gathering more useful data for planning.

Advancing integrative computing, CIP contributes new approaches for integrating multiple sources of human and machine intelligence to support planning. Lydia Chilton and I designed Frenzy [7], a collaborative planning tool for making sessions that draw on the expertise of many community members while still allowing for careful deliberation and discussion. To gather diverse paper groupings while also addressing global constraints (e.g., a paper can only be assigned to one session), Frenzy presents different *actionable feedback* at different phases of problem solving by using an automated system to encourage independent crowd work early when constraints are 'loose,' and smaller group work later when constraints are 'tight.' In this way, Frenzy provides a model for flexibly moving between *crowdware* and *groupware* modes of collaboration to support changing needs over the course of a community-based problem solving process.

In order to leverage the wealth of information collected from a CIP process, Juho Kim and I developed *community-informed, mixed-initiative interfaces* that empower organizers to make effective planning decisions using community input, machine intelligence, and their own tacit knowledge. For example, to account for noisy and incomplete community-sourced data, the Cobi scheduling tool [13] empowers organizers to make informed scheduling decisions based on optimizer-suggested moves and their own tacit knowledge.

2 Producing Globally Effective Behaviors for Physical Crowdsourcing

The second track of my research aims to enable *physical crowdsourcing systems* that motivate more people to contribute to solving local, communal problems. To do this, we are designing interactions and mechanisms that engage individuals at convenient moments in which they can make valuable system contributions.

The growth of mobile devices in recent years has helped to bring about *physical crowdsourcing systems* that connect people to tasks and lower barriers to participation. Citizen science and community sensing projects recruit volunteers to track migration patterns and report city problems. Peer economies such as TaskRabbit and Uber coordinate the provision of physical resources on-demand. Despite successes, existing systems are limited by the fact that contributions are gathered via one of two approaches: *opportunistic* or *directed*. Opportunistic approaches receive input from users when and where they decide to contribute. This can lower participation barriers, but has the downside that systems cannot make demands for particular data to be collected or particular tasks to be completed. Directed approaches prompt users for specific input to advance system goals, which often require contributions that are largely outside people's existing routines and thus require strong incentives. As a consequence, existing models must often leave tasks incomplete or rely on monetary incentives to ensure effective system outcomes.

As PI of an NSF Cyber-Human Systems award and a Microsoft FUSE Labs award, I am working with Darren Gergle, Yongsung Kim, and others to advance a hybrid approach to physical crowdsourcing that brings about the best elements of both opportunistic and directed approaches. Our approach develops computational ecosystems that permit access to a large number of people who when directed to make small, convenient contributions through their existing routines, contribute actions that jointly achieve system goals.

In one direction, we are developing interaction techniques that broaden participation by breaking down larger goals into smaller physical tasks that can be completed through people's existing routines. This led to our developing a set of *low-effort sensing* interactions that allow people to report events and provide small pieces of additional information in just seconds [1, 5]. We also considered ways to tap into people's physical routines directly. Through work on *habitsourcing* [16], we developed *sensing through actuation*, a new interaction technique for adding immersion and enjoyment into habit-building apps (e.g., for running) in ways that elicit physical actions from which we can infer environmental data (e.g., the locations and prevalence of certain types of trees). While making inferences from people's motion activities is typically a hard machine learning problem, with our approach the problem is made much simpler because we cue users to take specific actions within pre-specified timeframes. To support prototyping and testing mobile app interactions in realistic use cases, I used an **NSF CISE Research Initiation Initiative award** to develop methods and tools for building and testing mobile app prototypes in the wild [6].

In a complementary direction, our recent and ongoing work develops new decision-theoretic frameworks for deciding how to opportunistically recruit a well-motivated and minimally disrupted crowd to efficiently achieve system goals. Specifically, we are developing (1) a *4X* framework for scaffolding data collection through phases that continually expand data fidelity and spatial coverage; (2) a *hit-or-wait* [14] framework for deciding whether to notify helpers of tasks on their route by reasoning jointly about system needs and a helper's changing patterns of mobility; and (3) a *opportunistic supply management* framework for optimizing and adjusting community-level policies for recruiting helpers; (4) a *continual crowdsourcing* framework that monitors, identifies, and connects problem-solving needs to available resources. **By simultaneously reasoning about both system needs and the goals, interests, and availability of potential helpers, these decision-theoretic frameworks enable a flexible approach for connecting helpers to tasks that allows potential helpers to contribute opportunistically when they desire, in situations when it is convenient to them, valued by the system, and that ultimately lead to globally effective solutions.** More generally, these frameworks advance ecological thinking by offering principled approaches for managing contributions across a computational ecosystem, in ways that support ecosystem members and proper ecosystem function.

