VINEET PANDEY   SOCIAL COMPUTING SYSTEMS FOR COMPLEX WORK

My research advances the design of social computing systems by integrating learning and collaboration to enable complex work such as generating and evaluating scientific theories. Over 600 people from 30 countries have self-organized to generate theories about the human microbiome and test them by running experiments. My research raises the question: how can global communities create knowledge that meets their goals without waiting for experts to lead? My research prototypes collective systems for large-scale problems.

Many people have strong personal motivations and contextual insights. To create knowledge, they need mental scaffolds for organizing complex work, domain knowledge to compose and execute the steps, and ways to ask for help. Professional scientists benefit from conceptual knowledge, professional training, pre-existing organizational structure for collaboration, and direct access to resources. Currently, citizens lack these resources.

My doctoral research demonstrates how we might draw on people's diverse background knowledge, interest, and micro-expertise to expand scientific knowledge and push it in new directions. More specifically, the *Gut Instinct* platform that I have built instantiates these ideas enabling participants of the American Gut Project (the world's largest crowdfunded citizen science project) to generate and experimentally investigate hypotheses (Figure 1). 344 volunteers from 27 countries created 399 hypotheses about their health and the gut microbiome. Remarkably, microbiome scientists rated a fifth (75) of these hypotheses to have a scientifically valuable insight about a topic not covered by existing published work. Volunteers fleshed out 60 of these hypotheses into complete experimental designs. My entire work (code + data) is open source so others can edit, build, and experiment.

## Scaffolding Citizen-powered Knowledge Work

Citizens have a different background than professional scientists; they have unique personal experiences but lack the years of domain training. To create computational systems that leverage their strengths and mitigate the lack of training, I focus on domains where the science is nascent, highly contextual, and personally motivating. Synthesizing the crowdsourcing literature and my experience highlights three challenges: poor signal-to-noise from crowds due to lack of training; inefficient collaboration without careful attention; and poor results (or no results at all) unless experts lead. To address these concerns, my work introduces and evaluates peer production architectures and procedural learning.

The conceptual insight in my research is *creating learning abstractions for social computing by integrating conceptual knowledge and procedural skills*. To instantiate these ideas, I developed the *Gut Instinct* platform for citizen science [3–5]. Using *Gut Instinct* is like having a mentor available to guide you as you create hypotheses and experiments and then working with a global community to refine and test your work. My systems divide an activity (like experimentation) into

Idea   Design   Review   Run   Results

Figure 1 *The Gut Instinct platform enables anyone to transform their intuitions to hypotheses and then design and run experiments to test them [3-5]. Gut Instinct integrates conceptual learning embedded via short lectures and software-guided procedural learning to enable designing and reviewing experiments. Participants from around the world join experiments, follow instructions, and provide data in response to automated data collection reminders.*

separate *design, review,* and *run* tasks that embed learning for rapidly creating designs from lived experiences, solicit community contributions to improve and generalize original design, and automate routine tasks for correctness (Figure 1). *Gut Instinct*'s domain of choice for early usage is the microbiome. The microbiome offers an area with a lot of popular excitement, many opportunities for simple experiments, and a literature with more questions than answers. While everyone has a gut full of microbes, their causal influences remain largely unknown.
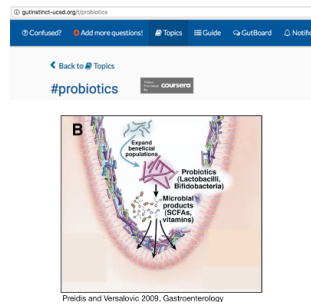
Figure 2 *Docent integrates conceptual and procedural learning; the combination substantially improved quality of hypotheses in a 2x2 experiment with 344 online participants [4]. Participants generated 399 hypotheses; 19% were rated novel by microbiome experts*

Principles to Integrate Learning in Social Computing

While large creative activities overwhelm novices with their complexity, experts are less daunted because they see an underlying system that supports different parts of the activity. *Gut Instinct* mitigates this expertise gap with interfaces that divide activities into phases.

