The Social Structures of OSINT: Examining Collaboration and Competition in Open Source Intelligence Investigations

Yasmine Belghith

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Computer Science and Applications

Kurt Luther, Chair
Andrea L Kavanaugh
Christopher L North

May 11, 2021
Blacksburg, Virginia

Keywords: Investigations, Open Source Intelligence, Competition, Self-organizing, Crowdsourcing

The Social Structures of OSINT: Examining Collaboration and Competition in Open Source Intelligence Investigations

Yasmine Belghith

ABSTRACT

Investigations are increasingly conducted online by not only novice sleuths but also by professionals — in both competitive and collaborative environments. These investigations rely on publicly available information, called open source intelligence (OSINT). However, due to their online nature, OSINT investigations often present coordination, technological, and ethical challenges. Through semi-structured interviews with 14 professional OSINT investigators from nine different organizations, we examine the social collaboration and competition patterns that underlie their investigations. Instead of purely competitive or purely collaborative social models, we find that OSINT organizations employ a combination of both, and that each has its own advantages and disadvantages. We also describe investigators' use of and challenges with existing OSINT tools. Finally, we conclude with a discussion on supporting investigators' with more appropriable tools and making investigations more social.

The Social Structures of OSINT: Examining Collaboration and Competition in Open Source Intelligence Investigations

Yasmine Belghith

GENERAL AUDIENCE ABSTRACT

Investigations are increasingly conducted online by not only novice investigators but also by professionals, such as private investigators or law enforcement agents. These investigations are conducted in competitive environments, such as Capture The Flag (CTF) events where contestants solve crimes and mysteries, but also in collaborative environments, such as teams of investigative journalists joining skills and knowledge to uncover and report on crimes and/or mysteries. These investigations rely on publicly available information called open source intelligence (OSINT) which includes public social media posts, public databases of information, public satellite imagery...etc. OSINT investigators collect and authenticate open source intelligence in order to conduct their investigations and synthesize the authenticated information they gathered to present their findings. However, due to their online nature, OSINT investigations often present coordination, technological, and ethical challenges. Through semi-structured interviews with 14 professional OSINT investigators from nine different organizations, we examine how these professionals conduct their investigations, and how they coordinate the different individuals and investigators involved throughout the process. By analyzing these processes, we can discern the social collaboration and competition patterns that enable these professionals to conduct their investigations. Instead of purely competitive or purely collaborative social models, we find that OSINT organizations employ a combination of both, and that each has its own advantages and disadvantages. In

other words, professional OSINT investigators compete with each other but also collaborate with each other at different stages of their investigations or for different investigative tasks. We also describe investigators' use of and challenges with existing OSINT tools and technologies. Finally, we conclude with a discussion on supporting investigators with tools that can adapt to their different needs and investigation types and making investigations more social.

Acknowledgments

I am sincerely and deeply grateful to all the following people; your support has inspired me and has given shape to the work I describe in the following pages.

First, I'd like to thank my parents, Nadia and Hamadi, and my siblings, Cyrine and Skander. Their words of wisdom and their constant enthusiasm have carried me through the ups and downs of the past two years.

My advisor, Dr. Kurt Luther, for offering me his mentorship and a seat in the Crowd Intelligence Lab. From learning the ins and outs of the academic research world and the intricacies of its methodologies, to overcoming the hurdles along the way, he has offered his trust, patience, and guidance. I will continue to heed his valuable feedback and hope to follow in his footsteps in my next endeavor of pursuing a Ph.D.

My thesis committee, Dr. Andrea Kavanaugh and Dr. Chris North. Dr. Kavanaugh's insights into the field of Human-Computer Interaction and her advice back when I was an undecided undergraduate freshman have guided me to this path and she continues to offer a sympathetic ear and her encouragement through the years. Dr. North's guidance and feedback on my first project contribution as a graduate student has allowed me to gain confidence as a researcher.

I also want to thank the members of the Crowd Intelligence Lab, especially my co-authors, Tianyi Li and Sukrit Venkatagiri, who have helped me throughout different stages of my research endeavors in the past two years. Your work and our discussions have shaped my vision as a researcher and have given life to many important ideas in this work. It would not be where it is today without your help.

Virginia Tech for being my home for the past six years and allowing me to grow in all aspects of life, and its wonderful administrative and technical staff for making paperwork, filings, and all your work to enable my research seem like a breeze.

Lastly, I would like to thank my partner, Zaid, and my dear friends for giving me a good laugh when I needed it most, for your patience, reassurance, and for riding with me through thick and thin.

Contents

List of Figures

List of Tables

List of Abbreviations

CSCW	Computer Supported Cooperative Work
CTF	Capture The Flag
HCI	Human-Computer Interaction
OSINT	Open Source Intelligence

Chapter 1

Introduction

Novice sleuths and professionals leverage publicly available information online in their investigations, with varying success rates. Novice investigators are often successful in uncovering crimes, finding perpetrators, and helping to deliver justice [1, 2]. They have also supported crisis response efforts [3, 4]. However, there have also been several well-known incidents involving online and in-person vigilantism [5, 6]. This includes "naming and shaming," disclosure of highly personal details (i.e., doxing), and misidentification of individuals, most notably in the 2013 Boston Marathon bombing [5] and the storming of the U.S. Capital on January 6, 2021 [7].

In contrast to novices' mixed results, professional investigators have been more successful both in the court of public opinion and the court of law [3, 8]. For example, professional investigators have used OSINT to uncover human rights violations in Syria [9] and provide information to law enforcement on the storming of the U.S. Capitol, leading to multiple arrests [10, 11]. In addition, they have avoided being implicated in incidences of vigilantism.

One reason for professional investigators' successes — and far fewer, if any, mishaps — may be due to their use of open source intelligence (OSINT) data, techniques, and underlying philosophy. OSINT refers to data that can be gathered from publicly and legally available sources [12]. Not only is OSINT a type of data, but it also encompasses an entire field, with its own rules and techniques for collecting, verifying, and analyzing open source information to derive intelligence and fulfill a goal (e.g., identifying a suspect, finding a missing person, proving or disproving a statement) [13]. It has numerous applications, ranging from employee vetting [14, 15] to counter terrorism and human rights advocacy [16, 17].

OSINT professionals often come together as part of larger organizations or events and leverage two different tactics to conduct their investigations: competition and collaboration. We refer to both tactics as social OSINT. While prior work has focused on OSINT tools and techniques [13, 18], our work here focuses on the social aspect of OSINT. Prior work in CSCW has shown that the social, human infrastructure of an organization can be just as important as the technological infrastructure [19, 20, 21]. Our work here seeks to inform future OSINT investigations — conducted by novices and professionals alike — with the goal of making them more effective and more ethical.

With this motivation, we address the following research questions in this paper:

1. RQ1: What are organizers' and contributors' motivations, experiences, and attitudes towards social OSINT investigations? How do they define success?
2. RQ2: How do organizers plan and structure the OSINT investigations? What challenges do they face, and how do they manage them?

To address these questions, we recruited OSINT investigators from organizations and events across the social structure spectrum, ranging from purely competitive to purely collaborative social structures. Through semi-structured interviews with 14 professional OSINT investigators from nine different organizations, we describe their backgrounds, motivations, and the commonalities and differences between their social structures. We also describe the factors that enable their success, such as their solid foundation in ethics and security, and the challenges that they face, such as the unreliability of certain tools and the difficulty in verifying digital content.

We also find that there is no clear delineation between collaborative and competitive structures. Instead, organizations employ competitive strategies within their overarching collaborative structures and vice versa, each having different implications, such as collaboration enabling investigators to broadly share their expertise with other, while competition helping them refocus their efforts. In addition, we find that these social structures are influenced by power dynamics outside of the organizations. For instance, the organizations that they work within dictate access to contributors and resources.

Our paper makes the following contributions:

1. An in-depth description of the social structures that support OSINT investigations, as well as defining and characterizing social OSINT. Our findings also add more nuance to related work on competitions.
2. We enrich the current literature on investigations within CSCW by presenting recommendations and implications for structuring other open source intelligence work and citizen investigations, both online and offline.
3. We suggest three design recommendations to better support the OSINT community.

Chapter 2

Review of Literature

2.1 Open Source Intelligence Investigations

Open source intelligence (OSINT) investigations involve the collection and analysis of publicly available data to generate intelligence that addresses a specific need [13]. More recently, the rise of social media and increasingly digitally-mediated social interaction has democratized access to large amounts of personal information, and powerful tools for analyzing it [22]. OSINT investigations of digital traces and social media are regularly conducted in domains such as journalism [23], business [e.g., 14, 24, 25, 26], counter-terrorism [27], cybersecurity [28], and human rights advocacy [16, 17].

McKeown et al. [29] argue that the target of an OSINT investigation will generally shape the type of investigation that will be carried out, the type of data that will be gathered and analyzed, the levels of detail that will go into the investigation, the tools used, the investigators' behaviors and attitudes, and the various outcomes of that investigation (e.g., reports, forecasts, news articles, criminal proceedings).

Conducting OSINT investigations involves more than just the type of data or techniques used, however. According to practitioners, OSINT also comes with its own ethos [13, 30]. The OSINT ethos prioritizes transparency, frowns upon the use of subterfuge, and limits investigations to passive reconnaissance [e.g., 16, 31].

This ethos may be the reason for professional OSINT investigators' successes and reduced rate of ethical mishaps. Directly contacting law enforcement coupled with only engaging in passive reconnaissance greatly lowers the possibility of vigilantism — especially doxxing and misidentification. For example, while both novices and professionals sought to use OSINT to investigate the storming of the U.S. Capitol in 2021, some novices publicly misidentified individuals [7]. On the other hand, John Scott-Railton, an OSINT professional, shared his findings directly with the FBI [10] and encouraged his collaborators and followers not to publicly tweet unconfirmed names. Scott-Railton's work directly led to two arrests [11].

Despite the advantages of OSINT investigations, there are also challenges [29, 32, 33]. During an investigation, issues can arise from the ephemerality of open source information online [16] or because deep fakes, dis-, and mis-information are harder to verify [16, 33]. Another challenge is the possible exposure to sensitive materials, especially when investigating violence of any kind, that can result in secondary trauma [34]. Regardless of resources, Gill [35] states that the success of an OSINT investigations depends on its social structure: its members, their roles, and their training.

Our work here contributes a deeper understanding of the *social* structures within OSINT organizations. We also highlight OSINT professionals' varied backgrounds and motivations, as well as the social and technological challenges that they face during the course of their investigations.

2.2 Investigations in Computer Supported Cooperative Work

Prior CSCW research has focused on top-down [e.g., 36, 37], bottom-up [e.g., 38, 39, 40], or hybrid investigations [21].

While all three investigation types include similar stages, such as collecting and analyzing information towards a specific goal (e.g., the opening of a criminal case, the identification of a suspect), top-down, law enforcement-led investigations are more commonly studied within CSCW. Here, access to information is limited by law enforcement. Prior work has focused on the design of tools to support collaboration and coordination between law enforcement agents [36, 37]. Sometimes, these top-down investigations benefit from members of the community or neighborhood residents passively providing information [4]. For example, Lewis and Lewis [41] examined a community's use of CLEARPath, a website that enables residents to "serve as an information sharing vehicle" between the police and the community, and found that residents used the forum to strengthen their social ties and discuss collective action. Brush et al. [42] proposed augmenting the potential for crime prevention through a digital neighborhood watch, linking neighbors' security cameras and alerting police of suspicious activity. Additional research surfaced the importance of civic engagement and communication, online and offline, between the police and communities in crime prevention [39, 40]. Sachdeva and Kumaraguru [43] recommend the design of technology that will increase interactions between police and residents, such as platforms to post concerns that require police response and attention.

On the other hand, bottom-up, novice-led investigations are typically self-organized by crowds, usually online, who coordinate their efforts and combine their diverse knowledge to conduct the different stages of an investigation. CSCW researchers have examined these crowd mobilizations on social media. Crowds take on varied roles, such as information diffusion during crises [e.g., 44, 45, 46], or conducting data analysis and validation in citizen science projects [47, 48]. Huang et al. [38] examined how crowds of online volunteers analyzed photos related to the Boston Marathon Bombings in the effort of identifying the perpetrators; however, this effort resulted in the infamous misidentifcation of a suspect. More recently, Arif et al. [46] studied mechanisms used by crowds to correct online information on social media about crisis events such as a rumored flight hijacking and the Paris Attacks, showing that, while crowds do share rumors, they undertake different strategies and attempts to correct them. Additional prior work also demonstrated that coordinated and directed crowds can augment an investigation's potential [e.g., 21, 49, 50, 51].

