Supporting High-Stakes Investigations with Expert-Led Crowdsourcing

Sukrit Venkatagiri
Virginia Tech, Arlington, VA
sukrit@vt.edu

Abstract

Expert investigators possess advanced skills and experience, but they face limits on their time and attention. In contrast, crowds of novices can be highly scalable and parallelizable, but lack expertise. My dissertation explores ways to combine experts' deep domain knowledge and experience with the speed and scale of crowds, an approach I call expert-led crowdsourcing (ELC). I study: 1) how novice crowds can augment expert investigators' work practice; 2) the ethical tensions of conducting an ELC investigation for sensitive, real-world investigations; 3) how capture-the-flag competitions increase inter-team collaboration in ELC investigations; and 4) how different teamwork structures affect intra-team collaboration in ELC investigations.

Expert investigators in fields such as journalism, law enforcement, and human-rights advocacy are tasked with conducting investigations of increasing scope but with dwindling resources. Before experts can report on their findings, they must thoroughly investigate every claim and piece of evidence. This includes inspecting evidence for clues, tracing and verifying the provenance of each piece of evidence, interviewing eye-witnesses, and corroborating facts using digital tools. These labor-intensive tasks can take hours to days, and may not always prove fruitful. They also do not scale easily, meaning that successful investigations are limited by experts' time and attention.

An alternative approach that has seen success is leveraging the powerful and adaptive capabilities of distributed online crowds. While some crowdsourced investigations have resulted in successes, such as locating missing persons, catching criminals, and supporting crisis response efforts, they are perhaps better known for high-profile failures involving vigilantism, or its online form, "digilantism." These include misidentifying individuals involved in the 2013 Boston Marathon bombing and the 2020 George Floyd protests, among others. Despite this criticism, crowdsourced investigations continue to flourish and evolve.

In this dissertation, I propose an approach called expert-led crowdsourcing (ELC) that combines experts' deep domain knowledge and experience with the speed and scale of crowds to effectively, ethically, and efficiently scale-up high-stakes investigations.

Borrowing from the traditional top-down model of crowdsourcing, ELC involves experts that provide evidence materials to the crowd, focus the crowd's efforts on specific topics, and reserves the authority of drawing final conclusions for experts. Borrowing from the bottom-up self-organizing crowd model, ELC relies upon intrinsically motivated novices coming together to help solve investigations.

Unlike either of these traditional models, ELC gives novice crowds largely unfettered access to real evidence materials in a controlled environment, facilitating trust and security between the expert and the crowd. The crowd is also provided with formal training in investigative techniques and tasked with analyzing, interpreting, and even questioning the data. In doing so, they can potentially speed up and scale up an expert investigator's work practice.

My dissertation focuses on the domains of journalism, law enforcement, and human rights-advocacy, where professionals avoid inaccurate statements, unethical or vigilante conduct, and slow completion or response rates. Designing a successful ELC investigation in high-stakes domains not only requires effectiveness, and efficiency, but also ethical conduct. I explore these dimensions through three research questions:

- **RQ1**: How can novice crowds effectively augment expert investigators' work practice?
- **RQ2**: What are the ethical tensions in expert-led crowdsourced investigations?
- **RQ3**: How do competition, collaboration, and teamwork structures affect efficiency in an ELC investigation?

Completed Research

GroundTruth: Augmenting Expert Image Geolocation with Crowdsourcing and Shared Representations The first part of my dissertation tackles how to effectively scale up expert investigators' work practice through crowdsourcing. I focused on a key step in many experts' investigative process, called *image geolocation*. It is a sensemaking process that involves verifying the exact location where photo or video imagery was taken. This manual task may take hours or days and does not scale easily. Further, computer vision

attempts at automating this process are insufficiently accurate, placing photos within within 200km of the correct location less than 30% of the time.

In this study, I answer **RQ1** by developing GroundTruth (Venkatagiri et al. 2019), a web-based system that helps expert investigators geolocate images with a crowd. I proposed an approach, called *crowd-augmented expert work*, that extends Heer's idea of shared representations between humans and intelligent agents (Heer 2019), and used it to facilitate crowd-supported expert image geolocation. GroundTruth allows visual traits and context to be communicated between experts and novice crowds performing a complex sensemaking task: image geolocation. The system augments but does not replace experts, while still promoting correctness, and it requires "neither perfect accuracy nor exhaustive modeling of the user's tasks to be useful."