Novices lack knowledge of what to do and how

Success with complex creative activities (like scientific work) requires both conceptual knowledge (facts) and procedural knowledge (how to do things). To help novices start contributing, my work introduces conceptual learning embedded in short lectures, and software-guided procedural learning. The *Docent* hypothesis generation system links conceptual learning (through online videos) with question-asking (by giving peers feedback) (Figure 2) [4]. My work demonstrates that prompting participants to explicitly connect personal observations with existing knowledge increases work quality. Because learning complex activities overwhelms working memory, the *Galileo* experimentation system chunks related elements and embeds procedural training [5]. Chunked presentation reduces cognitive load and offers a just-in-time plug for holes in knowledge (Figure 3).

Many hands make light work, but novices need clear contribution opportunities
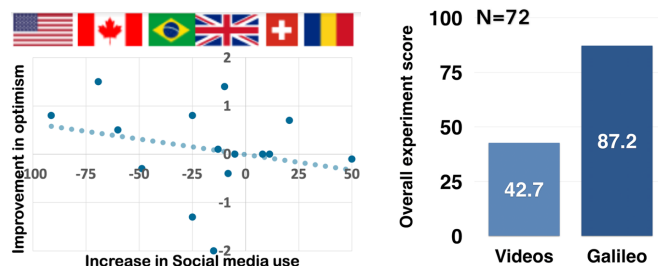
The crowdsourcing literature offers many good verification approaches for tasks with clear right or wrong answers – like whether two images represent the same product or what street number is written on a sign. However, verifying knowledge work necessitates a different approach because it requires making situationally-appropriate choices. *Gut Instinct* divides multi-party collaboration into complementary tasks and supports them using different contribution mechanisms (like adding a question, editing a response) and roles (like experimenter, reviewer, participant). This provides people the flexibility to choose how much they'd like to contribute. Finally, *Gut Instinct*

Figure 3 *Open Humans community (openhumans.org) designed and ran an experiment with participants from 7 countries and found that decreasing social media use improved optimism. Additionally, controlled experiments showed Galileo significantly improves experimental quality [5]*

vides people the flexibility to choose how much they'd like to contribute. Finally, *Gut Instinct*

automatically manages multiple activities to reduce bias and experimenter/participant workload, such as randomized placement of people into conditions, maintaining anonymity, and collecting and cleaning data.

Research Agenda

I plan to build systems that *democratize expertise* and provide ways to meaningfully embed computation in society. This section outlines future research ideas that open up multiple directions for designing new systems and for contributing novel theory.

1 Learning tools for end-users to perform higher-order scientific work

Learning has always been lifelong. Rapid change and the ready availability of online resources make it even more so. My work seeks to place learning experiences at the right time for people to use them. This offers both theoretical and practical benefits. The petri dish for my dissertation research has been guiding novices through doing experiments. In the learning sciences, Bloom's taxonomy shows a hierarchy of ways of engaging knowledge, from remembering facts to evaluating theories. Traditionally, this diagram invites a discussion of classroom learning objectives. For me, such an order implies a research trajectory. How might we design online learning environments for people to move their engagement with knowledge up the hierarchy? This has the potential to diversify the stakeholders and contributors to our future society.

Data science for personal benefits provides a domain that is structurally similar to microbiome science; it is personally motivating, contextually rich, and nascent. Type 1 diabetes researchers at UCSD and I are discussing plans to enable a community of patients to build their own personalized glucose recommendation engines. Leaders of Precision Medicine Initiative, Lyme Disease patients, and maternal health communities have expressed interest in these ideas. I intend to continue collaborating with academic researchers (such as American Gut Project collaborators), industry groups, and lead user communities (like Open Humans) to design and deploy such tools.