Our work here contributes to the growing body of literature within CSCW focused on understanding and supporting investigations. While prior work has focused on studying top-down investigations led by professionals, or bottom-up investigations led by novices, our work here focuses on *bottom-up* investigations led by *professional* OSINT investigators. Prior work has focused on how investigators leverage collaboration to scale up their investigations. We extend this work to study how individuals OSINT professionals leverage not only collaborative but also competitive efforts to conduct their investigations.

2.3 Competitions in Human-Computer Interaction

Using contests, a form of competition, to "reach a broad audience of people with various backgrounds, skills, and expertise has a long tradition." [52] Such competitions have played a major role in the development of innovations such as digital televisions and the first manned space mission to Mars, and are proposed by corporations, governments, or even non-profit organizations [52].

Within CSCW and HCI, researchers have explored how social technologies support and inhibit various forms of competition, ranging from innovation contests and hackathons [53, 54], and games and gamification [55, 56, 57], to self-competition [58]. Researchers observe an increased level of immersion and motivation when competition is present [59, 60]. For example, gamification is commonly used as an effective and purposeful incentive mechanism for users of CSCW systems; such design examples have been used in crowdsourcing [55], innovation communities [54] and other platforms. Yu et al. [61] combined intrinsic incentives, generally associated with collaboration, and extrinsic incentives, usually associated with competition, in several experiments on a crowdsourcing platform and found that both were important in motivating participants; however, some incentives could potentially undermine others. Similarly, Tausczik and Wang [54] examined open innovation contests on Kaggle, and found that only a small percentage of participants, mainly ones doing moderately well in the contest, shared code. They found that sharing code only improved individual, and not collective, performance. Tausczik and Wang recommend careful consideration when combining these approaches which can lead to greater benefits than using either alone. Another example from Hutter et al.'s work [52] focuses on the simultaneous combination of collaboration and competition in community-based design contests, where contestants are encouraged to communicate with their competitors and found that communititors (i.e. people who collaborated and competed) won the design contest and earned the most awards.

Because of the strength of online community ties, researchers argue that extrinsic incentives, such as winning a contest, are no longer the only motive to participating in organizations and events. There are additional intrinsic incentives related to community building, which, in turn, increase participation and enhance the quality of work submitted. Hutter et al. name this phenomenon "communitition" based on a similar concept in business named "co-opetition." [52] A more recent example comes from Morschheuser et al.'s work on the concept of cooperative gamification, a structure requiring positive goal interdependence between players, which they suggest could be a promising approach for crowdsourcing and other CSCW systems [55].

Distinct from prior work focused solely on competition or collaboration, we study both competitive and collaborative OSINT organizations to compare and contrast their strengths and weaknesses. In particular, we focus on their background and motivations, investigative processes and roles and responsibilities, comparing and contrasting their training practices and regimens, and their division of labor.

Chapter 3

Method

3.1 Recruitment and Participants

When recruiting participants, we aimed for a breadth of domain applications and basic social structures of the investigations. We identified a number of organizations and events that carry out open source intelligence investigations in various domains (e.g., international crime, environmental issues, national security and public safety, human rights violations, theoretical investigations) and social structures (i.e., competitive or collaborative) (see Table 3.1). Some of the organizations conduct real-world investigations, usually in partnership with another entity (e.g., law enforcement, media, NGOs, governments) while others create theoretical investigations, in the form of Capture The Flag (CTF) competitions or quizzes posted online, either using fabricated data or existing public information.

We sought to recruit, through purposive and snowball sampling [62, 63], at least two participants with different roles (i.e., organizer or contributor) from each organization and/or event. We began recruitment with purposive sampling, through direct email invitations of multiple organizers and/or participants who publicly mentioned belonging to one of the selected organizations or events, and continued with snowball sampling to include other organizers or contributors within their organization or event. Participants were compensated ($50 Amazon gift card) for taking part in our research study.

In total, we interviewed 14 participants (P1–P14), some of whom fulfilled the roles of both organizers and contributors in either different organizations or across different investigations carried in the same organization. The participants represented 14 different organizations and events (O1–O14). We mainly focus on the participants' experiences in nine organizations for which we recruited at least two participants (O1–O9).

The participants' locations included Asia (n=1), Europe (n=4), and North America (n=9). The participants identified mostly as men (n=9, n=5 women, n=0 nonbinary), and their ages ranged from 26 to 55 years, with a majority (n=6) falling in the 46 to 55 age range. Most participants were initially recruited based on their organizer role (n=12). However, during the interview process, some (n=4) explained either taking on a contributor role in some cases within their organization or taking part as a contributor in a different organization.

While all participants self-identify as open source investigators or having ties with open source intelligence investigations, their backgrounds range from security consultants (n=5), to journalists (n=5), including investigative journalists (n=2), to intelligence analysts (n=2), to geospatial analysts (n=1), and graphic designers (n=1). More details are provided in Table 3.2.

3.2 Data Collection

Participants completed a consent form and a demographics pre-survey, with Institutional Review Board (IRB) approval. We conducted semi-structured interviews between October 2020 and January 2021. Each interview was conducted remotely over Zoom and lasted a maximum of 60 minutes. During the interview, we asked the participant's about their professional background, their relation to OSINT, their motivations and definition of success when conducting OSINT investigations, and their investigative process, including their strategies

Table 3.1: Organization Codes and Descriptions *we present the overarching social framework based on our initial observations during the organizations' selection for recruitment, our understanding changes during the interviews, showing more complex frameworks, with collaborative organizations employing competitive concepts and vice versa.

Organization	Social Framework*	Domain Application
O1	Competitive	Jeopardy style contest (CTF) with changing themes and fabricated data
O2	Collaborative	National security and public safety investigations
O3	Collaborative	Human rights violations investigations
O4	Collaborative	Injustices and crime investigations in Africa
O5	Collaborative	War zones, human rights violations, and criminal investigations
O6	Competitive	Jeopardy style contest (CTF) investigating the lives of real volunteers
O7	Competitive	Jeopardy style contest (CTF) investigating missing person cases
O8	Competitive	Geo-location Quizzes
O9	Collaborative	Child trafficking and exploitation investigations
O10	Collaborative	Economic and cyber crime investigations
O11	Collaborative	OSINT news and trainings
O12	Collaborative	Cybersecurity and OSINT training
O13	Collaborative	Investigations for domestic violence victims
O14	Collaborative	Corporate social engineering investigations

for collecting, verifying, analyzing, and when applicable, preserving and disseminating, the information. Following that, we inquired about their investigations' social structures; we asked about the coordination of individuals during the process, their roles and responsibilities, and their typical tasks, as well as their technology usage and needs. While the main focus of our study is the social structure of these OSINT investigations, we believe that asking participants about their entire investigative process helps us understand the broader view of how different stages of the investigation are conducted and how different aspects of the investigation come into play which ameliorates our understanding of the social dynamics in action [64]. Participants P6 and P7 were interviewed simultaneously; all other participants were interviewed separately.

All interviews were audio and video recorded with participants' consent. Automated transcripts were generated by Zoom and manually corrected line-by-line. Throughout all interviews, we also maintained typed notes. All participant and organization names have been anonymized.

3.3 Data Analysis

We used a theoretical thematic approach to analyze the interview data [65]. Given that we are mainly interested in the social aspects of OSINT investigations, Braun and Clarke's qualitative methodology allows us to provide a more detailed and nuanced analysis of such group of themes within the data [65]. We used the Dedoose software to carry our qualitative analysis, creating a code tree based on themes extracted from our research questions and previous discussions between authors. Some initial themes included "division of labor structure", "organizer's support for contributors", and "training or practice regimen". As the analysis progressed, all authors periodically discussed observations about the data and iterated through the codes to capture interesting nuances and themes. Some of the later themes added included "power dynamics outside of the organization", "community ties'', and "contributors' support for other contributors".

We highlighted and annotated each interview transcript with one or more appropriate codes, resulting in a total of 1207 excerpts and 3652 code applications. The final code tree contained 44 codes, consisting of 7 main codes, and the rest being child codes.

3.4 Limitations

The field of OSINT and its domain applications are vast. While we attempted to capture some of that breadth through our recruitment techniques, we were unable to capture the totality of domain applications or investigative social structures in the organizations we sampled. Despite striving for a gender balance across our participants, we were unable to recruit more female participants. We also recognize an imbalance in the roles of participants recruited, having more organizers, as their association with their organization is generally made public.

Table 3.2: Participant Demographics. *We used an open-ended question to ask participants what gender they identified as; we received two response types: ”man” (M) and ”woman” (W). **O/C denotes participants who assumed both the role of an organizer and the role of a contributor, either in different organizations or across different investigations in the same organization.

P	Gender*	Age Range	Location	Role(s)**	Organization(s)	Occupation
P1	M	26-35	Asia	Organizer	O1	Security Consultant
P2	M	46-55	North America	Organizer	O2	Intelligence Analyst
P3	W	46-55	North America	Organizer	O3	Journalist
P4	M	-	Europe	O/C	O4/O5	Investigative Journalist
P5	M	26-35	North America	Contributor	O2/O10	Intelligence Analyst
P6	W	46-55	North America	Organizer	O6	Security Consultant
P7	M	46-55	North America	Organizer	O6	Security Consultant
P8	M	46-55	North America	O/C	O7/O11/O12	Security Consultant
P9	M	36-45	Europe	O/C	O8	Journalist
P10	W	36-45	Europe	Organizer	O8	Journalist
P11	M	26-35	Europe	O/C	O4/O5	Investigative Journalist
P12	W	36-45	North America	Contributor	O7/O8/O9/O13	Graphic Designer
P13	M	46-55	North America	Organizer	O9/O14	Security Consultant
P14	W	26-35	North America	Organizer	O3	Geospatial Analyst

Chapter 4

Findings

4.1 Motivations and Definitions of Success

The participants' motivations and successes, while encompassing diverse fields and investigations, share common themes, including education, giving back to the community, policy and social change, combating criminal activity, and the thrill of solving a case. In terms of education, some want to promote security education, specifically cybersecurity and computer security awareness among the public, as P6 and P7 mentioned: *"It was a fun way of trying to educate people about types of information that were out there [...] so that they have a greater awareness of what they're sharing, what their friends and family are sharing."* Others focus on educating people about their physical and psychological safety when conducting investigations, and teaching them the ethical implications of their work; *"success to us [is] that students have all the tools, including how to take care of themselves psychologically, physically and do it ethically"* (P3). Some participants also want to teach others how to foster their analytical and research skills when contributing to investigations, as P2 stated: *"It is always exciting to know that you're doing something that's having a contribution to national security and public safety; more exciting than that was helping these analysts develop their skills [...] as long as they develop strong analytical and research skills and come out of the program knowing how to effectively leverage social media, then I consider that to be a success."*

Others are motivated by the desire of giving back to various communities by volunteering their time and skills to help victims escape their abusers, find missing people, and rescue trafficked children. They also give back to the OSINT community, from which they have learned before, by enabling newcomers to meet more established members of the community through events they organize or chat rooms they administer. For example, P12 who is an administrator of a Discord server where OSINT community members conduct small investigations said: *”we’ll do kind of a live walkthrough [of an investigation] so new people can see how to do it and how people who have been in OSINT a while think, and we kind of guide them along. It’s almost like you’re teamed up with [...] a mentor.”* Some participants are excited by the dynamic nature of OSINT and the thrill of solving their investigations. P11 said *”it’s like a game, sort of trying to follow [...] that chain of evidence and you get addicted to it. It’s almost like a drug addiction in a way. That rush of adrenaline you get when you find something, you want to do that again.”* Lastly, some of our participants also mentioned policy and social change as a drive to their endeavor: *“once upon a time, I thought success would be ‘you found the bad guys’, now I think it’s about getting people to work together on new issues, [...] something that could last two or three years and end up in policy.”* (P4).

4.2 Investigation and Preparations

4.2.1 Training and Practice Regimen

Some of our participants mentioned encountering initial challenges when attempting to get formal training in the field because of the lack of resources for OSINT training online. For example, P1 mentioned, *“Well, in the starting there was not a lot of materials [...] so I will say that was one of the initial challenges, but I think that also led me to working hard* *in this space."* However, across our interviews, we observe an increase in resources to learn and train specifically on OSINT techniques such as the Berkeley Protocol on Digital Open Source Investigations mentioned by P14, First Draft News' training content mentioned by P9, and the Verification Handbook mentioned by P8 [66].