3 Creating Readily Available Learning Experiences on the Web

While the last two sections focused on designing computational ecosystems to advance and scale solutions for community-based problem solving, these next two sections focus on the learning of complex skills. The third track of my research aims to make authentic learning experiences broadly available so that many learners can gain the skills needed to perform complex, professional work. I discuss our efforts to transform professionally-made artifacts into an authentic learning resource in the domain of web development.

Aspiring web developers are turning to online resources to teach themselves to code, but existing learning platforms primarily teach syntax or provide practice on constrained tutorial examples. With few resources to support the progression from writing functional code to writing professional-quality software, we consider the use of professional websites as a potential resource for authentic learning. Professional websites offer rich details missing from training examples, embed the programming concepts and implementation techniques used by professionals, and are continually updated as new solutions arise. But despite the abundant availability of front-end code, professional examples are difficult for learners to understand.

As PI of an NSF Cyberlearning grant, I am working with Eleanor O'Rourke, Josh Hibschman, and others to advance the creation of Readily Available Learning Experiences (RALE) that transform the entire web of professional examples into a resource for learning programming concepts and design patterns. To achieve these goals, we are designing a learning ecosystem that embeds software scaffolds that support self-directed learning, so that learners can more effectively investigate examples, manage their learning process, assess what they have learned, and drive further investigation.

As a first step toward RALE, we created new tools for helping web developers find relevant source code. We introduced Unravel [11], a tool that helps developers find entry points for understanding UI features on professional websites in 50% less time than when using Chrome Dev Tools. We also introduced Telescope [12], a tool which constructs a composite view of the less than 150 lines of relevant code that produce a UI interaction out of the 10k-100k lines of code on professional websites.

While Unravel, Telescope, and prior tools reduce the complexity of exploring professional code, these solutions target experienced developers and are insufficient for helping novice developers build a conceptual understanding of how code constructs work together to implement a feature. To overcome such difficulties, we designed *learner-centric developer tools* that embed sensemaking scaffolds informed by the learning sciences to explicitly support novices building such conceptual models. We created Isopleth [10], a web-based platform for making sense of complex code constructs and hidden asynchronous relationships in professional web applications. We also created Ply [15], a web inspector that computes visual relevance to highlight the relationship between CSS code and observable effects on professional webpages. **Results from four studies show that using these tools helped novice developers build significantly more accurate conceptual models of how professional features are implemented, and helped to surface expert design patterns across a diverse range of complex, professional websites.** In other words, by designing effective scaffolds for self-directed learning, we are able to help novices learn professional concepts from complex, professional artifacts. In contrast, existing approaches can only help experienced developers use such materials, and can only help novices learn basic concepts from constrained tutorial examples.

Our work on RALE is beginning to realize the vast opportunities for supporting authentic learning from professionally-made artifacts. In continuing work, **we are advancing the design of *mixed-initiative scaffolds*, in which learner-created artifacts are used in conjunction with automated methods and interface affordances to realize the desired learning scaffolds.** This integrative computing approach seeks to overcome the limitations of fully automated approach, which cannot scaffold the entire self-directed learning process of building conceptual models, practicing concept implementation, and applying learned concepts across examples. Advancing mixed-initiative solutions provides a compelling direction for enabling many learners to pursue mastery in complex domains when existing approaches do not scale the same learning benefits, and to generally enable solutions to complex problems when automated approaches fall short.

4 Scaling Research Training to Cultivate Self-Directed Learners

The fourth track of my research aims to expand our ability to train novice researchers to self-direct complex work. My goal is to cultivate self-directed learners, significantly increase the number of students who receive authentic research training, and lower the barrier to participating in authentic research experiences.

Apprenticeship, or 1-on-1 mentoring, provides a powerful model for training young researchers, but imposes a heavy orchestration burden on a single faculty researcher who can only mentor a handful of students at a time. Novice researchers lack *regulation skills*, i.e., cognitive, motivational, emotional, metacognitive, and strategic behaviors for reaching desired goals and outcomes. Leading a research project requires students to develop regulation skills to plan research work and to adopt effective help-seeking and collaboration strategies to overcome challenges. Training and support for developing such skills is necessary but time-consuming for mentors and difficult to support with software alone. As a consequence, most students are relegated to rote tasks and only a few advanced students are trained to direct complex work themselves.

