

Tailoring Generative AI to Augment Creative Leadership in Capture-The-Flag Development

Anirban Mukhopadhyay

Department of Computer Science & Center for HCI,
Virginia Tech
Blacksburg, VA, USA
anirban@vt.edu

Kurt Luther

Department of Computer Science & Center for HCI,
Virginia Tech
Alexandria, VA, USA
kluther@vt.edu

Abstract

Capture-the-Flag (CTF) competitions are valuable for training in cybersecurity and investigative skills, but their development is time-consuming and requires skilled staff. This paper explores how generative AI can augment creativity and collaboration to streamline CTF development. Using a Research-through-Design (RtD) approach, we develop CTFBot, an AI-powered assistant to support leadership behaviors such as planning, clarifying, monitoring progress, and problem-solving. Grounded in Distributed Leadership (DL) theory, CTFBot enhances collaboration while preserving human creativity. A pilot study reveals challenges in maintaining engagement and providing support through conversational user interfaces. This work offers insights into AI-assisted collaboration for team-based creative tasks.

Keywords

OSINT, Creativity, Collaboration, Distributed Leadership, Research through Design

ACM Reference Format:

Anirban Mukhopadhyay and Kurt Luther. 2018. Tailoring Generative AI to Augment Creative Leadership in Capture-The-Flag Development. *Proc. ACM Meas. Anal. Comput. Syst.* 37, 4, Article 111 (August 2018), 5 pages. <https://doi.org/XXXXXXX.XXXXXXX>

1 Introduction

Capture-the-Flag (CTF) competitions are widely recognized as an effective method for building investigative skills, particularly in cybersecurity [7, 8, 11]. CTFs present participants with a series of challenges across various topics, often simulating real-world scenarios faced by professionals in the field. These challenges involve uncovering hidden, unguessable strings known as "flags," which are revealed only when the correct solutions are found. Participants can submit each flag to a scoring server, earning points based on the difficulty of the challenge.

Authors' Contact Information: Anirban Mukhopadhyay, Department of Computer Science & Center for HCI, Virginia Tech, Blacksburg, VA, USA, anirban@vt.edu; Kurt Luther, Department of Computer Science & Center for HCI, Virginia Tech, Alexandria, VA, USA, kluther@vt.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM 2476-1249/2018/8-ART111
<https://doi.org/XXXXXXX.XXXXXXX>

CTFs are especially relevant for Open Source Intelligence (OSINT) due to the need for an adversarial mindset during investigations and the sense of urgency, as highlighted by Belghith et al. [3]. OSINT involves collecting and analyzing publicly available data to address specific investigative needs. For example, investigative journalists use it to verify the authenticity of videos from conflict zones such as Ukraine, while cybersecurity experts rely on OSINT techniques to uncover data breaches and prevent social engineering attacks. Despite its growing importance, existing OSINT training resources often lack interactivity and fail to replicate the complexity of open-ended and dynamic investigations, limiting their effectiveness [2, 3, 14].

Developing OSINT-based CTFs is a creative process that requires human cognition, domain expertise, and problem-solving skills. It is challenging due to the significant effort and specialized skills required to create realistic and engaging challenges [7, 14]. To address this, teams are often formed to divide the workload and leverage diverse expertise. However, this introduces new challenges in coordination, aligning on difficulty levels, and maintaining educational value. Distributed leadership through project management tools has been attempted as a solution [12, 16], allowing multiple team members to share responsibility, but it provides limited support with planning and implementation of associated tasks. Here, generative AI could help by providing scaffolds for challenge creation, streamlining collaboration, and structuring leadership behaviors to enhance workflow efficiency.

However, integrating generative AI into CTF development requires careful consideration to ensure it enhances creativity and collaboration without compromising critical thinking and leadership dynamics. This paper explores how generative AI can support critical thinking and decision-making throughout the CTF development process by promoting leadership behaviors to improve team collaboration and the quality of challenges.

2 Method

We apply a Research-through-Design (RtD) method, which is a practice-based research methodology that generates knowledge through iterative design, prototyping, and evaluation of artifacts [5, 17]. RtD allows researchers to explore evolving problems like collaboration by actively engaging with the design process.