In addition to this increase in resources, all of our participants indicate adopting some common strategies to train themselves and others in this field. A common strategy among all interviewees is learning by doing. At least five of our participants mention either learning or teaching the basic skills and techniques through formal courses that provide hands on training; P3 explained that *"we did live election monitoring and we had 60 students and a lot of students in our class came and did the live monitoring and kind of got their feet wet doing that."* Others encourage learning by participating in organizations that provide daily trainings such as O8, or competitions, such as O7, or even picking a target of interest and trying to apply some techniques by themselves and with others, as P12 suggests, *"I think outside of just competing in the [O7] events which is good for just learning how to think outside of the box, on the fly, most of what I have learned from has been like blogging and doing my own investigations into random like scams and stuff."* Some incorporate game elements into their training to motivate themselves and others; for example, P10 said *"I know having fun and learning is the best combination, so while I was going to a training, I took some pictures from where I was [...] [then] published [the picture] on Twitter and people started to answer my question. I was always asking the question 'where did I take [the picture]?'"*

Another consensus among our participants is that, because of the range of OSINT skills and applications, OSINT experts tend to be generalists, as P8 put: *"It's like mastering, you know, all of the languages in the world, there are some people that get really good at a lot of them, but most of us, you know, we pick what we need to work on and we master those as much as* *we can, but there's always stuff outside of our area of expertise."* To address this challenge, learning from and with others is a common training regimen. For example, P8 remembered a co-investigator sharing that a sea has a higher level on the horizon than an ocean and *"his experience helped me get better at validating and verifying [image geolocation]"*. As another example, P6 spoke excitedly about a competition she and P7 organized where *"the people who were competing are standing around asking each other, 'well, how did you get that answer?'[...] They're explaining and sharing the different open source sites that they found that other people didn't know anything about."* P10 heavily encourages people who participate in O8's quizzes to share their various approaches to finding the answer, as the goal is *"learning from others and working together with others."* We observe through these experiences that learning happens not only through collaborating with other investigators but also through competing against them. Some participants also share their knowledge through blog posts to a larger public, as P12 laughingly shared: *"Somehow I fell into like... I'm like the ship OSINT person, or marine OSINT. I wrote a blog one time and now everybody contacts me about it and so I've kind of become this person who just like absorbs marine OSINT, or maritime information but it's not easy to find."* Through sharing her acquired expertise in that specific area with the broader public, P12 became a recognized expert in the OSINT community.

Participants emphasized several important aspects of training one's own self and others. P2 and P14 highlighted that the OSINT techniques used are completely dependent on the type of data collected and it is important to understand the scientific foundations of the technique when, for example, applying social science processes to verify information, *"[as] if you were looking at the primary source in historical research"*, (P2) or conducting a *"geospatial analysis"* (P14). However, learning and practicing the techniques is not enough; OSINT investigators need to understand the ethical and legal frameworks around their usage of these techniques, and practice security procedures. P2 recommended to train all team members in the appropriate legal guidelines, such as 28 CFR (Code of Federal Regulations Title 28) which regulates the collection of *”information on U.S. persons”* when working with law enforcement, and *“some basic operational security procedures that are meant to protect themselves as well as the work they’re doing [like] crash courses on VPNs, virtual machines and the dark web”*. P9 focused on ethical questions such as: *“‘Should I try to reset someone’s password to find out if he’s registered at this platform?’”* Lastly, participants emphasized critical and analytical thinking skills and the “spirit” of OSINT rather than focusing on tools and techniques; in P6’s words, *“it’s more about teaching and reinforcing a mindset than it is teaching and reinforcing techniques because the techniques will absolutely change.”* We delve into more details about this perception of tools in Section 4.2.4.

4.2.2 Investigation Preparation

Across interviews, we observe that all organizers formulate an investigation plan. For some, the plan is detailed and robust, and for others more minimal, depending on the involvement of a partner or a client in the investigation, the number of contributors or investigators, the timeline, and the potential target. P3 recounts this process:

> *any investigation we do, whether it’s sort of quick turnaround or longer term, we try to have an investigation’s plan and so that [...] is actually written out and it involves aspects of, kind of, ‘what is our objective? What are, you know, some of the risks involved with this cybersecurity-wise or resiliency-wise? [...] What is our objective with our partner? What’s the expectation of the partner? What will the deliverables be?’ And then ‘What are the steps along the way to get there? What’s the capacity of our team? Do we have the right language skills for this?*

Do we have the tech skills?

P14 pointed out, however, that such a plan might not always be followed if *"an opportunity comes up"*. For organizations or events (such as O2, O3, O13, and O14) that tend to conduct most of their investigations with partners or clients, these associated or outside organizations or individuals will sometimes provide an investigation topic of interest, and even some baseline information, which help kick-start a concrete investigation plan. *"We have a general idea of what organizations were interested in"*, P2 said. *"[T]he other factor that determined where we looked really had to do with whether we thought exploiting social media would yield useful information or actionable intelligence."* P13 added that, when hired for an investigation, the client provides the name and email address of the target, *"so we know we're kind of targeting the right individual."*

For some organizations, especially the ones that construct investigations either as simulations or using existing public information (such as O1, O6 or O8), the pre-investigation stage involves the fabrication of such data and its dissemination online. Some develop an overarching theme tying the different challenges or questions of the investigation, forming a *"network of decoy companies"* (P1), others just snap the right picture — *"now I already walk around and I spot something, and I say 'perfect for a quiz'"* (P10). For example, O6 recruits volunteers who act as targets for the investigation which takes the form of a Capture The Flag (CTF) competition. Flags, in this case, are questions about those targets, and the answers are provided by the targets themselves during the pre-investigation stage. P6 and P7 explain that they research those answers themselves and rank the difficulty of finding them, assigning them different point values for the CTF.

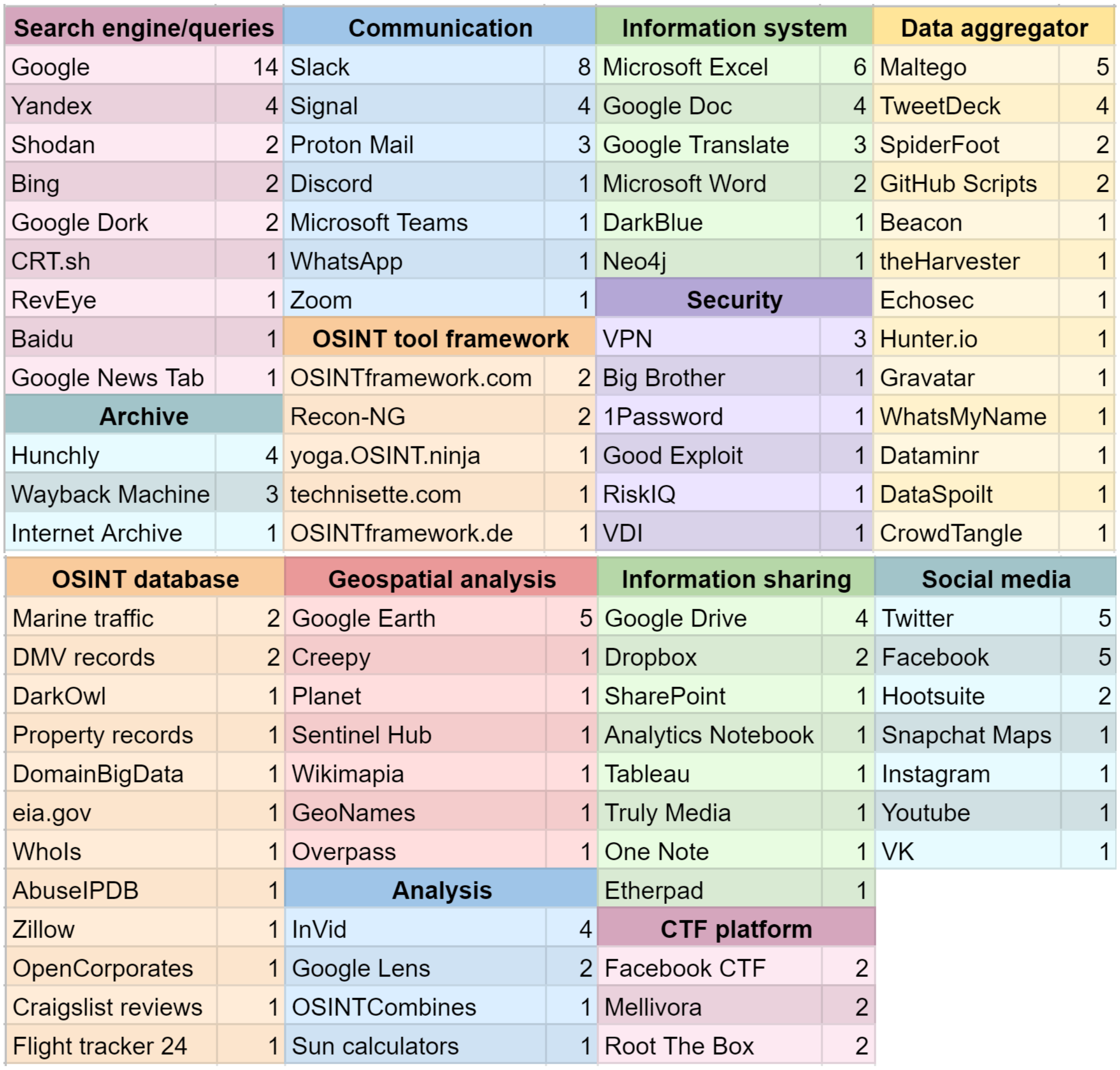
The pre-investigation stage defines the structure of the investigation, including the extent to which different stages, such as preservation of the information or its publication, are needed. For example, O3 attempts to archive all of the intelligence collected as digital evidence for potential legal proceedings, while O6 omits the entire preservation stage since its objective is only to provide a simulation for the duration of the CTF. More importantly for our study, these preparations shape the involvement of different individuals in the investigative process, the possible division of labor, and the rules and training individuals will have to abide by. Participants describe these preparations as an iterative process, learning from previous investigations they conducted. P3 and P14 both described O3's pre-investigation stage as constantly evolving and reliant on the students they recruit in the organization. For example, P14 said that *"we constantly are changing how to do it because it's always a lot to teach. Initially we got people in teams right from the outset, and we had, I think, six or eight different teams, on six or eight different investigations, it was just full on."*. More recently, they use a different model with far fewer teams. P3 mentioned adding a staff member to all student teams, as they have learned that entrusting an entire investigation to a student leading a team of students can be overwhelming. For P1, the iteration happens when creating the different CTF challenges and realizing that players were catching on to P1's ideation process, which helped players solve the challenges faster. In order to avoid this, P1 started having *"different minds coming in"* and bringing diverse thinking to the creation of challenges.

4.2.3 The Social OSINT Cycle

As mentioned previously, the structure of the open source intelligence cycle consists of: content discovery, verification, preservation, and publication. We observe that many challenges arise from the content discovery and verification stages as they often require the use of various tools and techniques and the synthesis of all data into a cohesive whole from which actionable intelligence can be extracted. Investigators try to remedy those challenges by combining the wisdom of different experts or other OSINT community members for these stages through crowdsourcing and competitions. As P11 shared, for one of the investigations conducted by O4, *"we needed more people, and so, we ended up bringing in 15 people all together, working, but most of them, it wasn't full time... it was quite intense but then once the findings were made, then they don't have to work on it."* These investigators are only solicited part-time for their various domain expertise on specific verification tasks that become an obstacle for the main investigators, such as geolocating and chronolocating videos in P11's example. In addition, O1, O6, and O7's CTF competitions only include the content discovery and verification stages. After findings are finalized, competitors neither archive nor publish them publicly. O1 and O6 discard the data after the winner is announced as their goal is mainly to provide an avenue for people to develop their skills; however, O7's organizers share the intelligence collected with law enforcement and encourages contributors not to discuss their findings online.