2 Social computing environments that model richer work

Domain experts make creative contributions like writing articles, curating museums, leading teams, and more. How might we enable a federation of novices to do these activities? As in science, the number of experts is relatively small and their training relatively homogenous. I intend to develop techniques to conceptually divide activities into tasks amenable to learning, community inputs, automation, expert help, and other techniques I haven't explored yet. For instance, how might we integrate people's contextual expertise in existing data collection and management workflows?

3 A science of social computing systems

Social computing systems leverage ideas from multiple disciplines and require substantial engineering. A science of such systems can support early-stage researchers in the face of high-barrier to entry and multiple failure modes. How can we rapidly prototype, debug, and evaluate social computing systems? How might systems support co-design by users? How might we leverage similarities between Bloom's taxonomy of learning and the hierarchy of social computing roles? Finally, what are the limits to solving problems with social computing? Principles derived from psychology, organizational behavior, and previous systems provide a starting point [6], but they can also be poor predictors of emergent behavior in new systems. We need a combination of theory, prototyping tools, and benchmarks. For instance, one approach might be to categorically separate measures for system evaluation (e.g. do people collaboratively create better questions using *Docent*?) from feature evaluation (e.g. do people improve another user's question using *Docent*'s edit feature?). A clear separation might help system designers sort the evaluation components in order of importance, assign different quality thresholds (e.g. controlled experiments vs observational), and communicate overall evaluation effectively to the community.

To summarize, my dissertation research created systems for complex work by drawing on insights from interactive systems, social computing, and learning theory. I have been the primary system designer for all the papers in my dissertation. In building and evaluating my platform, I have collaborated with 6 communities (including microbiome enthusiasts, Open Humans) and 27 co-authors from 8 fields including microbiology [3–5], cognitive science, learning psychology [2], and systems [1,7]. Working with multiple domain experts has taught me the importance of finding common ground and developing a shared vocabulary. In return, my collaborators have brought their diverse insights to human-computer interaction work; they have also taken HCI techniques home. Some of them now use needfinding and low-fidelity prototyping techniques before beginning complex software development. Many diverse efforts, including Precision Medicine Initiative (allofus.nih.gov), Zooniverse (zooniverse.org), and Foldscope Microcosmos (foldscope.com), are interested in using our work to diversify and deepen citizen contributions. I intend to engage in research that solves difficult problems and helps people. I plan to do this in collaboration with computer scientists, biologists, designers, and the wonderful communities that make this research possible.

References

1. Concerto: a high concurrency key-value store with integrity. 2017. Arvind Arasu, Ken Eguro, Raghav Kaushik, Donald Kossmann, Pingfan Meng, Vineet Pandey, and Ravi Ramamurthy. In *Proceedings of the 2017 ACM International Conference on Management of Data (SIGMOD)*, 251–266.

2. Framing feedback: Choosing review environment features that support high quality peer assessment. 2016. Catherine M Hicks, Vineet Pandey, C Ailie Fraser, and Scott Klemmer. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 458–469.

3. Gut Instinct: Creating Scientific Theories with Online Learners. 2017. Vineet Pandey, Amnon Amir, Justine Debelius, Embriette R Hyde, Rob Knight, and Scott Klemmer. In *2017 CHI Conference on Human Factors in Computing Systems (pp. 6825-6836). ACM.*

4. Docent: transforming personal intuitions to scientific hypotheses through content learning and process training. 2018. Vineet Pandey, Justine Debelius, Embriette R Hyde, Tomasz Kosciolek, Rob Knight, and Scott Klemmer. In *Proceedings of the Fifth Annual ACM Conference on Learning at Scale*, 9.

5. Galileo: Scaling Citizen-led Experimentation with a Procedural Training Platform. 2019. Vineet Pandey, Tushar Koul, Chen Yang, Daniel McDonald, Rob Knight, and Scott Klemmer. *In Submission*.

6. Building Successful Online Communities: Evidence-based social design. 2011. P. Resnick and R. Kraut. *MIT Press, Cambridge, MA.*

7. An HCI View of Configuration Problems. 2016. Tianyin Xu, Vineet Pandey, and Scott Klemmer. *arXiv preprint arXiv:1601.01747.*