We build on our previous work where we employed an RtD approach in developing collaborative CTFs (CoCTFs) for debunking social media misinformation [14]. Through multiple deployments over a semester with 46 students, CoCTFs evolved to provide participants with greater agency through competition while addressing challenges like information silos and groupthink using collaborative

features such as flag sharing and crowdsourcing. Our work is also informed by Benjamin et al. [4], who used RtD in an educational context and demonstrated its value as a "rapid response methodology" for engaging with fast-evolving, disruptive technologies like generative AI.

In this study, we employ RtD to develop CTFBot, an AI-powered agent designed to support leadership behaviors in creating high-quality OSINT-based CTFs collaboratively. We describe the four stages of the RtD process used to design the CTFBot in the following sections: (1) Framing the problem; (2) Conceptual development and design goals; (3) Ideation and development of higher-fidelity prototypes; and (4) Deployment, iteration, and evaluation.

2.1 Framing the Problem

We ran a formative study with five undergraduate students who developed and hosted an OSINT-focused scenario-based CTF to identify challenges in CTF development. We ran a focus group study at the end of two months to collect feedback. The study revealed major challenges, including scenario consistency across flags, balancing creativity with learning goals in flag design, task tracking, and getting structured feedback. The team struggled with maintaining a cohesive narrative, avoiding overlap in flag creation, and efficiently distributing responsibilities. These challenges highlighted the need for better leadership structures.

2.2 Conceptual Development and Design Goals

We leveraged Distributed Leadership (DL) as a theoretical framework to address these challenges [6]. By distributing leadership tasks across team members, DL reduces reliance on a single leader while enhancing engagement, communication, and problem-solving. Prior work, such as Luther et al.'s Pipeline system, demonstrates how technology can facilitate task-oriented leadership behaviors -- planning, clarifying, monitoring, and problem-solving -- essential for team effectiveness [12]. We posit that by supporting these behaviors through technology, we can better support collaboration and ensure higher-quality CTF challenges.

2.3 Ideation and Development of Higher-Fidelity Prototypes

We explored AI-driven solutions to support leadership in collaborative creative work. We identified generative AI, particularly ChatGPT, as a useful tool for content generation, structuring ideas, and getting feedback during the formative study. He et al. [9] found value in using generative AI to support group facilitation and enhance the diversity of perspectives and ideas during creative brainstorming. Building on these ideas, we created an AI agent, CTFBot, to support leadership behaviors through structured AI actions. The bot was designed to act as a coordinative artifact [10] that helps teams navigate the CTF creation process while mitigating common collaboration bottlenecks.

The bot's functionality was built around four key AI actions: Planning, Clarifying, Monitoring Progress, and Problem-Solving. The *Planning* action generates templates and concept maps to help teams structure their work. The *Clarifying* action suggests OSINT tools and flag ideas, referencing knowledge sources like real-world

investigation reports and guides, and balancing creative brainstorming with realistic scenario-building. To track progress, the bot's *Monitoring* action summarizes updates from uploaded working documents and suggests the next steps to keep the team aligned. Lastly, the *Problem-Solving* action provides early-stage feedback on flag difficulty, assesses hint quality and generates multiple solution pathways to refine challenges. These AI-driven features were designed to reinforce task-oriented leadership behaviors described by Yukl et al. [15] and help teams achieve both outcome goals (high-quality flags) and process goals (effective collaboration).

Building on Slack bots used in team-based scenarios [13], we tested the idea of agentic workflows through a Discord bot that can be invoked in team channels. CTFBot was implemented using Node.js and connected to an Open AI Assistant via API. The Open AI Assistant is powered by the state-of-the-art GPT-4o model [1]. The assistant has access to tools like code interpreter that can run code and display results, and file search which can help retrieve relevant parts of the knowledge base to augment responses to user queries. The bot describes the steps taken, including any tool used, in its response. Figure 1 shows a screenshot of a prototype Discord bot named "OSINT Assistant" being used by participants for early testing.

2.4 Deployment, Iteration, and Evaluation

We ran a two-week pilot study with students developing their own CTFs in a semester-long OSINT university course. There were 20 graduate students divided into five teams of four members. CTFBot was integrated into Discord team channels, providing structured AI assistance throughout the development phases. We provided a set of prompts to try out during class apart from freely using the bot. We collected feedback through surveys at the end of classes to iterate on its usability and usefulness. We iterated on the prompt powering the Open AI assistant to perform the AI actions (provided in Section A). We did not conduct a summative evaluation because the bot did not fully meet the design goals and presented usability challenges.