4.2.4 Tool Perceptions

All of our participants mentioned using tools and technology to conduct their investigations, as listed in Figure 4.1. The tools mentioned during our interviews are used for content discovery, organizing and visualizing information, analyzing and archiving the data, and for security purposes. For the content discovery stage, we observe the usage of multiple search engines such as Google (n=14), Yandex (n=4), and Shodan (n=2), social media platforms such as Twitter (n=5) and Facebook (n=5), OSINT databases such as DMV records (n=2) and Flight Tracker 24 (n=1), as well as, data aggregators which gather related information from different sources (e.g., Maltego, SpiderFoot) or from the same source (e.g., TweetDeck for Twitter data, CrowdTangle for Facebook data). In order to organize, visualize and derive intelligence from the raw information gathered, participants used information systems such as Microsoft Excel spreadsheets (n=6), Google Docs (n=4), and Analytics Notebook (n=1), and



Search engine/queries		Communication		Information system		Data aggregator	
Google	14	Slack	8	Microsoft Excel	6	Maltego	5
Yandex	4	Signal	4	Google Doc	4	TweetDeck	4
Shodan	2	Proton Mail	3	Google Translate	3	SpiderFoot	2
Bing	2	Discord	1	Microsoft Word	2	GitHub Scripts	2
Google Dork	2	Microsoft Teams	1	DarkBlue	1	Beacon	1
CRT.sh	1	WhatsApp	1	Neo4j	1	theHarvester	1
RevEye	1	Zoom	1	**Security**		Echosec	1
Baidu	1	**OSINT tool framework**		VPN	3	Hunter.io	1
Google News Tab	1	OSINTframework.com	2	Big Brother	1	Gravatar	1
Archive		Recon-NG	2	1Password	1	WhatsMyName	1
Hunchly	4	yoga.OSINT.ninja	1	Good Exploit	1	Dataminr	1
Wayback Machine	3	technisette.com	1	RiskIQ	1	DataSpoilt	1
Internet Archive	1	OSINTframework.de	1	VDI	1	CrowdTangle	1

OSINT database		Geospatial analysis		Information sharing		Social media	
Marine traffic	2	Google Earth	5	Google Drive	4	Twitter	5
DMV records	2	Creepy	1	Dropbox	2	Facebook	5
DarkOwl	1	Planet	1	SharePoint	1	Hootsuite	2
Property records	1	Sentinel Hub	1	Analytics Notebook	1	Snapchat Maps	1
DomainBigData	1	Wikimapia	1	Tableau	1	Instagram	1
eia.gov	1	GeoNames	1	Truly Media	1	Youtube	1
WhoIs	1	Overpass	1	One Note	1	VK	1
AbuseIPDB	1	**Analysis**		Etherpad	1		
Zillow	1	InVid	4	**CTF platform**			
OpenCorporates	1	Google Lens	2	Facebook CTF	2		
Craigslist reviews	1	OSINTCombines	1	Mellivora	2		
Flight tracker 24	1	Sun calculators	1	Root The Box	2		

Figure 4.1: Visualization of frequency of tool mentions by participants organized by tool category.

they shared that information between each other using information sharing platforms such as Google Drive (n=4), Dropbox (n=4), or SharePoint (n=1). Participants also mentioned more specialized tools they use for different analysis strategies such as Google Earth (n=5) and Sentinel Hub (n=1) to geolocate photographs, videos, or track airstrikes or other events. Other analysis tools include InVid (n=4) which allows investigators to verify and analyze videos through their metadata, copyright information, and other features, and Google Lens (n=2) which allows investigators to visually analyze images and reverse image search them. Participants also preserve their gathered data and intelligence using archiving tools such as Hunchly (n=4) which preserves screenshot of every page visited during the investigation or the Wayback Machine (n=3) which allows investigators to archive content while making it available to other investigators.

Figure 4.1 lists a total of 90 tools mentioned by our 14 participants. Two of the reasons for this large amount of different tools used are the topic of the investigation and the nature of the data gathered which dictates which analysis strategies will be needed, as P11 stated: *"Depending on, you know, what topic we're investigating or what we need to do, we'll have a [different] set of tools."* P14 added that she relies more frequently on flexible tools *"that I can repeatedly use in an investigation, something that can really leverage those data points, with a range of different information sources."*

While all of our participants use tools and technology to conduct their investigations, most expressed that they are careful in approaching tools. Because of the dynamic nature of the open source intelligence field and the constant changes in regulations, restrictions, and even layouts on different online platforms, participants do not want to over-rely on a technology that quickly becomes obsolete. P13 shared that *"APIs change so often,[...] I found tools to be less useful than doing it manually so that presents a problem cause a lot of OSINTers [...] rely on tools, which means that they get faulty or bad data."*

Many of our participants shared their slight disappointment when newcomers to their organizations, or the OSINT community in general, seem primarily attracted to the tools: *"I see these new people coming in, and they want to know all the tools, 'what are all the tools?', 'what tools should I use?', 'what do I need to know?' and they don't learn the trade craft behind it"* (P12). P3 added that students join O3 thinking *"'Oh, this is a tech heavy thing' but it's really not. It's about fact finding [...] that part of it has to be really emphasized in any class."*

P2 agreed and shared that he does not spend much effort teaching students any particular tool set as they will definitely change, and instead focuses on the tradecraft which does not. Tools entail various limitations for our participants; many become obsolete, lack flexibility across online platforms or analysis techniques, or do not uphold certain investigations' legal standards. Some are also expensive, or come with a steep learning curve. Focusing on them can prevent a new member of the OSINT community from fostering essential and portable skills, such as critical and analytical thinking.

Despite these cautionary tales, participants find substantial value in using tools to coordinate between the many individuals involved in the investigation and to support their organizations' social structure. We observe that participants successfully employ general-purpose, and usually open source, tools and adapt them to their investigative needs. Many participants, including P11, P12, and P13, cited Slack, Discord, and Signal group chats and channels as valuable communication tools. P12 explains that it can be a beneficial way to receive quick feedback on findings, or bounce off ideas, and that she *"thrive[s] in that kind of situation."* P6 and P7 go even further, saying that their CTF contestants sometimes prefer text chat communication over voice conversations during the competition, as it may allow them to stay focused during the time-sensitive event: *"We had this one team of three people, and they sat at a table near us the whole time[...] with their earphones. I mean they had music going. And they were just staring for three or four hours, all three, and I don't think they ever talked to each other... Yeah, they were really intense."*

Using OSINT also requires being able to share that information between investigators. Participants described using SharePoint, Google Drive folders, and spreadsheets, for this purpose. P11 shared: *"we rely quite a bit on spreadsheets ourselves, so we do a lot of Excel spreadsheet [work], all the time. Especially one of my colleagues, [name] [...] For each investigation, he does what we call '[name] epic spreadsheet', because it's a huge spreadsheet* *with [...] every video of an incident [and its location, chronolocation]..."* In addition, P3 and P11 also mentioned the Wayback Machine as a means to archive data that can be retrieved by others at a later time. However, cross-platform sharing, especially if investigators have to use different tools for data analyses, can become somewhat unmanageable. P8 explained, *"I can't tell you the number of students that come into class that tell me 'hey, I have to use One Note and it sucks,' 'I have to use Microsoft Office and it sucks,' 'I use Etherpad,' 'I use a Google Docs,' 'I use a spreadsheet,' 'I use a mind map.' There's all these different ways of documenting and yet none of them is great for sharing."* Facing these challenges, some participants described building their own tools to fulfill their investigative needs. For example, P10 said *"I developed [a collaborative platform for the analysis and verification of digital content] with my team, and also other organizations use it. So we can also collaborate on this together."*

The OSINT ethos applies to our participants' perceptions of the tools they choose to use. Many participants valued transparency in their tools, preferring tools that clearly explained their working process. For example, P14 explained validating the reliability of a tool before using it:

> *"whether you're, like, say, sampling a whole lot of social media profiles, what has been left out? Can the source be altered? Is there a hash code that can be attributed to the evidence but also the process? [...] there's all sorts of AI and advanced tools that I think pass over a lot of power to the tool [...] and it's really important to getting back to one of the categories of classification, what's being left out?"*

P8 elaborated on the importance of this transparency when comparing two versions of a tool created, a command line version and a web application version: *"You run it from the* *command line as a Python tool and you'll get the results. If you use the website out there, it'll do the same thing, but because of something called CORS, Cross Origin Resource Sharing, your web browser might give you false negatives."* Because of this aim for transparency of the process, some participants preferred open source tools that are flexible and modifiable. Many of them would rather build their own tools, as P1 shared: *"I really don't want to work with tools because they work in their own, you know, defined manner [...] [so I] prefer writing my own [tools] as and when things are required."*

When asked about their tools "wish list", almost all participants had at least one idea to share. Many proposed tools that would optimize some of their investigative tasks such as collecting and/or organizing data, cross referencing data points across multiple sources (e.g. finding a user's profile on various social media sites), or performing certain analyses on collected data (e.g. social network analysis, image manipulation...). For the purpose of our study, we focus on technological ideas that would support or improve the current social structure of OSINT investigations. P8 eagerly shared his need and vision for an open source *"case management software dedicated to open source intelligence"*: *"[It] would absolutely be a winning piece of software because, I mean, there are so many people [...] one of the things that they say is, you know, 'how do I do OSINT in a team, in a group?' because... 'How do I decrease the redundancy?' [...] It's that last piece of managing the entire case that's really, really missing."* Another example is P4's idea for preserving collaborative open source intelligence that has already been generated on certain platforms by OSINT experts: *"[a tool that would] scrape the entire Twitter for Google Maps mentions and put them on a map. [...] So that all the history of geolocation work that I've done on Twitter over the past five years could be then plotted on a map in collaboration with every other open source investigator"* (P4).

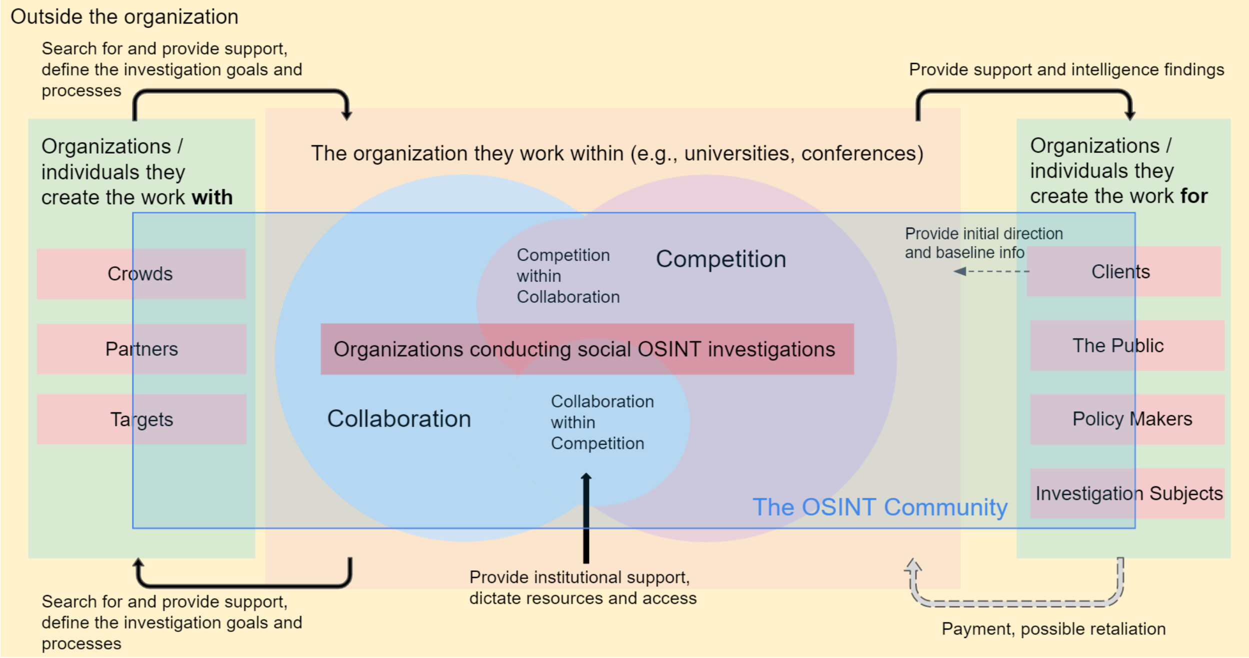
4.3 Social Dynamics

While we recognize that open source intelligence work can be done by a lone analyst, participants still described ways they work with other people as critical to the success of their investigations, and they provided multiple examples of how their investigations are socially structured, employing various collaborative and/or competitive concepts and formations, and various scales, ranging from smaller teams to larger online communities.

In the following sections, we start by presenting the social structure of these investigations within the multiple organizations, focusing on the collaboration, competition or a combination of both processes implemented or observed by the participants, and how these frameworks support the contributors throughout the investigative process (Fig. 4.2). We then describe the presence of entities outside the participants' organizations, such as other organizations they work within, with or for, and how their presence or involvement can affect the investigative process.

4.3.1 Roles and Responsibilities within the Organization

Each organization features different positions with different roles and responsibilities, some organizations are more rigid and hierarchical in their structure of positions, while others are more flexible and less formal. For example, O2, O3, O4, and O9 feature organizer positions with specific titles such as *"Lab Director"*, *"Team Coordinator"*, and *"Executive Producer"*. O1, O6 and O8 do not feature official position titles other than *"organizer"* and *"participants."* O7 features a *"judge"* position in addition to the last two. This difference in hierarchical structure partly depends on the overarching framework, collaborative or competitive in nature, that the organization chooses to implement. We dive into more details about this in Section 4.3.2.



Outside the organization
Search for and provide support, define the investigation goals and processes
Provide support and intelligence findings
Organizations / individuals they create the work **with**
The organization they work within (e.g., universities, conferences)
Organizations / individuals they create the work **for**
Crowds
Partners
Targets
Competition within Collaboration
Competition
Provide initial direction and baseline info
Clients
The Public
Policy Makers
Investigation Subjects
Organizations conducting social OSINT investigations
Collaboration
Collaboration within Competition
The OSINT Community
Search for and provide support, define the investigation goals and processes
Provide institutional support, dictate resources and access
Payment, possible retaliation

Figure 4.2: Diagram depiction of the Social Dynamics involved within and outside the organizations conducting OSINT investigations. These organizations sit in the middle of the diagram, in an interplay between collaboration and competition. We present organizations and/or individuals they work within, with and for and the different interactions/dynamics between them. We overlay the boundaries of the OSINT community. (Dashed arrows denote possible, but not certain, interactions.)