GroundTruth consists of three shared representations as system components: (1) an expert-created aerial diagram to help share context with the crowd, focus their attention, and overcome their spatial reasoning limitations; (2) a gridded map overlay specified by experts that generates microtasks for crowd workers, indicating where they should search, while providing the expert an overview of crowd progress; and (3) a heatmap displaying expert and crowd decisions which quickly and at-scale indicates to the expert where their own time and attention is best spent.

I conducted a mixed-methods evaluation of GroundTruth involving a think-aloud protocol, log analysis, and semi-structured interviews with 11 experts working with 567 crowd workers from Amazon Mechanical Turk. I found that GroundTruth effectively merges the benefits of both expertise and crowdsourcing, demonstrating the feasibility of crowd-supported expert image geolocation using shared representations. Experts worked with crowds in real-time to narrow the search area substantially, and frequently succeeded in geolocating the image. Experts were also excited by the idea of incorporating GroundTruth into their toolset since it provided features that were not available to them.

CrowdSolve: Managing Tensions in an Expert-led Crowdsourced Investigations GroundTruth showed that crowds can effectively support experts' investigative process using shared representations. However, the system was evaluated in the lab and was focused on a specific investigative task. Further, while some crowdsourced investigations have resulted in successes, they are perhaps better known for high-profile failures, including misidentifying individuals involved in the 2013 Boston Marathon bombing. In the second part of my dissertation, I answer **RQ2** by studying the ethical challenges of conducting a real-world crowdsourced investigation led by law enforcement investigators (Venkatagiri, Gautam, and Luther 2021).

I conducted an ethnographic study of CrowdSolve, a crowdsourcing event that blended top-down guidance by law enforcement experts with bottom-up participation by a crowd of more than 250 amateur sleuths. Over the course of four days in October 2019, experts and the novice crowd collaborated in a co-located setting in Seattle, Washington, USA to investigate two decades-old unsolved murder cases. I describe how CrowdSolve represents a third model of crowdsourced investigations, called *expert-led crowdsourced investigations*.

Using Lee et al.'s lens of human infrastructure (Lee, Dourish, and Mark 2006), I examined two features of the event that enabled success for all of its stakeholders: using a controlled environment and allowing each stakeholder group to contribute meaningfully. However, within any human infrastructure, there is the potential for friction, because various stakeholders have conflicting motives and actions (Tatar 2007).

I found that the CrowdSolve organizers managed the tension between openness and security and privacy. On the one hand, opening up the cases and sharing as much information with attendees as possible would maximize the chances of discovering new leads. On the other hand, the organizers had to weigh the privacy considerations of the victims and their families, as well as law enforcement's desire to maintain a viable legal case and avoid tainted juries.

I also found that attendees' enthusiasm for true crime blurred the boundary between entertainment and reality. This enthusiasm led to invested participation, but it occasionally bordered on fetishization. Attendees also desired closure, but were limited due to the nature of criminal investigations. The organizers' unconventional yet beneficial decision to have the victims' families at the event supported attendees' dual desires for altruism and immersion. The families' presence helped attendees empathize, heightened the stakes of the event, and strongly motivated attendees to work hard and generate leads.

CrowdSolve successfully leveraged the complementary strengths of experts and novice crowds. First, the organizers secured the participation of law enforcement experts who ran training and discussion sessions where they provided high-level guidance and leadership to keep the crowd focused on making progress. Second, the event activities were parallelized and led by an expert, taking full advantage of the 250–strong crowd. As a result, I found that the experts not only helped attendees learn, but also that attendees applied this knowledge to generate new and useful leads.

By framing volunteer crowd work as an act of fandom, leaders and requesters can come to see crowds as more than just interchangeable human processing units (HPUs). As in games with a purpose, organized efforts, like CrowdSolve, allow participants to indulge in their passions while also contributing meaningfully to society.

Research in Progress

QuriOSINTy: Combating Misinformation through Capture-the-Flag Competitions My prior work showed that shared representations of expert processes can enable crowd workers to effectively augment expert work practice. I have also shown that expert-led crowdsourcing (ELC) can be used to conduct high-stakes, real-world investigations — while mitigating the possibility of sensitive information being leaked. However, I found at CrowdSolve that there was duplication of effort across teams and that there was minimal shared context between teams. This may have

reduced the efficiency of the event. I also found that attendees' motivations bordered on fetishization.

In the third part of my dissertation, I answer **RQ3** by studying how to make ELC investigations more efficient by reducing intra-team redundancies and increasing inter-team collaboration. I focus on debunking misinformation through *open-source intelligence* (OSINT) techniques. OSINT relies on publicly available information online, thus enabling crowds to conduct real-world investigations without requiring access to law enforcement case files. By focusing on debunking misinformation, I can easily measure success.