3 Preliminary Results

3.1 Strengths and Limitations of CTFBot

Based on limited feedback collected from multiple teams working on CTF development, the chatbot demonstrated strong capabilities in analyzing and ranking flags, showing adaptability to increasingly specific questions about flag design. The bot also performed well with shorter, focused queries to improve the quality of flags and suggest creative names for fictitious elements in the design. However, teams faced challenges with the bot's image generation capabilities and handling of complex queries, often requiring multiple attempts to get useful responses. Technical issues included repeated and overly long responses and difficulties in processing uploaded files. Some users noted that the bot sometimes provided vague responses and struggled with detailed implementation steps.

3.2 Challenges in Collaborative Use of CTFBot

While integrating an AI-powered bot into a team communication platform like Discord seemed promising, our study revealed significant limitations in effectively combining team collaboration,

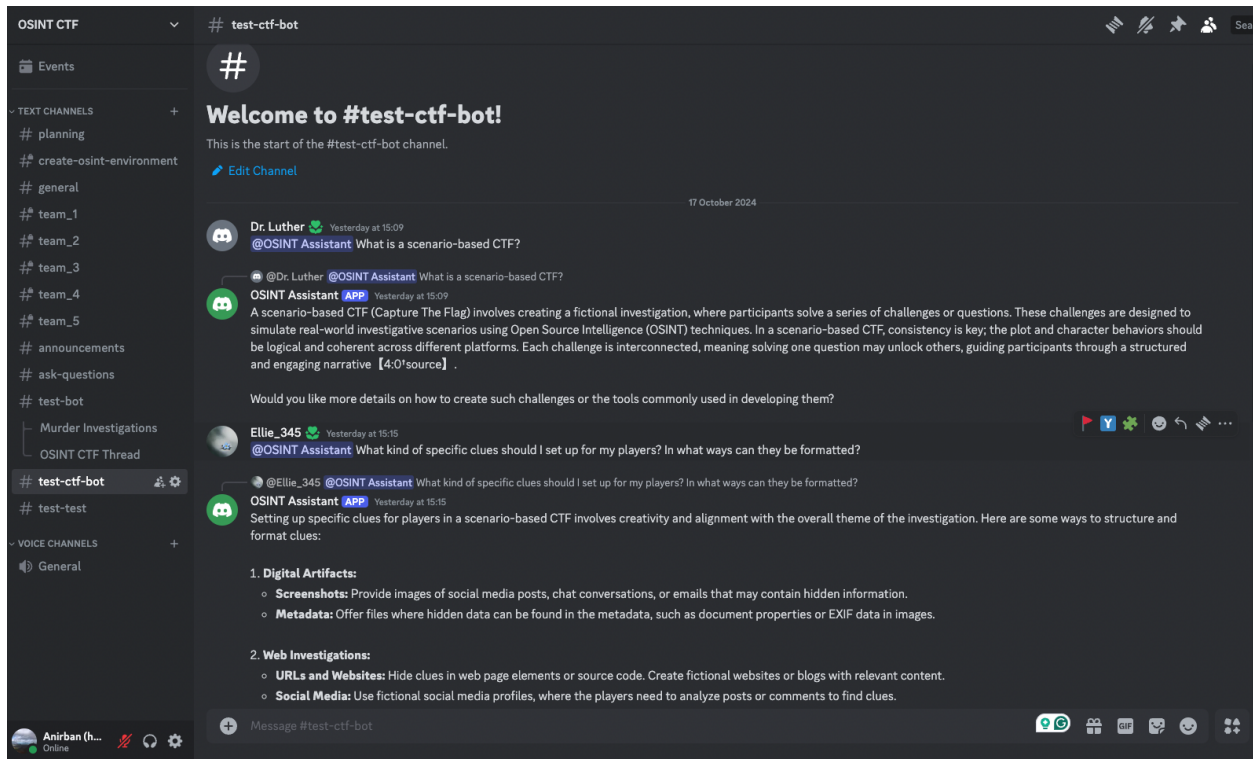


Figure 1: Screenshot of the prototype Discord bot named OSINT Assistant

structured AI support, and a shared workspace for Generative AI assistance. The bot’s usage remained constrained to specific tasks rather than seamlessly integrating into the overall workflow. Teams struggled to incorporate AI-generated results into their discussions, making it unclear how the bot’s outputs influenced decision-making or final design.