Based on the experiences shared by the participants, we notice that the longer an individual has been involved in the organization, and as their knowledge in the OSINT field expands, the more opportunities for that individual's role to evolve and become increasingly central to the organization. *"We all started out as participants in there,"* said P9, who became an organizer in O8. In O8, we observe the presence of lurkers, and their presence is appreciated by the organizers. P10 explained that *"what is nice about Twitter is you don't need to have an account to just watch and I know that a lot of people are following [O8], but they are too shy to participate, but they are learning from just watching it and reading, and this is also beautiful."* She further explains her reluctance to switch O8 from Twitter to another platform from fear of inadvertently excluding some lurkers. Understanding the OSINT community as a community of practice, we notice that Legitimate Peripheral Participation [Lave] applies here; newcomers start with low-risk tasks and slowly gain a level of mastery and become central to the community.

We also observe, across many interviews, the presence of individuals who use an online identity or internet persona with a pseudonym to participate in investigations and organizations. Using a moniker and preserving one's anonymity online, a very common practice in the infosecurity and cybersecurity communities, does not prevent them from belonging to the OSINT community; in fact, some have become central members of the community and have been mentioned by many of our participants during the interviews as fully-fledged and influential members who are valued for their work and publications. P10 adds that *"it doesn't matter what your Twitter name is or where you're coming from"* because *"everyone has the same goal, to find the solution."*

4.3.2 Collaboration Within Organizations

Collaboration traditionally tends to be more present than competition in OSINT work. Explaining the reason behind O2's more collaborative structure, P2 said: *"part of it is a recognition that almost all the work done in the U.S. intelligence community now, analytic work, is done in a team-based environment, so we want [the students] to be familiar with working in a team-based environment and develop these team-working skills, but it's also just a sound analytic practice, that if you have multiple perspectives, it helps eliminate or at least counteract cognitive bias."* P3 elaborated on collaboration as a cultural aspect of the open source community: *"[the] open source community generally is super collaborative and that's what I love about it and I think anything we all do in this space should be emulating that; collaborating with different partners, collaborating across disciplines, covering different* *sectors...etc."* This emphasis on collaboration may be due to the fact that the field of OSINT cannot be mastered by a single or small group of individuals, and therefore, working with others and accepting help from others leads to more successful investigations. As P11 said *"just working on your own, you end up missing a lot of information [...] you can spend hours trying to investigate the story, but [...] you talk to someone else and other open source investigators and they might have another idea that you haven't thought about."* P5 added that being able to leverage other people's subject matter expertise is *"a giant skill whenever it comes to open source investigation."*

Our interviews support that organizations that choose a more collaborative framework overall, tend to feature a more defined structure and hierarchy in their roles and responsibilities, such as having official titles for staff positions, and team members reporting to a team leader. There also tends to be a considerable need for coordination between investigators, mainly to reduce the redundancy in effort and advance the investigation more efficiently. P14 explained that *"It usually works best when there's a professional staff member with a [Graduate Student Researcher] or with an undergrad to define the team and to help structure the tasks, and to make decisions about what the parameters would be."* We speculate that these organizations define a clear structure and hierarchy in order to support the large number of individuals usually involved in their investigations, with teams ranging from four or five individuals to upwards of 30. However, those bigger teams tend to split up into smaller groups; *"When you're getting together a room of 30 people trying to document every airstrike in [country], sometimes you need to break them down in teams of five"* (P4). P4 also pointed out, however, that occasionally investigations can happen spontaneously, without much structure and/or coordination discussed; *"it's been that investigators from Twitter, that just had a mutual drive and a mutual passion and say, 'oh my god, let's get these people', and there's no roles discussed there's no hierarchy or team leaders, it really is just a group of people that want to* *do good in the world.”*. P2 also warned that having defined hierarchical roles can sometimes lead to certain challenges, such as *“a team lead who is doing all the work themselves or isn’t providing sufficient direction or is not allowing people to participate as much as they should.”*

On the other hand, these organizations tend to feature less explicit or strict rules about how to carry out specific tasks, which tools or techniques to use, or which sources of information to explore, and rely more heavily on the expertise and creativity of the contributors. As P3 recounted: *“we were empowering the students to [be] the experts and to be the innovators and that [...] was a great, great model because the students didn’t go look to us and go okay, ‘help me figure this out.’ ”* P14 elaborates on some of the benefits of empowering contributors to be creative: *“we like that atmosphere that everyone feels like it’s a little bit more free [...] by seeing what’s possible they then start to realize that [the information] they’re sitting on is really valuable.”*

Strong communication In more collaborative settings, good communication is a requirement to enable individuals to work together effectively for lengthy investigations, especially when those individuals do not have the same domain expertise or background. In keeping with that, some participants mention that they attribute more effort to the process over the product. For example, P9 attributed more importance to contributors sharing their methods in solving the quiz than the correct answer. P3 elaborated that the importance is in showing *”these are the steps that I went through, and this is what we can show, and this is what I know and this is what I don’t know. [...] That transparency is critical to the open source process.”* When working with other investigators, the methodology needs to be transparent, with detailed and structured documentation in order for all investigators to be of the same mind and communicate more precisely. For example, P5 explained that his target profiles are very robust and thorough when working in a team environment, but very minimal and only comprised of *"little notes"* when he is working alone.

Strong bonds Along with robust communication, many participants value strong bonds between investigators. There is a push for individuals investigating together to foster a friendly relationship which improves the quality of their communication, of their resilience and in turn of their work and their sense of community. P4 mentioned building those strong bonds as the *"perfect way"* to work on investigations. Similarly, P5 recalled structuring the tasks of one of O2's investigations based on his teammates' preferences, strengths and weaknesses: *"I think it takes a good amount of knowledge on the people that you work with [to do that]."* P9, P10 and P12 added that contributors who communicate with others will slowly build *"some sort of relationship"* (P9) and slowly become *"part of the family"* (P10). Some participants, including P4 and P11, mentioned that breaking down bigger groups into smaller teams aids in the creation of such bonds, and in keeping communication lines open between people. P11 shared that during one of O4's big investigations, as more contributors were added to the Slack channel, public conversations in the main chat were decreasing, while private messages were increasing. As a solution, he split this *"big collaboration"* into smaller groups so *"people are then more comfortable to talk and express"*. P4 also explained his preference for smaller teams: *"you build a bond with people as well which is important, rather than just 'you do this', 'you do that',[...] it's more 'hey, we're out to do well in the world and we have a small group of dedicated people that can work together.'"* P8 adds that smaller teams can also be beneficial in harnessing the power of sole performers or lone analysts in a collaborative setting by having *"teams of one"* encompassed in a workflow with bigger teams.

Collaboration within competition In a number of organizations, collaborative strategies appear at a smaller scale, within a more competitive framework. Especially in team-based competitions, as implemented by O6, O7 and sometimes O12, members of the same team have free access to each other's skills and domain expertise, and emphasize strong and constant communication because of the contests' time constraints. For example, P12 recalled that during O7's CTF, *"they give you like seven people that you're looking into, so each one of us, [in the team], will pick one person and we'll work on them for like an hour and then, if we hit a dead end we'll switch people just to kind of keep it going."* Her team also sets up a Slack or Discord channel for each target and each teammate *"will post in all [the target] details, what they're finding and other people will kind of comment on it. So it is a collaborative environment..."*

4.3.3 Competition Within Organizations

Many of our participants consider competition a powerful motivator, keeping contributors engaged in the investigation and focused on the task at hand by gamifying some of the investigative process, and creating a sense of urgency. One benefit of competition, and especially gamification, is motivating people to learn OSINT skills who might not otherwise be interested. P6 and P7, for example, implemented competition in the form of a CTF and *"sort of gamified the education process."* P8 added that this sort of structure is *"getting a huge number of people introduced to [the] OSINT world, it's getting a lot of people interested in investigation, it's getting them into the process."* P1 also mentioned that sometimes companies or other organizations encourage their employees to participate in CTFs and acquire new skills that way. However, a number of participants, including P4, P8, and P11, were quick to warn that this kind of motivation does not necessarily correlate with better investigative results, and there are some things to keep in mind in order to successfully implement competitive strategies in the investigative process. Specifically, P4 shared that, while competition is *"a wonder for projects"*, *"the idea of independent competition is not* *so great because that's what happens in intelligence agencies, they silo information and they don't reach out to each other."* P1 adds that the prize of the CTF also plays a role in extrinsically motivating people to participate, which leads us to suggest that some players may be more interested in the prize than the investigative process and might lead some to cheat their way to the prize.

Competition also creates a sense of urgency that might encourage contributors to work more efficiently. P11 believes that *"a healthy dose of competition can help, definitely, people move faster and... refocus at times because, also again, the open source investigation, you can go down so many rabbit holes... [competition] can definitely push you... As soon as you switch on the competitive mindset, you might be more focused and because you've got the time pressure."* This efficiency can be critical for time-sensitive investigations, such as breaking news events, missing persons cases, and criminal manhunts.

Another, more subtle, form of competition we observe in our participants' experiences is competing for attention or recognition. For instance, some contributors or organizers have the desire to be the first to publish the findings of their investigation, with journalists not wanting to *"get scooped"* (P11), or want to showcase the results of their investigations and the skills they employed,*"natural[ly] wanting to show outputs that different teams have [like stories, reports, or legal memos]"* (P14). This motivation, according to P10 and P11, is sometimes linked to the *"ego of the individual"* (P11), and whether they care for public recognition of their achievements, or can also be a personal challenge.

Nevertheless, a smaller number of participants, while recognizing the benefits of competitive strategies, do not find competition to be a motivator for them. Some prefer the social interactions enabled by more collaborative models. P8, while playing in O7's CTF, shared that *"from the competitor point of view, it was... it was pretty darn isolating, [...] as I already mentioned, I'm a very collaborative person"* and P10 agreed that she is happier when collaborating with others.

In contrast to organizations that support a more collaborative framework, we find that those that support a more competitive framework overall tend to feature a less defined structure and hierarchy in roles and responsibilities within teams conducting investigations. While organizers have defined roles and responsibilities, team members often do not have an official team leader; each team usually has a team name displayed on the scoreboard and decides on their own structure or lack thereof. P6, P7, and P12 all shared that players *"get to choose who's on their team"* (P6 and P7) and *"it's organized as far as the specific teams feel like they want to organize it"* (P12). P12 elaborated, saying that *"the more seasoned teams [in CTFs] have a structure set up where maybe one person is digging deeper into all of the missing people and someone is just doing surface level submissions for points and I think when people have done these competitions a few times, they start to figure out that you have to have a system like that to get the big points."* We speculate that this is partly due to the smaller sizes of teams, ranging from two to four individuals, participating in shorter, more time-constrained, investigations. While sometimes teams of one are allowed in competitive events, according to participants, they are rare.

On the other hand, we observe more explicit and strict rules about which sources of information are acceptable, which tools and techniques are allowed, how information should be submitted and how points are awarded. Rules related to acceptable sources and submission steps aim to keep the competitors focused on contributing useful information during the time-constraint, while other rules related to the point system and tools allowed strive to maintain a level playing field among contestants. For example, P8 mentioned that for O7's CTFs, *"the teams need to submit the URL and why they think it's important, [...] their reasoning or analysis behind it."* P6 and P7 also said that players *"could only have two tries"* for O6's challenges. Many of the participants explain that maintaining a level playing field is important. Sometimes players will *"start, you know, just breaking the rules a little bit"* (P1) or *"hacking the game and making it unfair for the rest of the participants"* (P6). P12 adds that O7's judges may accept submissions differently if the rules are not standardized, which can create some tension during the competition.

Participants who construct investigations as simulations said that trying to emulate real-world investigations and techniques in a CTF-style competitive event can be difficult, and there is a difference between methodologies that are successful in a CTF and in the real world. While P1 tries to create challenges that hone in on real-world OSINT skills, P8 said that *"if [people] take that same methodology that they use to win the CTF, and they try to apply that within a business setting in a real OSINT environment they're going to absolutely fail"* because for many CTFs the goal is *"to submit as many flags as possible, which is different than doing an open source intelligence investigation."* There are also competitive strategies used against other teams that are more specific to CTFs, such as a lack of communication in order to mislead other competitors on a certain team's progress. For example, P6 and P7 mentioned that *"we've had people hold on to flags and drop them at the last minute to drive some of the other competitors crazy."* Some teams employ *"smack talk"* (P7) in an attempt to demoralize other teams.

