Social Computation

Vineet Pandey

November 15, 2021

I am a **Human-computer interaction researcher**. My areas of interest include digital health, social computing, and accessibility. I design, build, and evaluate social computing systems and end-user tools that support communities in performing complex knowledge work like experimentation. My research draws on the knowledge and needs of experts (microbiologists and neurologists) and communities (fermenters and people with neurological disorders).

2005 Machines struggled to label images, translate phrases, and fold protein structures. As access to the internet expanded, Luis von Ahn [1] saw an opportunity: people could do what machines failed at. Under the hood, such *human computation* approaches also curated massive labeled datasets that have bootstrapped important technological successes.

2021 People cannot match the scale and latency at which machines label images, fold proteins, and translate phrases. Meanwhile, social computing has exploded. People globally collectively spend millions of hours every minute online. I see a massive opportunity for *Social Computation*.

I propose **Social Computation as a class of computational and social computing techniques that enable community-led complex knowledge work**. Social Computation techniques build on communities' contextual knowledge, social structures, and lived experience. To support motivated communities, tools and platforms embed techniques that provide conceptual and procedural guidance [2], collaboration roles [3], and designs for institutional impact [4] [5]. Online spaces are typically designed by drawing on the presuppositions and relationships that mark offline systems[6]. Social Computation provides an alternate vision where communities need not just donate data to answer experts' questions. Rather, they can perform complex knowledge work themselves.

Widely considered hot-spots of misinformation and hoax, internet platforms demonstrate the challenging, complex relationship between science and society. My research seeks to improve this situation. In my doctoral and postdoctoral work at UC San Diego and Harvard University, I have built systems for Social Computation for science and telemedicine. **Contributing principles to social computing, accessibility, and digital phenotyping tools, my research demonstrates several firsts**. First, 344 volunteers from 27 countries converted their lived experience to hypotheses about the microbiome; experts rated 19% of such hypotheses as novel. Second, communities ran controlled experiments with global participation to test such hypotheses. In particular, A fermentation community tested whether drinking kombucha helps maintain normal stool consistency with a between-subjects experiment. Third, people with neurological disorders provide accurate and reliable estimates of motor impairments with a few minutes of at-home use of a tool. This work provides the first *in-the-wild* empirical evaluation of multiple motor impairment measures from accessibility literature.

My research has received multiple accolades and interest beyond my immediate research area of Human-Computer Interaction. My doctoral research was awarded the **School of Engineering Exemplary Ethical Engineering Award** at

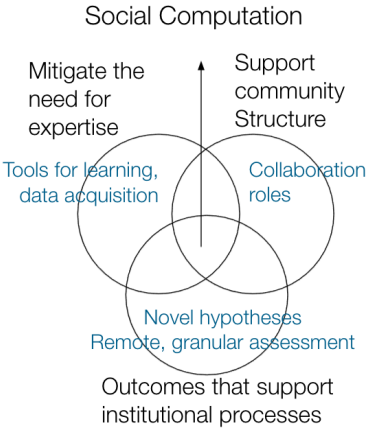

Figure 1: Social computation provides tools for deeper individual contributions, platforms that support community structures, and outcomes that are relevant for institutional processes.

[1] L. von Ahn. *Human Computation*. PhD thesis, Carnegie Mello University, 2005

[2] **Vineet Pandey**, J. Debelius, E. R. Hyde, T. Kosciolek, R. Knight, and S. Klemmer. Docent: Transforming personal intuitions to scientific hypotheses through content learning and process training. In *Proceedings of the Fifth Annual ACM Conference on Learning at Scale*, pages 1–10, 2018

[3] **Vineet Pandey**, T. Koul, C. Yang, D. McDonald, M. Price Ball, B. Greshake Tzovaras, R. Knight, and S. Klemmer. Galileo: Citizen-led experimentation using a social computing system. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2021

[4] **Vineet Pandey**, N. C. Khan, K. Z. Gajos, and A. S. Gupta. At-home use of a computer-based pointing task accurately and reliably estimates motor impairments. In *Submission*, 2022

[5] N. C. Khan, **Vineet Pandey**, K. Z. Gajos, and A. S. Gupta. Free-living motor activity monitoring in ataxia-telangiectasia. *The Cerebellum*, pages 1–12, 2021

[6] Jim Hollan and Scott Stornetta. Beyond being there. In Proceedings of the SIGCHI conference on Human factors in computing systems, pages 119–125,1992

my PhD institution. I have been invited to present research at multiple venues including **American Society of Microbiology**; **Precision Medicine Initiative**; and **MIT Innovation Lab**. I am grateful that multiple classes have including my research in their curriculum.

Future research directions include 1) online platforms that prototype different community-expert configurations; 2) computational tools for deeper knowledge mechanistic insights; and 3) techniques for better scientific work and discourse online. My goal is to make Social Computation systematic for a broad range of communities and experts to draw benefits from. and any researcher to use and broaden to more domains.

Achieving Social Computation: Tools, Platforms, Institutions

> "Part of what distinguishes science from cognition more broadly is the **cultural accumulation of tools and knowledge** that can support information gain in ways that go well beyond naive exploration." [7]

Synthesizing existing literature and my research suggests that technological efforts for complex knowledge production succeed when they intervene at three levels: individual, community, and institutional.

1. **Deepen individual contributions** with tools that mitigate the need for expertise. People need to know the genre of work and implement it correctly. While learning resources are distributed across the internet, they are rarely integrated with the task. People lack knowledge for *what* to do and *how*. Tools for designing experiments or collecting appropriate data can either provide such learning resources or change the representation the activity to a domain that mitigates the need for training.