As PI of an NSF Cyberlearning award, I am working with Matt Easterday, Liz Gerber, Leesha Maliakal, and others to design Agile Research Studios (ARS) [19], a computational ecosystem for research training that disperses control of the learning process across the entire community. ARS: (1) adapts agile methodologies from software development and design for the purposes of learning regulation skills and conducting independent research; (2) designs social structures for training and support without reliance on a single mentor; and (3) provides new technologies that support learning to plan, seek help, and reflect.

ARS extends the scale and capacity of a community to produce and learn in ways that a mentor cannot possibly support on their own. By reconceptualizing the role of students as self-directed learners and distributing support across the community, ARS frees up mentors to respond to challenges they can best address and to teach regulation skills. Applying integrative computing and ecological thinking, its underlying design addresses the core computational bottlenecks in scaling by shifting the roles of faculty members, graduate students, and undergraduates to support their respective needs and those of other ecosystem members. For example, by developing processes and tools such as pair research [17]—which matches students to help one another across diverse learning needs—while also scaffolding and encouraging helping behaviors more generally, ARS intends to grow a supportive community and helping culture over time through practice.

The Design, Technology, and Research (DTR) program that I founded and direct at Northwestern used the Agile Research Studios model to train 70 students (7 PhDs, 63 undergraduates) who led 35 independent research projects, won 33 undergraduate research grants, published 14 student-led papers and extended abstracts at top HCI conferences (CHI, UIST), and won 6 major student research competitions, placing first at ACM CHI in the past two years. Beyond supporting students to produce research, early results [19] show that the ARS model supported students developing planning skills and increased students' help-seeking dispositions as they learned to more effectively self-direct complex work.

Research Impact and Future Directions

My research, to date, has resulted in 32 major publications with 1000+ citations, leading to an h-index of 16 (using Google Scholar). My work has been published at premier venues in HCI (CHI, UIST), AI (AAAI, IJCAI), Social and Crowd Computing (CSCW, HCOMP), and the Web (WWW). It is supported by 4 awarded NSF grants on which I serve as PI, a Microsoft FUSE Labs research award, three Murphy Society Awards, and an Office of the Provost Award for Digital Learning. To broaden the reach of my research, I have given invited talks at nine major research institutions and at the AAAS annual meeting. I have also submitted a contributed article to CACM on *"Designing Computational Ecosystems to Advance Human Values"* for review. **The article discusses the need and opportunities to look beyond individual technologies, and provides a conceptual framework for designing computational ecosystems that scale solutions to complex human problems and advance desired human values.**

While my work thus far has focused on forming new computational ecosystems, there is also a need to manage and evolve existing ecosystem solutions as a community grows and its needs change. In continuing work, I am interested in designing systems with *ecosystem-level intelligence* that can monitor, coordinate, and restructure the interactions across a computational ecosystem to improve its function globally over time. Building on the success of ARS, I recently submitted an NSF CAREER proposal on *"Mixed-Initiative Orchestration Technologies for Advancing Community-Based Learning and Problem Solving in Research Communities."* As communities and the technologies that support them grow in size and complexity, it becomes difficult for any community member or component technology to keep track of the multitude of needs and to effectively enact mentoring strategies and to coordinate and distribute support. Addressing this gap, **I proposed a *mixed-initiative ecosystem management framework* for managing needs and resources globally across diverse ecosystem interactions in a highly-distributed manner.** Using this framework, I plan to develop *ecosystem-level mechanisms* for (1) identifying needs and connecting to opportunities for mentoring and collaboration; (2) supporting continual learning and skill development; and (3) responding to changing problem-solving needs by adjusting the use of existing resources and building new capacity. To expand the broader impacts of my work, I will also develop training materials for designing computational ecosystems, and host workshops on designing integrative human-machine systems at leading conferences.

Broadening my domain of study, I have begun to design a computational ecosystem for connecting family, friends, and strangers. Despite having a plethora of social technologies for connecting at a distance, meeting both our desire for direct engagement and our need to fit these interactions into our busy lives remains a challenge. My work with Darren Gergle, Ryan Louie, and others aims to enable *computer-facilitated opportunistic interactions* that provide new opportunities for engaging in shared activities and that foster new and existing communities. I plan to submit an NSF proposal on this work in the coming year.