Competition within collaboration Nonetheless, some competition is valued by participants when incorporated into real-world investigations, and more collaborative settings, as long as communication lines between "opposing" teams remain open and people continuously share their progress and skills with others. Beyond the motivational benefits, competition can help create sounder investigative arguments by having an "opponent" poke holes in the analysis of another investigator, known as "red teaming." P11 explained that, while O4 conducts investigations collaboratively, once the investigation is complete *"we'll bring another* *open source investigator who wasn't working on the story and that person will go through the story with the idea: 'I need to break that story. I need to find a hole in that story. I need to find a mistake.'"* It can also, perhaps surprisingly, help investigators cope with emotionally distressing work facilitating their progress. For example, P4's organization conducts human rights investigations using OSINT methods, and *"to keep that fun we set up little challenges like capture the flag; [...] which is terrible to think of because you're doing human rights cases and you're looking at bombings, but at the same time, if you can gamify that competition, you can get better results."*

Learning through competition Even in these more competitive settings, we find that members of the OSINT community still learn from each other. P6 and P7 were happy to see that during a rundown of the results after their CTF *"some people [said] 'can somebody tell me how you got so and so?' and all of a sudden this conversation went on, and I think it was about DMV records, [...] they were like, 'oh man, that's awesome. I gotta try that technique the next'"* However, P12 mentioned that this aspect is still lacking in O7: *"you don't get a rundown of what everyone has found afterwards, how they found it, because they just give it to law enforcement and a lot of it never gets [shared back with the contestants]."*

4.3.4 Supporting Organization Members

Throughout the investigations, organizers support their contributors in different ways. From our interviews, we observe that organizers train their contributors and provide them with exercises that are designed to guide them through the investigative process and allow them to build not only the skills and mindset needed, but also the confidence to do OSINT work on their own. P11 recounted training others through case studies *"showing them the thread of evidence, [...] how they can go step by step, both about finding the story, but also investigating* *it. That's really helpful because it builds confidence."* P5, as a contributor in O2, shared that cultivating his project step by step and slowly graduating to become a subject matter expert *"was really cool."* P2 and P8 also mentioned that it is helpful for contributors to include their interests in those exercises and projects, in order to motivate them.

Resilience and safety Another dimension in which organizers support their contributors is resilience and safety. Open source work can be long-lasting, frustrating and sometimes even traumatizing. Organizers emphasize the importance of teaching patience, rotating contributors on projects, and promoting mental health by fostering strong bonds within the organization, mandating therapy sessions, or even blacklisting certain areas of the internet. While working on an investigation related to the COVID-19 pandemic, P2 said *"that team was so busy, we started rotating people through it every week [...] it was such an intense work environment, we didn't want people in that for too long."* P3 added that *"user generated content can be dramatic and traumatic in ways just as traumatic as coming into contact with trauma firsthand in some ways because it's so intimate."* In order to reduce the exposure to such content, P4 advised to *"watch it on a phone, watch it in black and white, turn off the sound, don't get immersed into this."* In the case of O9's investigations, P13 shared that all investigative work has to be done on their proprietary VDI (Virtual Desktop Infrastructure) software as it is set up to safeguard the volunteers *"from getting any illegal material by mistake on their computer,"* and encourage restraint; *"everything is being logged and tracked also, we have like Big Brother watching so that way, no one can do anything illegal that we're not able to see."*

Organizers are also attentive to their contributors' needs and suggestions about potential improvements to the investigative structure, as P6 and P7 explained: *"one of the reasons we change things around all the time is because we listen to what people are telling us, and* *we say, 'what would make it better?'"*

Contributors supporting each other Aside from the supportive infrastructure established by the organizers, contributors also support each other by building strong community ties and maintaining friendly relationships. P14 excitedly recounted *"the commitment to and really genuine interest to create a really nice atmosphere with the students. Our students do quite a bit of like getting to know each other, and when we were in person they'd go out salsa dancing."* P3 elaborated on the benefits of community building, citing it as a resiliency method that also helps establish a solid ethical foundation, *"[making] everyone feel, you know, part of this team and connected and don't leave anyone behind."* Other objectives of creating this pleasant environment are to get better at working together, to encourage the participation of different team members, and to learn from each other. P12 mentioned that the community and other contributors have been very supportive of her learning and growing as an open source investigator, and that she *"will always find somebody to bother"* for help. P9 shared that within the contributors *"there are a few people who are really good and they hold back a little bit,[...] because they don't want to spoil [the answer] for the others,"* and instead they support others by pointing them in the right direction.

4.3.5 Outside the Organizations

Having examined the social dynamics and structures within the different organizations our participants belong to, we now turn to certain power dynamics that affect the OSINT investigations from outside the organization. First, we present our observation about organizations they work within, with and/or for and how these entities impact the investigative process, then we present certain ties with communities that receive the results of the investigation, or even in some cases with the targets of the investigations, before showcasing the existence of different cultures within the OSINT community.

Organizations within organizations Some participants mentioned that their organizations are hosted within an overarching institution or event. For example, O2 and O3 are part of two different universities, O8 is hosted on Twitter, while O1, O6 and O7 are hosted by a number of different conferences. As a result, these institutions' or events' rules, regulations and structures impact the organizations' investigations and structure or sometimes dictate the resources they have access to or the contributors allowed to join. P2, P3, P5 and P14 pointed out that all staff members or organizers of O2 and O3 are employees or faculty members of their respective universities, while the contributors are recruited from a pool of qualified students (e.g., students from a specific major or students who have completed pre-requisite classes). P3 stated that, while O2 is open to students from various disciplines, it is based in a department of the university, encouraging a larger number of that department to join O2. One of the challenges surfaced by the participants is *"a lot of also quality control, because students [who join], they become great at it and then they graduate"* (P3). P11 indicated a difference in access to resources between O4 and O5, with O4 having stricter rules when it comes to crowdsourcing the help of the broader OSINT community and sharing investigative information on social media. P1, P6, and P7 all mentioned that they need to coordinate O1 and O6's events with the conference organizers and that participation is usually restricted to the conference attendees.

Organizations they work for As mentioned in a previous section, some organizations tend to conduct their investigations with or for other institutions (e.g., law enforcement, media, NGOs, government agencies). We observe a difference in relationship between entities the organizations work with, and entities they work for. In the latter case, those entities tend to be perceived as *"customers"* (P2) or clients for whom the organization is providing a product or service. These institutions may provide direction or baseline information during the pre-investigation stage; however, they are less involved in the different components, tasks, and intricacies of the investigative process. P2 shared his perspective on working for such entities:

> *[O]ne of the sort of myths of the intel cycle is that the customer will always provide you with specific questions about [what] their needs are. Most of the time, they don't do that. In a lot of cases, consumers of intelligence regard it as a free good. It's just something that shows up magically in their inbox and they don't really give much thought to what comes behind it. So in addition to those specific taskings, we're going to spend a lot of time thinking about 'what is going on in the world that at least should be of interest to our stakeholders, even if they don't realize that it should be of interest to them?'*

Organizations they work with On the other hand, institutions that organizations work with tend to act more like *"partners"* (P3). Even though these partners will still consume the intelligence provided by the organization, they tend to be more involved throughout the investigative process, providing assistance in several stages and helping with training. P3 prefers conducting investigations with partners of O3, saying: *"[NGO name] is great and has been our best partner over time because they have researchers around the world that really need extra support. They know what they want, they sometimes come to us and say, you know 'something's happening in Cameroon, we have five videos. Can you verify these?' and then the students will take a deep look at those."* We discern that participants indicate more balance and compromise happening between the partners' needs and the organization's, than with the clients'. P14 talked about one of O3's partners providing investigative support: *"they have four — which I thought was quite a lot — four different people independently review* *the geolocation, which I think is good, so the pressure isn't on the students and they've got other professionals and contractors."*

Even when organizations are not necessarily working on a certain investigation with another entity, there are community ties between contributors belonging to both institutions which encourage them to share information or provide assistance. P4 and P10, among other participants, said that they follow many other members of different organizations conducting OSINT investigations, which allows them to share or receive leads, ask for help, or even *"have quizzes during lockdown and stuff like that together"* (P4).

Subjects of the investigation Participants shared that, in some cases, the subjects of the investigation are aware that an investigation is being conducted on them and provide their consent, such as in the case of O6 that recruits volunteer targets for their CTF, or O13 that is tasked to assist *"domestic violence victims"* (P12) by restricting their information online. P6 and P7 prepare their *"voluntargets"* by trying to put them at ease and explain the process in detail, while also connecting them to previous volunteers who act as references. P12 provides the victims with regular updates about the investigation, and involves them in the verification process. P5, however, pointed out that he uses alias accounts on social media to be granted access to certain Facebook groups; in this case, subjects of the investigation are not made aware that an investigation is happening. In P13's corporate investigations, while the company is aware that an investigation is being conducted on their employees, the employees themselves are not.

The public Once the investigation is complete, a number of organizations (such as O3, O4, and O5), share their findings publicly (e.g., by publishing a news story or producing a documentary). There are a number of potential consequences, including individuals or policy makers being influenced by the findings, subjects of the investigation being placed in the limelight, other OSINT analysts trying to poke holes in the findings, or even retaliation by outside actors. P4 elaborated on some of the reasons for meticulously reviewing every step of the investigation: *"it's going to be digested by the wizards on Twitter that say, 'well, your geolocation's wrong here, your open source's wrong and your evidence is wrong.' And people in [country] have patriotic open source analysts, happens about [another country] too, [...] that will look at this stuff and say, 'I'm going to take this thing apart.'"*

P4 added that another social aspect to consider is the invisible influence from outside actors that affects what we see and what is being investigated; when speaking about receiving a lead into an investigation, he said: *"it may have been sent to you for a reason of trying to stir the pot... to influence something that's actually already ongoing, and I think that's very dangerous because it's something that's very hard to verify with proof."*

Cultures and values in the OSINT community From a broader perspective, our interviews highlight the existence of different cultures and values within the OSINT community. Different participants placed more or less value the ethical implications of their work, the interdisciplinary nature of the OSINT field and joining experts of diverse backgrounds in this space, and receiving online recognition for their work. We can also discern some of these values latent in their definitions of success and their motivation to conduct OSINT investigations. A number of participants attributed certain values, such as the desire for online recognition, attention, and credit to the *"tech bro"* (P3), or more specifically, the *"BrOSINT"* (P14) culture. P3 and P14 mentioned more men being attracted to this side of the OSINT community because of their driving interest in the technological side of OSINT, and less in the *"human rights"* (P3) side. P11 brought up some of the tensions that can arise from these different values colliding: *"There was like an interesting tension when we* *were working on [investigation title] between open source investigators who wanted to publish. They made the findings, they wanted to publish them straight away [on their Twitter accounts], and obviously my boss was like, 'no, no, we can make a video about it.'"*

On a separate but related front, participants credited people with different values joining a certain organization to the way that organization frames or defines itself and its investigations. P3 exemplified the latter by stating: *"we found when we have classes where we get more men, and we have classes about human rights, where we get all women, so…but it's like framing is important, like, how are you framing this: is it a tech bro thing or is it a human rights thing?"* A similar phenomenon is discussed by Elliot Higgins, founder of Bellingcat, an online open source collective of researchers, investigators and citizen journalists [23]. He noted a lack of (in particular, gender) diversity since the early days of open source investigations and women investigators drawing disproportionate criticism. He drew parallels to similar gender dynamics in the online gaming community:

> Gaming became so central to their lives that some came to expect that human relationships operated by similar rules: if you meet the objectives, you can have what you want. Once life proved more complicated, a faction grew resentful, twisting online camaraderie into a self-pitying fraternity that vented its spite through digital bullying.

Pushing against these attitudes, Higgins had achieved gender parity in his hiring of Bellingcat staff by 2019.

Chapter 5

Discussion

We described the various personal, interpersonal, and organizational factors that shaped our participants' investigative process. We now consider what our results mean for the social structure of OSINT investigations. More specifically, we reflect on division of labor and social structures as well as technology designs that can inform future open source intelligence efforts and benefit members of the OSINT community.

5.1 Reflections on the Division of Labor

5.1.1 Combining Collaboration and Competition

As previously introduced in our related work, combining collaboration and competition can motivate individuals to participate in innovation contests and crowdsourcing tasks, and enhance the quality of work submitted [52, 54]. Our findings show that the open source intelligence community also employs both concepts in different combinations when conducting their investigations. Even in more competitive settings, participants report a desire to give back to the community, or to learn from other members, motives that align with Hutter et al.'s study of collaboration in design contests [52]. However, previous research also demonstrates that competition, and some extrinsic incentives, can sometimes inhibit collaboration, but the degree to which collaboration is inhibited depends on contests' design [54]. These studies recommend attributing an extrinsic incentive to collaboration, by rewarding competitors who exhibit collaborative behavior throughout the contest [52, 67], or by designing positive goal interdependence in the game, making the success of one player positively correlated with the success of another [55]. We propose that these approaches may also benefit the CTF designs of more competitive OSINT organizations we examined, such as O1, O6, or O7, by reducing the amount of duplicated effort by different teams, and incentivizing contestants to share some of their expertise with other teams. We also suggest that fully or partially sharing the answer to a flag with the rest of the competitors after it has been found by a team can reduce the effort in deduplicating the competitors' submissions after the CTF, and discourage teams from siloing useful information.