In Spring 2021, I developed QuriOSINTy, a web-based system, using an iterative design-based research approach with a class of 40 students. QuriOSINTy divides the investigative process into four distinct phases: discovery, verification, archival, and reporting. QuriOSINTy enables crowd workers to view each other's work output, thus reducing redundancies within an investigation. Further, the system leverages a "capture the flag"-style (CTF) points-based setup to motivate teams to compete against each other to score the most points. To promote both competition and collaboration across teams, teams get points for the quality of their work as well as for building upon other's work.

The class used QuriOSINTy several times and I made multiple changes based on students' feedback and my observations. I conducted semi-structured interviews with 6 different teams, consisting of 16 participants in total. These teams had varied performance, ranging from last place to being in the top two ranks in the final CTF competition. I am now analyzing transcripts from the interviews as well as inspecting system log data. My preliminary findings show that teams performed better when: (1) they consisted of individuals who enjoy competition; (2) tried multiple different tactics instead of just one; and (3) took up tasks that suited their strengths. I also found that more teams were motivated more by altruism (i.e., combating misinformation in the real-world) than competing to win the CTF.

Increasing Intra-Team Efficiency through Liberating Structures In the first two parts of my dissertation, I studied how to design effective and ethical ELC investigations. In the third part of my dissertation, I studied how to help crowds collaborate more efficiently within an investigation. This final part of my dissertation evaluates how different leadership structures affect efficiency.

Prior work has shown that the leadership structure of a team can influence its performance. However, ELCs are different from traditional, homogeneous collaborations because experts simultaneously *lead* and *participate in* the collaboration. Further, experts not only possess greater domain expertise than crowds, but also perform the same and a superset of the crowd's tasks. Experts' own work is also shaped and redirected based on the crowd's results streaming in real-time. Hence, leadership structures may not be as effective with ELCs as they are with more homogeneous collaborations. In the final part of my dissertation, I will answer **RQ3** by evaluating how different teamwork structures affect the efficiency of an ELC investigation. I will evaluate four different liberating structures (Lipmanowicz and McCandless 2013) — microstructures that enhance relational coordination and trust —– over a semester-long class with 30 students working with expert investigators. I will conduct a mixed-methods evaluation consisting of surveys, document analysis, and semi-structured interviews to understand how well participants performed under each teamwork structure, as well as their experiences with each. My findings will inform the design of systems that instantiate teamwork structures that increase the overall efficiency of ELC investigations.

Anticipated Contributions and Goals

My dissertation informs the design of more effective, ethical, and efficient crowdsourced investigations. More specifically, I seek to contribute three mixed-methods evaluations of two web-based systems and various teamwork structures, as well as an ethnographic study of a 250-person crowdsourcing event. Throughout, I highlight the challenge and opportunities faced in conducting effective, ethical, and efficient expert-led crowdsourced investigations. I also contribute two conceptual contributions for enabling effective, ethical, and efficient expert-crowd collaborations in high-stakes investigations. The first conceptual contribution is the idea of shared representations for crowd-augmented expert work that enables novice crowds to effectively contribute to experts' investigative practice. The second is the expert-led crowdsourcing (ELC) framework. ELC blends top-down guidance by law enforcement experts with bottom-up participation by a crowd of novices.

I hope to receive feedback on the framing of my dissertation and gathering feedback for my final study (RQ3). Another goal of mine is to speak with senior researchers' about their experiences in academia and industry and how they built out a broader research narrative post-graduation.

References

Heer, J. 2019. Agency plus automation: Designing artificial intelligence into interactive systems. *Proceedings of the National Academy of Sciences* 116(6):1844–1850.

Lee, C. P.; Dourish, P.; and Mark, G. 2006. The human infrastructure of cyberinfrastructure. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work*, CSCW '06, 483–492. New York, NY, USA: Association for Computing Machinery.

Lipmanowicz, H., and McCandless, K. 2013. *The surprising power of liberating structures: Simple rules to unleash a culture of innovation*. Liberating Structures Press Seattle, WA.

Tatar, D. 2007. The design tensions framework. *Human–Computer Interaction* 22(4):413–451.

Venkatagiri, S.; Thebault-Spieker, J.; Kohler, R.; Purviance, J.; Mansur, R. S.; and Luther, K. 2019. Groundtruth: Augmenting expert image geolocation with crowdsourcing and shared representations. *Proc. ACM Hum.-Comput. Interact.* 3(CSCW).

Venkatagiri, S.; Gautam, A.; and Luther, K. 2021. Crowd-

solve: Managing tensions in an expert-led crowdsourced investigation. *Proc. ACM Hum.-Comput. Interact.* 5(CSCW1).