As computing deepens its reach into every aspect of society, there is a real opportunity for it to advance the richness of human life, but it would be naive to think that technology alone can solve the most pressing human problems or advance the values we hold dearest. Perhaps paradoxically, broadening the impact of computing technologies will require us to recognize the inherent limits of individual technologies and the need for conceptual and technical advances in integrative thinking. In this way, the study of computational ecosystems charts a course through which we can better align the systems we build with the values we hold. It calls on us to envision solutions in which people and technology are part of a larger ecology, and that bring to bear all of our human and machine abilities to deal with the complexities of the world.

References

[1] Zachary Allen. Gaze: Using mobile devices to promote discovery and data collection. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, pages 97–102. ACM, 2015.

[2] Paul André, Haoqi Zhang, Juho Kim, Lydia Chilton, Steven Dow, and Rob Miller. Community clustering: Leveraging an academic crowd to form coherent conference session. In *Proc. HCOMP 2013*, 2013.

[3] Anant Bhardwaj, Juho Kim, Steven P. Dow, David Karger, Samuel Madden, Robert C. Miller, and Haoqi Zhang. Attendee-sourcing: Exploring the design space of community-informed conference scheduling. In *Proc. HCOMP 2014*, 2014.

[4] Anthony S Bryk, Louis M Gomez, Alicia Grunow, and Paul G LeMahieu. *Learning to improve: How America's schools can get better at getting better*. Harvard Education Press, 2015.

[5] Stephen Chan, Nicole Zhu, and Haoqi Zhang. Tapshare: Participatory communitysensing through physical gestures and mobile sensors. Technical report, Northwestern University, 2015.

[6] Kevin Chen and Haoqi Zhang. Remote paper prototype testing. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 77–80. ACM, 2015.

[7] Lydia Chilton, Juho Kim, Paul Andre, Felicia Cordeiro, James Landay, Dan Weld, Steven Dow, Rob Miller, and Haoqi Zhang. Frenzy: Collaborative data organization for creating conference sessions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014.

[8] George W Furnas. Future design mindful of the MoRAS. *Human-Computer Interaction*, 15(2):205–261, 2000.

[9] Atul Gawande. Atul Gawande: How do we heal medicine. *TED video*, 19:19, 2012.

[10] Josh Hibschman, Darren Gergle, Nell O'Rourke, and Haoqi Zhang. Isopleth: Supporting sensemaking in professional web applications. *Submitted to ACM Transactions on Computer-Human Interaction (TOCHI)*, 2018.

[11] Josh Hibschman and Haoqi Zhang. Unravel: Rapid web application reverse engineering via interaction recording, library detection, and source tracing. In *Proceedings of the 28th annual ACM symposium on user interface software and technology*, UIST '15. ACM, 2015.

[12] Josh Hibschman and Haoqi Zhang. Telescope: Fine-tuned discovery of interactive web ui feature implementation. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pages 233–245. ACM, 2016.

[13] Juho Kim, Haoqi Zhang, Paul André, Lydia B. Chilton, Wendy Mackay, Michel Beaudouin-Lafon, Robert C. Miller, and Steven P. Dow. Cobi: A community-informed conference scheduling tool. In *Proc. UIST 2013*, 2013.

[14] Yongsung Kim, Darren Gergle, and Haoqi Zhang. Hit-or-wait: Coordinating opportunistic low-effort contributions to achieve global outcomes in on-the-go crowdsourcing. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 96. ACM, 2018.

[15] Sarah Lim, Josh Hibschman, Nell O'Rourke, and Haoqi Zhang. Ply: A visual web inspector for learning from complex professional examples. In *Proceedings of the 31st Annual Symposium on User Interface Software and Technology (to appear)*. ACM, 2018.

[16] Katherine Lin, Henry Spindell, Scott Cambo, Yongsung Kim, and Haoqi Zhang. Habitsourcing: Sensing the environment through immersive, habit-building experiences. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pages 639–650. ACM, 2016.

[17] Robert C Miller, Haoqi Zhang, Eric Gilbert, and Elizabeth Gerber. Pair research: matching people for collaboration, learning, and productivity. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 1043–1048. ACM, 2014.

[18] Hope Reese. Mastery of AI has been 'harder than expected' and 'future is uncertain,' says Microsoft's AI chief. *TechRepublic*, 2015.

[19] Haoqi Zhang, Matthew W Easterday, Elizabeth M Gerber, Daniel Rees Lewis, and Leesha Maliakal. Agile research studios: Orchestrating communities of practice to advance research training. In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 45–48. ACM, 2017.