2. **Support existing community structures**, motivation, and participation levels. Complex work can overwhelm people with the number and variety of tasks. To succeed at collaboration, technological interventions need to be directed towards some–e.g. progressive–forces over others [8]. Communities that accomplish impressive feats exhibit resilience, self-organization, and hierarchy [9]. Such communities demonstrate dynamic, contextual decision-making; they *act*; and they possess an organizational structure. While many online communities are massive, few support complex work with appropriate roles and support contribution mechanisms. My research suggests supporting motivated communities with explicit roles that reuse prior knowledge and provide scaffolds for micro-expertise.

3. **Produce outcomes that support institutional processes**. My postdoctoral research has taught me a key lesson: long-term success and value of community-led work relies on its success in leveraging existing institutional expertise and meeting existing needs. Many novel research tools and processes are not designed with the final beneficiaries needs in minds; others have noted it too [10]. To leverage communities' strengths and make their outputs useful, my research focuses on domains where communities can complement institutional work. The human microbiome offers an area with a lot of popular excitement, many opportunities for simple experiments, and a literature with more questions than answers. The science is *nascent*, *personally motivating*, and *contextual*. Experts know little, people care, and individual

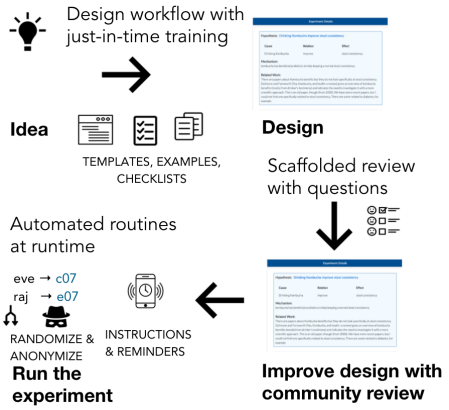

Figure 2: GALILEO supports communities in designing and running between-subjects experiments to test intuitions

Figure 3: A global fermentation community used GALILEO to design, review, and run a week-long experiment testing the effect of kombucha on stool consistency. The flags represent participants' nationality.

[7] Schulz, L. E. (2012). Finding new facts; thinking new thoughts.

[8] K. Toyama. *Geek heresy: Rescuing social change from the cult of technology*. PublicAffairs, 2015

[9] D. H. Meadows. *Thinking in systems: A primer*. Chelsea Green Publishing, 2008

[10] M. Roberts, D. Driggs, M. Thorpe, J. Gilbey, M. Yeung, S. Ursprung, A. I.Aviles-Rivero, C. Etmann, C. McCague, L. Beer, et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate forcovid-19using chest radiographs and ct scans.Nature Machine Intelligence,3(3):199–217,2021

differences are important. While everyone has a gut full of microbes, its causal influences remain largely unknown. The world could benefit greatly from a more comprehensive understanding of the microbiome, what influences its composition, and the impact our gut has on our health. Neurological disorders provide another area of enquiry. Clinical assessments evaluate progression but tracking early and subtle changes is difficult. Interpretable, granular, quantitative measures that support institutional needs can be useful.

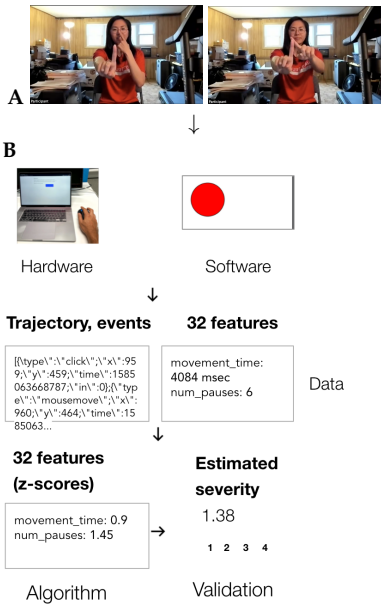

Figure 4: A) A demonstration of Finger-Nose task over video conferencing. B) HEVELIUS AT HOME tracks features. Families with neurological access Hevelius at home using a mouse and a browser on a personal computer. Activities comprise pointing tasks and reports from participant and caregivers.

References

N. C. Khan, **Vineet Pandey**, K. Z. Gajos, and A. S. Gupta. Free-living motor activity monitoring in ataxia-telangiectasia. *The Cerebellum*, pages 1–12, 2021.

D. H. Meadows. *Thinking in systems: A primer*. Chelsea Green Publishing, 2008.

Vineet Pandey, J. Debelius, E. R. Hyde, T. Kosciolek, R. Knight, and S. Klemmer. Docent: Transforming personal intuitions to scientific hypotheses through content learning and process training. In *Proceedings of the Fifth Annual ACM Conference on Learning at Scale*, pages 1–10, 2018.

Vineet Pandey, T. Koul, C. Yang, D. McDonald, M. Price Ball, B. Greshake Tzovaras, R. Knight, and S. Klemmer. Galileo: Citizen-led experimentation using a social computing system. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2021.

Vineet Pandey, N. C. Khan, K. Z. Gajos, and A. S. Gupta. At-home use of a computer-based pointing task accurately and reliably estimates motor impairments. In *Submission*, 2022.

K. Toyama. *Geek heresy: Rescuing social change from the cult of technology*. PublicAffairs, 2015.

L. von Ahn. *Human Computation*. PhD thesis, Carnegie Mello University, 2005.