Correspondingly, we propose that competitive strategies can benefit more collaborative OSINT organizations, as we observe in O2, O3 or O4, by limiting groupthink and inaccuracy blindness, reducing feelings of immersion in traumatic content, and encapsulating certain investigative tasks. Kane et al. [68] posit that prompting collaborators to "exert discriminatory thinking and analysis" towards their teammates' work could help them detect inaccurate information. For example, in our interviews, we found that adopting a competitor's mindset helped members of O4 generate well-grounded investigative arguments and avoid retaliation by outside actors once their investigations' results were made public. Gamifying certain stages of the investigation can reduce feelings of immersion in user-generated content, especially when it is traumatic such as in human rights investigations. While such gamification could help mitigate secondary trauma among participants exposed to upsetting content, such decontextualization risks adverse consequences such as trivializing or misconstruing its meaning. Competition, we found, involves more explicit rules and the encapsulation of a task with clear instructions on how to win. Taken together, these characteristics could increase the potential of certain tasks to be crowdsourced to analysts outside the organization.

5.1.2 Tool Design Implications

Our findings present participants' attitudes towards tools and technology, emphasizing a measured realism when approaching various specialized tools but embracing their use nonetheless. However, we also report their successful use and adaptation of more general-purpose, typically open-source, tools to coordinate their social OSINT efforts. Participants also expressed a need for tools that support or improve the current social structures of their organizations and events, such as a multi-user case management platform. We present a number of recommendations based on these results.

OSINT investigations represent a complex and creative process, requiring the adaptability of a multitude of tools at various stages of the investigation. Given this requirement, we highlight the concept of appropriation in design and relate it to the domain of OSINT investigations. Gonzales et al. [69] surface the importance of designing *appropriable* tools that can accommodate changing workflows and adjust well with other tools being used. In addition, tools such as shared workspaces and collaborative sensemaking systems have been shown to increase awareness of others' activities and progress, benefiting analytical tasks [70, 71]. With these prior works in mind, we propose that tools supporting social OSINT investigations should allow users to define their own investigative process without centralizing all the tasks in one platform. Our first recommendation is a tool that would act as a dashboard for the investigation, allowing users to visualize the current activities being completed, and check their progress status, and as a meeting point where investigators can upload their data from different tools for others to download. In contrast, we also recognize that, in some cases, investigators may not want to share the information they have, as in the case of journalists concerned about getting scooped. Therefore, we recommend that tools promote social translucence into the investigative process without inhibiting the potential for competitive strategies to be implemented by the users. For example, a user would be able to see the investigative stages a competitor has completed and a summary of their findings without having access to the details of the activity they are currently working on or the exact data they have gathered. Platforms such as CrossCheck by First Draft News [72], or Check by Meedan [73] which enable journalists to collaborate on the verification of information, demonstrate that journalists can be encouraged to collaborate, especially on open source information.

OSINT investigators conduct various analyses, employing different tools highly dependent on the data they collect. To alleviate the burden of sharing their data across tools, we suggest an open source standard for tool interoperability at different levels of abstraction [74, 75], facilitating the importation and exportation of data from one tool to another. As an example use case, we propose a mapping software that allows users to export GPS coordinates, then to import them into a flight tracking software to automatically track flights within the vicinity of those coordinates.

Our findings demonstrate that OSINT investigators rely on crowdsourcing content discovery and verification tasks when they face obstacles and they do so by recruiting outside analysts on social media platform, notably Twitter. However, social media content can be ephemeral and eclipsed by newer content, as mentioned by P4. Inspired by our interview responses, we suggest providing investigators with a crowdsourcing platform that would be directly link them to the community of experts on Twitter, for example, while preserving and categorizing their work which would alleviate the burden of registering on a different platform and take advantage of an already thriving social network of experts.

5.2 Making OSINT Investigations More Social

Previous research has shown that scaling and speeding up investigations was possible by involving more people, such as via crowdsourcing [21, 49, 50, 51]. Our findings build on this research by showing that, not only engaging the help of more people, but making OSINT investigations more diverse and more social, increases the benefits to the investigation and to the investigators. Cultural diversity has been shown to improve creative outcomes [76] and decision making [77]. In addition, other research in collective intelligence has demonstrated a positive correlation between the number of women in a group and that group's problem-solving ability [78]. Considering prior research and our findings, we suggest that promoting more diversity in both demographics and domain expertise has the potential to generate more robust results for OSINT investigations. However, prior research has also found team diversity to be detrimental if individuals are not able to achieve common ground [77, 79]. Similarly, our findings show that some tensions can arise from having individuals who hold different values within the OSINT community on the same team, such as between members of the so-called ”brOSINT” culture and other investigators otherwise driven by social justice. These dynamics are further complicated by empirical research showing that in general, men tend to be more attracted to competitive environments, while women tend to be more drawn to collaboration (e.g., [80]). We speculate that carefully combining competitive and collaborative strategies to appeal to both the intrinsic and extrinsic motivations and values of these groups can engender more diversity and embrace these differences, and setting clear goals and intermediate stages for the investigation during the planning phase can potentially remedy tensions that arise.

Our findings demonstrate that social OSINT investigations have the potential to be more ethical than individual ones, as individual analyst can be prone to oversight and employing current technological tools does not remedy those issues. DeGrassi et al. [81] found that diverse groups are more likely to make ethical decisions than homogenous groups and individuals. In addition, our study show that,in the case of O9, all investigative work is logged and tracked on proprietary machines and accessible to all investigators assigned to the case; this accountability encourages restraint from committing crimes. O3 organizers also emphasize the importance of training members in the ethical implications of OSINT investigations. Based on these factors, we suggest that increasing the size and diversity of groups conducting investigations can potentially lead to restraint from criminal or vigilante behavior [82] and to more ethical investigations.

Finally, a lesson learned from the organization we studied is that encouraging the establishment of strong social ties, more specifically friendships, between investigators can provide advantages beyond sensemaking and productivity by strengthening the available support systems and building resilience into investigative teams. Kessler et al. [83] posit that careful conversation with a sympathetic peer can help cope with stress, which is also a PTSD mitigation technique employed by other OSINT investigators who focus on human rights violations [16]. Growing a strong social support network between investigators, and even including members who are versed in therapeutic techniques, can assist in shielding members from secondary trauma. We suggest that more organizations conducting OSINT or other investigations can benefit from this model.

Chapter 6

Conclusion

The rise of social media and increasingly digitally-mediated social interaction has democratized access to large amounts of personal information. This information has been leveraged by both professionals and novice sleuths online. While novice sleuths conducting crowdsourced investigations online can often lead to vigilantism and other mishaps, we observe the rise of a community of professional OSINT investigators successfully conducting bottom-up investigations online. These professional OSINT investigators leverage competition and collaboration to conduct their investigations more efficiently and more ethically; we refer to both approaches as *social OSINT*.

The first contribution of this thesis is an in-depth description of the social structures that support OSINT investigations, detailing the various personal, interpersonal, and organizational factors that shaped them, which extends the current literature on investigations. We also define and characterize the term ”social OSINT” and detail the competitive and collaborative strategies employed by OSINT investigators and their implications; this concept adds more nuance to previous work on competitions.

The second contribution of this work is a list of recommendations and implications for structuring other open source intelligence work and citizen investigations, both online and offline. These recommendations extend the current literature on investigations as well, and can be generalized to both, bottom-up investigations model employing crowds of novices online and the more recent hybrid investigations model in which crowds are led by experts.

We expect the latter two models to continue growing and gaining importance in the coming years.

Finally, this thesis contributes three design characteristics and three design recommendations for tools and technological systems to better support the OSINT community but also any professionals leveraging OSINT data in a multitude of domain applications. By leveraging previous research on crowdsourcing and sensemaking and adapting previous findings in the development of tools, our recommendations can assist the OSINT community to overcome some of their current challenges.

With increasing amounts of online data being generated, and access to such data being democratized, people are empowered to use that content for various motivations. Professional OSINT investigators are a successful example of using such data towards diverse goals, ranging from uncovering human rights violations around the globe, to fighting child trafficking and exploitation, to debunking disinformation. This work examines the social structures enabling the OSINT community to accomplish their feats and the potential for various concepts such as crowdsourcing, sensemaking, and design appropriation, among others, to further empower and support OSINT investigators and other citizen investigations.

Bibliography

[1] "Stop child abuse – trace an object," Feb 2019.

[2] C. Dewey, "Crowdsourcing may have solved a 20-year-old cold case," *The Washington Post*, 2015.

[3] N. J. LaLone, J. Kropczynski, and A. H. Tapia, "The symbiotic relationship of crisis response professionals and enthusiasts as demonstrated by reddit's user-interface over time," in *ISCRAM*, 2018.

[4] D. Dailey and K. Starbird, "Journalists as crowdsourcerers: Responding to crisis by reporting with a crowd," *Computer Supported Cooperative Work (CSCW)*, vol. 23, pp. 445–481, Dec 2014.

[5] J. Nhan, L. Huey, and R. Broll, "Digilantism: An analysis of crowdsourcing and the boston marathon bombings," *The British journal of criminology*, vol. 57, no. 2, pp. 341–361, 2017.

[6] D. Trottier, "Digital vigilantism as weaponisation of visibility," *Philosophy & Technology*, vol. 30, no. 1, pp. 55–72, 2017.

[7] M. Kornfield, "The wrong ID: Retired firefighter, comedian and Chuck Norris falsely accused of being Capitol rioters," *Washington Post*, 2021.

[8] E. Yardley, A. G. T. Lynes, D. Wilson, and E. Kelly, "What's the deal with 'websleuthing'? news media representations of amateur detectives in networked spaces," *Crime, Media, Culture*, vol. 14, no. 1, pp. 81–109, 2018.

[9] Amnesty International, “Syria: ‘Nowhere is Safe for Us’: Unlawful Attacks and Mass Displacement in North-West Syria,” tech. rep., May 2020.

[10] R. Farrow, “An air force combat veteran breached the senate,” 2021.

[11] G. Myre, “How online sleuths identified rioters at the capitol,” 2021.

[12] Michael W. McLaughlin, “Using open source intelligence software for cybersecurity intelligence,” June 2012.

[13] H. J. Williams and I. Blum, “Defining second generation open source intelligence (osint) for the defense enterprise,” tech. rep., RAND Corporation Santa Monica United States, 2018.

[14] Ryan Hunt, “Thirty-Seven Percent of Companies Use Social Networks to Research Potential Job Candidates, According to New CareerBuilder Survey - Apr 18, 2012,” Apr. 2012.

[15] A. Broughton, B. Foley, S. Ledermaier, and A. Cox, “The use of social media in the recruitment process,” p. 81, 2014.

[16] *Digital Witness: Using Open Source Information for Human Rights Investigation, Documentation, and Accountability.* Oxford, New York: Oxford University Press, Feb. 2020.

[17] L. Kermode, J. Freyberg, A. Akturk, R. Trafford, D. Kochetkov, R. Pardinas, E. Weizman, and J. Cornebise, “Objects of violence: synthetic data for practical ML in human rights investigations,” *arXiv:2004.01030 [cs]*, Apr. 2020. arXiv: 2004.01030.

[18] S. C. Mercado, “Sailing the sea of OSINT in the information age,” tech. rep., American Psychological Association, 2004. type: dataset.

[19] C. P. Lee, P. Dourish, and G. Mark, “The human infrastructure of cyberinfrastructure,” in *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work*, CSCW ’06, (New York, NY, USA), p. 483–492, Association for Computing Machinery, 2006.

[20] M. Dye, D. Nemer, J. Mangiameli, A. S. Bruckman, and N. Kumar, “El paquete semanal: The week’s internet in havana,” in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2018.

[21] S. Venkatagiri, A. Gautam, and K. Luther, “Crowdsolve: Managing tensions in an expert-led crowdsourced investigation,” *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW1, pp. 1–30, 2021.

[22] M. Glassman and M. J. Kang, “Intelligence in the internet age: The emergence and evolution of Open Source Intelligence (OSINT),” *Computers in Human Behavior*, vol. 28, pp. 673–682, Mar. 2012.

[23] E. Higgins, *We Are Bellingcat : An Intelligence Agency for the People*. London: Bloomsbury Publishing, 2021.

[24] J. L. Calof and S. Wright, “Competitive intelligence: A practitioner, academic and inter-disciplinary perspective,” *European Journal of Marketing*, vol. 42, pp. 717–730, Jan. 2008. Publisher: Emerald Group Publishing Limited.