Teams were co-located and had access to multiple communication channels, reducing their reliance on a shared chat-based interface for coordination. As a result, participants found that working through a chatbot often felt disconnected from their primary workflow. Follow-ups from team members were not always collaborative — users frequently created separate threads to maintain context, but this fragmented discussions and AI-generated results. This lack of continuity made it difficult for members to track conversations, integrate AI outputs effectively, and stay aligned throughout the CTF development process.

Moving forward, we plan to redesign the bot based on this feedback and our observations. Improvements will focus on enhancing the user experience through structured scaffolds and menus integrated into the chat interface to guide interactions. We will work to clarify the bot’s role and capabilities more explicitly to users. We are also exploring the balance between proactive interventions (where the bot anticipates and suggests next steps) and reactive support (where it responds to specific queries) to create a more intuitive and supportive development environment.

4 Discussion

4.1 Designing AI to Augment Human Cognition Through Leadership Behaviors

We found that generative AI can enhance CTF development by generating problem templates, crafting engaging scenarios, suggesting refinements, and evaluating challenge difficulty. However, integrating AI into creative and investigative tasks raises concerns about diminishing human agency and critical thinking. In our case, a key objective of the CTF development process was to help students learn OSINT tradecraft and refine their investigative reasoning skills. Here, automated workflows risk students consuming AI-generated content rather than actively developing their problem-solving abilities. To foster meaningful learning, human control and thoughtful decision-making must remain central to the process.

AI should function as a co-creator rather than an autonomous decision-maker. By providing structured prompts, alternative perspectives, and analytical insights — rather than direct answers — AI can help maintain cognitive engagement. This aligns with the concept of redistributing leadership that we used in the study [12], where AI reinforces leadership behaviors such as planning, clarifying, monitoring progress, and problem-solving without replacing human expertise. By designing AI interventions that reinforce these behaviors, we can structure human-AI collaboration in a way that enhances both teamwork and cognition. The challenge remains in designing user experiences that encourage iterative refinement rather than passive acceptance of AI-generated outputs. The open

question here is: How can AI be used to scaffold leadership behaviors in a way that enhances team dynamics rather than creating dependency?

4.2 Embedding Generative AI into Collaborative Workflows

Our findings highlight the need for better integration of AI assistance within team workflows for creative tasks, moving beyond reliance on chat-based interactions. For example, in domains like business meetings or product design, AI could be embedded directly into collaborative tools — summarizing discussions, suggesting design alternatives, or tracking decisions in real-time — to support ongoing group work rather than functioning as a separate conversational agent. It will be important to investigate what alternative interaction paradigms can augment chat-based AI assistance.

We think that a lack of proactive AI support limited engagement and made the bot less useful in team workflows. While retrospective feedback helped users refine ideas and maintain creative control, relying solely on it led to missed opportunities for timely intervention. Real-time assistance, when well-integrated, could help teams structure tasks and stay aligned. To enhance usefulness, AI should proactively surface guidance during challenging phases of collaborative workflows, ensuring it supports rather than disrupts human cognition.

5 Acknowledgement

This work was partly supported by the Virginia Commonwealth Cyber Initiative.

References

- [1] 2024. OpenAI Platform. Retrieved 2024-09-11 from <https://platform.openai.com>
- [2] Michael Bazzell. 2016. *Open Source Intelligence Techniques: Resources for Searching and Analyzing Online Information* (5th ed.). CreateSpace Independent Publishing Platform, North Charleston, SC, USA.
- [3] Yasmine Belghith, Sukrit Venkatagiri, and Kurt Luther. 2022. Compete, Collaborate, Investigate: Exploring the Social Structures of Open Source Intelligence Investigations. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3491102.3517526
- [4] Jesse Josua Benjamin, Joseph Lindley, Elizabeth Edwards, Elisa Rubegni, Tim Korjakow, David Grist, and Rhiannon Sharkey. 2024. Responding to Generative AI Technologies with Research-through-Design: The Ryelands AI Lab as an Exploratory Study. In *Designing Interactive Systems Conference*. 1823–1841. doi:10.1145/3643834.3660677 arXiv:2405.04677 [cs].
- [5] Arne Berger, Sören Totzauer, Kevin Lefevre, Michael Storz, Albrecht Kurze, and Andreas Bischof. 2017. Wicked, Open, Collaborative: Why Research through Design Matters for HCI Research. *i-com* 16, 2 (2017), 131–142.
- [6] Richard Bolden. 2011. Distributed Leadership in Organizations: A Review of Theory and Research. *International Journal of Management Reviews* 13, 3 (2011), 251–269. doi:10.1111/j.1468-2370.2011.00306.x _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-2370.2011.00306.x>.
- [7] Gregory Conti, Thomas Babbitt, and John Nelson. 2011. Hacking Competitions and Their Untapped Potential for Security Education. *IEEE Security & Privacy* 9, 3 (May 2011), 56–59. doi:10.1109/MSP.2011.51 Conference Name: IEEE Security & Privacy.
- [8] Ahmad Haziq Ashrofi Hanafi, Haikal Rokman, Ahmad Dahaqin Ibrahim, Zul-Azri Ibrahim, Md Nabil Ahmad Zawawi, and Fiza Abdul Rahim. 2021. A CTF-Based Approach in Cyber Security Education for Secondary School Students. *Electronic Journal of Computer Science and Information Technology* 7, 1 (Oct. 2021). doi:10.52650/ejcsit.v7i1.107 Number: 1.
- [9] Jessica He, Stephanie Houde, Gabriel E Gonzalez, Dario Andrés Silva Moran, Steven I Ross, Michael Muller, and Justin D Weisz. 2024. AI and the Future of Collaborative Work: Group Ideation with an LLM in a Virtual Canvas. In *Proceedings of the 3rd Annual Meeting of the Symposium on Human-Computer Interaction for Work*. 1–14.
- [10] Ming-Tung Hong, Jesse Josua Benjamin, and Claudia Müller-Birn. 2018. Coordinating Agents: Promoting Shared Situational Awareness in Collaborative Sense-making. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '18 Companion)*. Association for Computing Machinery, New York, NY, USA, 217–220. doi:10.1145/3272973.3274059
- [11] Stylianos Karagiannis, Elpidoforos Maragos, and Emmanouil Magkos. 2020. *An Analysis and Evaluation of Open Source Capture the Flag Platforms as Cybersecurity e-Learning Tools*. doi:10.1007/978-3-030-59291-2_5 Pages: 77.
- [12] Kurt Luther, Casey Fiesler, and Amy Bruckman. 2013. Redistributing leadership in online creative collaboration. In *Proceedings of the 2013 conference on Computer supported cooperative work*. ACM, San Antonio Texas USA, 1007–1022. doi:10.1145/2441776.2441891
- [13] Michael Muller. 2023. Exploring human-ai co-creativity under human control: Framing, reframing, brainstorming, and future challenges (workshop keynote). In *CEUR Workshop Proceedings*, Vol. 3547.
- [14] Sukrit Venkatagiri, Anirban Mukhopadhyay, David Hicks, Aaron Brantly, and Kurt Luther. 2023. CoSINT: Designing a Collaborative Capture the Flag Competition to Investigate Misinformation. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference (Pittsburgh, PA, USA) (DIS '23)*. Association for Computing Machinery, New York, NY, USA, 2551–2572. doi:10.1145/3563657.3595997
- [15] Gary Yukl. 2012. Effective leadership behavior: What we know and what questions need more attention. *Academy of Management perspectives* 26, 4 (2012), 66–85.
- [16] Jinlong Zhu, Zhenyu Liao, Kai Chi Yam, and Russell E. Johnson. 2018. Shared leadership: A state-of-the-art review and future research agenda. *Journal of Organizational Behavior* 39, 7 (2018), 834–852. doi:10.1002/job.2296 _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/job.2296>.
- [17] John Zimmerman and Jodi Forlizzi. 2014. Research through design in HCI. In *Ways of Knowing in HCI*. Springer, 167–189.