[25] D. Rouach and P. Santi, “Competitive Intelligence Adds Value:: Five Intelligence Attitudes,” *European Management Journal*, vol. 19, pp. 552–559, Oct. 2001.

[26] L. Šubelj, . Furlan, and M. Bajec, “An expert system for detecting automobile insurance fraud using social network analysis,” *Expert Systems with Applications*, vol. 38, pp. 1039–1052, Jan. 2011.

[27] D. of Homeland Security, “(U//FOUO//LES) DHS Terrorist Use of Social Networking Facebook Case Study | Public Intelligence,” Dec. 2010.

[28] Esteban Borges, “SecurityTrails | OSINT Framework: The Perfect Cybersecurity Intel Gathering Tool,” Jan. 2019.

[29] S. McKeown, D. Maxwell, L. Azzopardi, and W. B. Glisson, “Investigating people: a qualitative analysis of the search behaviours of open-source intelligence analysts,” in *Proceedings of the 5th Information Interaction in Context Symposium*, IIiX ’14, (New York, NY, USA), pp. 175–184, Association for Computing Machinery, Aug. 2014.

[30] M. Hanham and J. Shin, “Ethics in the age of osint innocence,” *Stanley Center for Peace and Security*, May 2020.

[31] “Search Party Rules.”

[32] S. Dyer and G. Ivens, “What would a feminist open source investigation look like?,” *Digi War*, Apr. 2020.

[33] A. S. Hulnick, “The Downside of Open Source Intelligence,” *International Journal of Intelligence and CounterIntelligence*, vol. 15, pp. 565–579, Nov. 2002.

[34] E. Baker, E. Stover, R. Haar, A. Lampros, and A. Koenig, “Safer Viewing,” *Health Hum Rights*, vol. 22, pp. 293–304, June 2020.

[35] P. Gill, “The way ahead in explaining intelligence organization and process,” 2018.

[36] R. Poelman, O. Akman, S. Lukosch, and P. Jonker, “As if being there: mediated reality for crime scene investigation,” in *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, CSCW ’12, (New York, NY, USA), pp. 1267–1276, Association for Computing Machinery, Feb. 2012.

[37] J. Alcaidinho, L. Freil, T. Kelly, K. Marland, C. Wu, B. Wittenbrook, G. Valentin, and M. Jackson, "Mobile Collaboration for Human and Canine Police Explosive Detection Teams," in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, (New York, NY, USA), pp. 925–933, Association for Computing Machinery, Feb. 2017.

[38] Y. L. Huang, K. Starbird, M. Orand, S. A. Stanek, and H. T. Pedersen, "Connected Through Crisis: Emotional Proximity and the Spread of Misinformation Online," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '15, (New York, NY, USA), pp. 969–980, Association for Computing Machinery, Feb. 2015.

[39] A. Israni, S. Erete, and C. L. Smith, "Snitches, Trolls, and Social Norms: Unpacking Perceptions of Social Media Use for Crime Prevention," in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, (New York, NY, USA), pp. 1193–1209, Association for Computing Machinery, Feb. 2017.

[40] S. L. Erete, "Engaging Around Neighborhood Issues: How Online Communication Affects Offline Behavior," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '15, (New York, NY, USA), pp. 1590–1601, Association for Computing Machinery, Feb. 2015.

[41] S. Lewis and D. A. Lewis, "Examining technology that supports community policing," p. 10, 2012.

[42] A. B. Brush, J. Jung, R. Mahajan, and F. Martinez, "Digital neighborhood watch: investigating the sharing of camera data amongst neighbors," in *Proceedings of the*

2013 conference on Computer supported cooperative work, CSCW '13, (New York, NY, USA), pp. 693–700, Association for Computing Machinery, Feb. 2013.

[43] N. Sachdeva and P. Kumaraguru, "Call for Service: Characterizing and Modeling Police Response to Serviceable Requests on Facebook," in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, (New York, NY, USA), pp. 336–352, Association for Computing Machinery, Feb. 2017.

[44] K. Starbird and L. Palen, "(How) will the revolution be retweeted? information diffusion and the 2011 Egyptian uprising," in *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, CSCW '12, (New York, NY, USA), pp. 7–16, Association for Computing Machinery, Feb. 2012.

[45] V. Wulf, K. Misaki, M. Atam, D. Randall, and M. Rohde, "'On the ground' in Sidi Bouzid: investigating social media use during the tunisian revolution," in *Proceedings of the 2013 conference on Computer supported cooperative work*, CSCW '13, (New York, NY, USA), pp. 1409–1418, Association for Computing Machinery, Feb. 2013.

[46] A. Arif, J. J. Robinson, S. A. Stanek, E. S. Fichet, P. Townsend, Z. Worku, and K. Starbird, "A Closer Look at the Self-Correcting Crowd: Examining Corrections in Online Rumors," in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, (New York, NY, USA), pp. 155–168, Association for Computing Machinery, Feb. 2017.

[47] A. Wiggins and Y. He, "Community-based data validation practices in citizen science," in *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, (New York, NY, USA), p. 1548–1559, Association for Computing Machinery, 2016.

[48] R. Tinati, M. Van Kleek, E. Simperl, M. Luczak-Rösch, R. Simpson, and N. Shadbolt, "Designing for citizen data analysis: A cross-sectional case study of a multi-domain citizen science platform," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, (New York, NY, USA), p. 4069–4078, Association for Computing Machinery, 2015.

[49] S. Venkatagiri, J. Thebault-Spieker, R. Kohler, J. Purviance, R. S. Mansur, and K. Luther, "GroundTruth: Augmenting Expert Image Geolocation with Crowdsourcing and Shared Representations," *Proc. ACM Hum.-Comput. Interact.*, vol. 3, pp. 107:1–107:30, Nov. 2019.

[50] T. Li, K. Luther, and C. North, "CrowdIA: Solving Mysteries with Crowdsourced Sensemaking," *Proc. ACM Hum.-Comput. Interact.*, vol. 2, pp. 105:1–105:29, Nov. 2018.

[51] T. Li, Y. Belghith, C. North, and K. Luther, "CrowdTrace: Visualizing Provenance in Distributed Sensemaking," p. 5.

[52] K. Hutter, J. Hautz, J. Füller, J. Mueller, and K. Matzler, "Communitition: The Tension between Competition and Collaboration in Community-Based Design Contests," *Creativity and Innovation Management*, vol. 20, no. 1, pp. 3–21, 2011. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8691.2011.00589.x.

[53] E. Porter, C. Bopp, E. Gerber, and A. Voida, "Reappropriating Hackathons: The Production Work of the CHI4Good Day of Service," p. 5, 2017.

[54] Y. Tausczik and P. Wang, "To Share, or Not to Share? Community-Level Collaboration in Open Innovation Contests," *Proc. ACM Hum.-Comput. Interact.*, vol. 1, pp. 100:1–100:23, Dec. 2017.

[55] B. Morschheuser, A. Maedche, and D. Walter, "Designing Cooperative Gamification:

Conceptualization and Prototypical Implementation," in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, (New York, NY, USA), pp. 2410–2421, Association for Computing Machinery, Feb. 2017.

[56] M. J. Rogerson, M. R. Gibbs, and W. Smith, "Cooperating to Compete: the Mutuality of Cooperation and Competition in Boardgame Play," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, (New York, NY, USA), pp. 1–13, Association for Computing Machinery, Apr. 2018.

[57] S. Lee, S. Lee, Y. Lee, S. Park, and J. Kim, "Effect of competition and collaboration in social network game on intimacy among players," in *Proceedings of HCI Korea*, HCIK '15, (Seoul, KOR), pp. 425–433, Hanbit Media, Inc., Dec. 2014.

[58] A. Michael and C. Lutteroth, "Race Yourselves: A Longitudinal Exploration of Self-Competition Between Past, Present, and Future Performances in a VR Exergame," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, (New York, NY, USA), pp. 1–17, Association for Computing Machinery, Apr. 2020.

[59] J. Park, A. Oh, and S. Kim, "Analysis of the Effect of Competition on Player Immersion and Engagement in a Mobile Game," p. 6, 2017.

[60] N. Yee, N. Ducheneaut, and L. Nelson, "Online gaming motivations scale: development and validation," p. 4, 2012.

[61] L. Yu, P. André, A. Kittur, and R. Kraut, "A comparison of social, learning, and financial strategies on crowd engagement and output quality," in *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, CSCW '14, (New York, NY, USA), pp. 967–978, Association for Computing Machinery, Feb. 2014.

[62] M. D. C. Tongco, "Purposive Sampling as a Tool for Informant Selection," *Ethnobotany Research and Applications*, vol. 5, pp. 147–158, Dec. 2007. Number: 0.

[63] I. Seidman, *Interviewing as Qualitative Research: A Guide for Researchers in Education and the Social Sciences*. Teachers College Press, 2006.

[64] M. J. Salganik and D. J. Watts, "Web-Based Experiments for the Study of Collective Social Dynamics in Cultural Markets," *Topics in Cognitive Science*, vol. 1, pp. 439–468, July 2009.

[65] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Research in Psychology*, vol. 3, pp. 77–101, Jan. 2006. Publisher: Routledge _eprint: https://www.tandfonline.com/doi/pdf/10.1191/1478088706qp063oa.

[66] "Verification Handbook: homepage."

[67] Y. Tausczik and M. Boons, "Distributed Knowledge in Crowds: Crowd Performance on Hidden Profile Tasks.," p. 10.

[68] A. A. Kane, S. Kiesler, and R. Kang, "Inaccuracy Blindness in Collaboration Persists,," p. 9, 2018.

[69] J. A. Gonzales, C. Fiesler, and A. Bruckman, "Towards an Appropriable CSCW Tool Ecology: Lessons from the Greatest International Scavenger Hunt the World Has Ever Seen," p. 12, 2015.

[70] G. Convertino, H. M. Mentis, M. B. Rosson, A. Slavkovic, and J. M. Carroll, "Supporting content and process common ground in computer-supported teamwork," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, (New York, NY, USA), pp. 2339–2348, Association for Computing Machinery, Apr. 2009.

[71] N. J. Pioch and J. O. Everett, “POLESTAR: collaborative knowledge management and sensemaking tools for intelligence analysts,” in *Proceedings of the 15th ACM international conference on Information and knowledge management*, CIKM ’06, (New York, NY, USA), pp. 513–521, Association for Computing Machinery, Nov. 2006.

[72] “CrossCheck: Our Collaborative Online Verification Newsroom.”

[73] “Check.”

[74] P. Ganguly and P. Ray, “Software interoperability of telemedicine systems: a CSCW perspective,” in *Proceedings Seventh International Conference on Parallel and Distributed Systems (Cat. No.PR00568)*, pp. 349–356, July 2000. ISSN: 1521-9097.

[75] D. C. Engelbart, “Knowledge-domain interoperability and an open hyperdocument system,” in *Proceedings of the 1990 ACM conference on Computer-supported cooperative work*, CSCW ’90, (New York, NY, USA), pp. 143–156, Association for Computing Machinery, Sept. 1990.

[76] H.-C. Wang, S. R. Fussell, and D. Cosley, “From diversity to creativity: stimulating group brainstorming with cultural differences and conversationally-retrieved pictures,” p. 10.

[77] P. Shachaf, “Cultural diversity and information and communication technology impacts on global virtual teams: An exploratory study,” *Information & Management*, vol. 45, pp. 131–142, Mar. 2008.

[78] A. W. Woolley, C. F. Chabris, A. Pentland, N. Hashmi, and T. W. Malone, “Evidence for a Collective Intelligence Factor in the Performance of Human Groups,” *Science*, vol. 330, pp. 686–688, Oct. 2010. Publisher: American Association for the Advancement of Science Section: Report.

[79] C. D. Cramton and P. J. Hinds, “An Embedded Model of Cultural Adaptation in Global Teams,” *Organization Science*, vol. 25, pp. 1056–1081, Jan. 2014. Publisher: INFORMS.

[80] M. Niederle and L. Vesterlund, “Do Women Shy Away From Competition? Do Men Compete Too Much?*,” *The Quarterly Journal of Economics*, vol. 122, pp. 1067–1101, Aug. 2007.

[81] S. W. DeGrassi, W. B. Morgan, and S. S. Walker, “Ethical Decision-Making: Group Diversity Holds the Key,” p. 15.

[82] “Is our growing obsession with true crime a problem?,” *BBC News*, Mar. 2019.

[83] R. C. Kessler, R. H. Price, and C. B. Wortman, “Social Factors in Psychopathology: Stress, Social Support, and Coping Processes,” *Annual Review of Psychology*, vol. 36, no. 1, pp. 531–572, 1985. _eprint: https://doi.org/10.1146/annurev.ps.36.020185.002531.