A AI Agent Prompt For Processing User Query

You are an AI assistant designed to support teams developing Open Source Intelligence (OSINT)-based Capture-The-Flag challenges. Your capabilities include the following 5 modules:

A. Planning

Rules: 1. Provide templates (not as a table) with headings wherever possible (e.g., templates for designing fictitious scenarios, flag features like relevant OSINT tools, real-world analog, and testing measures) to get the team started. 2. Show examples based on the templates. 3. Assist in breaking down the project into manageable tasks and subtasks.

B. Providing Feedback and Facilitating Brainstorming

Rules: 1. Ensure individual flag ideas are grounded in realistic OSINT practices and make use of OSINT tools and techniques. 2. Provide thought-provoking prompts or questions to inspire creativity and improvement.

Files: Utilize any initial flag ideas or documents provided by the team to offer tailored suggestions.

C. Monitoring Progress

Rules: 1. Provide a summary of task progress for the team, highlighting completed and pending tasks based on the status field. 2. Generate a report for each individual creator of flags (use the creator field to group the flags), summarizing their progress and highlighting the pending tasks. Use the status field. 3. Create a short plan for the team to complete the tasks by suggesting the next steps. 4. Highlight areas where the team can focus their efforts to complete the tasks by the end of the week.

Files: Access and index the team's flag ideas, task assignments, implementation plan, and status updates for a set of flags.

D. Maintaining Consistent Narrative for flags

Rules: 1. Summarize how the flag descriptions flow across the set of flags. Ask a follow-up question if the narrative is not clear, to gather details about the story behind the flags. 2. Provide feedback to maintain a consistent narrative for CTF based on the different flags

Files: Access the list of flags to understand the narrative of CTF.

E. Detecting Overlaps among Flags

Rules: 1. Analyze proposed flag details and implementation plans to detect potential overlaps by using the values from columns like Description, How to Solve, Skills Required, and Resources Required. Don't use difficulty level, hints, or owner names. 2. Alert team members to any overlaps and suggest collaboration opportunities.

Files: Access list of flags to identify overlaps.

F. Testing flags

Rules: 1. Provide detailed feedback on these aspects: 1a. Technical: Are the instructions relatively clear? 1b. Creativity: Is the flag embedded in the story-world and advancing the story? Is the description written in an interesting and engaging way? Is the puzzle relevant to OSINT? 2. Attempt to solve the flags and describe how you will solve it. Then mention if there a way to break the challenge or solve it easily through a work-around?

Files: Use the files with flag details provided by the team for testing and feedback.

““““

Wait. Think before you generate the answer. For each user request, generate results by following these steps:

1. Intent Recognition: Determine the user's intent based on keywords in their request.
2. Context Mapping: Map the intent to the appropriate module(s).
3. Rule Application: Apply the relevant rules associated with those modules.
4. File Utilization: If the user's message references an uploaded file, use its content to tailor your response. Reference the file content appropriately.
5. Response Generation: Print the Context Mapping step output at the top in a "[]" before the answer. Then provide a helpful and contextually appropriate response.
6. Always maintain a professional, supportive, and collaborative tone. Keep the responses short and actionable.

““““

This is an example of a User Request and how the Response should look like:

"Hey, we've uploaded a draft of our challenge ideas but aren't sure if there are any overlaps or conflicts. Can you check and let us know?"

Assistant's Processing:

Intent Recognition: The user wants assistance in detecting overlaps in their challenge ideas.

Context Mapping: Map to Module 3: Detecting Overlaps.

Rule Application: Analyze proposed ideas to detect potential overlaps. Alert team members to any overlaps. Suggest ways to collaborate or differentiate their work.

File Utilization: Use the uploaded draft of challenge ideas to perform the analysis.

Response Generation: Provide a detailed yet concise report on any overlaps found. Offer suggestions for addressing the overlaps.

Assistant's Response:

I've reviewed the draft of your challenge ideas that you uploaded. Here's what I found:

Overlap between Challenge 2 and Challenge 5:

Both challenges involve social media OSINT techniques focusing on Twitter data analysis. Suggestion: Consider combining these into a single, more comprehensive challenge or differentiating them by focusing on different platforms (e.g., Twitter vs. Instagram) or

different aspects of analysis. Overlap between Challenge 3 and Challenge 4:

Both are geolocation challenges using image metadata. Suggestion: Perhaps one challenge could focus on geolocation through metadata, while the other uses visual clues within images to determine location. Let me know if you'd like to discuss these suggestions further or need help brainstorming alternatives!"