Lessons from the Field: Practical Frameworks for CTF Competition Success

Dylan Middendorf*, Suman Bhunia*, Arthur Carvalho†
* Department of Computer Science and Software Engineering, Miami University, Oxford, Ohio
† Farmer School of Business, Miami University, Oxford, Ohio
Email: middendm@miamioh.edu, bhunias@miamioh.edu, carvalag@miamioh.edu

***Abstract*—Capture the Flag (CTF) competitions have become increasingly popular over the previous decade. This research aims to systematically approach organizing CTF competitions by developing two frameworks. The first framework utilizes design science research with a stakeholder approach to generate a definition of success. By defining success, an actionable set of design principles was derived to guide competition organizers. These design principles are structured to provide actionable insight to organizers to help solve the multi-objective problem of stakeholder requirements. The second framework identifies the challenge development life cycle through a temporal approach while describing the current best practices in challenge development. These models were then evaluated through a CTF competition case study hosted for an intra-collegiate audience. This competition utilized both frameworks to assess them in a critical light. The competition is hosted on cloud servers with appropriate security mechanisms to adhere to the industry standards. To our knowledge, this paper is the first to provide a design science approach for hosting a successful CTF. Based on the findings, a brief discussion on refining the frameworks and potential areas for improvement is provided.**

***Index Terms*—Capture the Flag, Computer Security, Design Science Research**

I. Introduction

Capture the Flag (CTF) competitions in computer security are structured exercises that challenge participants with technical problems, typically presented in a gamified environment. The goal is to solve these challenges by "capturing" a flag, representing the solution. The first ever CTF competition was hosted at DEF CON, a computer security conference, in 1996 [1]. This CTF competition provided participating teams with intentionally vulnerable infrastructure and two primary objectives: (1) secure the team's infrastructure and (2) identify and exploit vulnerabilities within other teams' servers.

Over the years, CTFs have become increasingly popular leading to several adaptations of the original game's style. For instance, the above format is known colloquially as an attack-defend CTF competition [2]. However, this format is not scalable for large competitions with hundreds or thousands of teams [3]. As a result, another format was developed, inspired by the popular television game show Jeopardy [4]. Specifically, this format organizes the challenges, or puzzles, within several categories, typically pulling from domains including *binary exploitation*, *cryptography*, *reverse engineering*, and *web exploitation*. Once a challenge has been solved, secret and unique data, known colloquially as flags, can be submitted for points, which vary depending on the challenge's difficulty, allowing for a global leaderboard to be created. In other words, the focus is shifted from vulnerable infrastructure to curated challenges. Several, less popular CTF formats have also been developed, involving defense-only and attack-only, or scavenger-hunt [5].

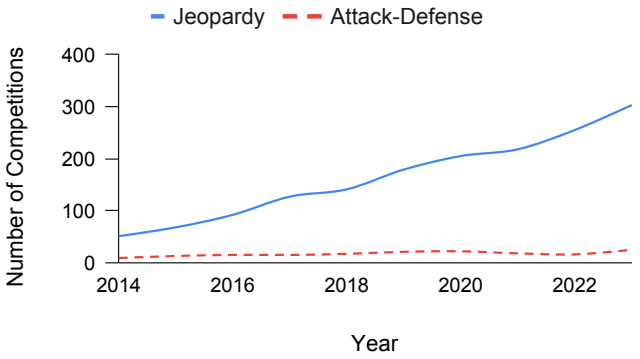


Figure 1: Estimated number of Jeopardy and attack-defense CTF competitions hosted per year between 2014 to 2023. The competitions were identified from CTFtime [6], a popular website containing information about CTF competitions, challenges, and teams. The number of Jeopardy-style CTF competitions appears to be increasing monotonically, while attack-defense CTF competitions remain consistent.

Over the past decade, the CTF ecosystem has grown considerably. This phenomenon can be observed through CTFtime's competition archives, with 51 Jeopardy competitions organized in 2014 and 303 competitions in 2023 (Fig. 1). As a result, most weekends now feature several events, hosted concurrently, reflecting a substantial increase in demand for high-quality and engaging competitions. This increasing demand for CTF competition quality burdens CTF organizers to utilize behavioral science research, even non-consciously, to identify the fundamentals when developing successful CTF competitions. Compounding these challenges is the limited research focused on developing and organizing CTF competition.

This research has been motivated by the absence of models and methods for organizing successful CTF competitions. Consequently, our research leverages the field of design science research (DSR) to model CTF competitions systematically, including stakeholder motivations and requirements. Afterward, the focus shifts to constructing two robust, heuristic frameworks aimed at providing actionable insight for compe-

tition organizers, helping guide the multi-objective problem of stakeholder requirements.

The frameworks developed within the research target the two primary levels of most CTF competitions: challenge-level and competition-level design. Specifically, the challenge-level framework will focus on the challenge development life-cycle and the best practices for developing challenges. On the other hand, the competition-level framework will focus on how to organize challenges within the CTF competition to satisfy all participating stakeholders.

Through the development of these two frameworks, our contributions aim to provide the following benefits not previously seen in research:

1) This research proposes a formalized approach to defining "success" in CTF competitions.
2) This paper introduces actionable design principles that help balance stakeholder requirements.
3) This research consolidates current best practices for CTF challenge curation.
4) Through DSR a diverse set of methods and perspectives help enrich future CTF research.

Due to the broad nature of CTF competitions and formats, this paper will primarily focus on open Jepardy-style CTF competitions. More specifically, these competitions form the backbone of the CTF ecosystem, providing a competitive environment for all competitors, novice to professional [7]. These frameworks can still be used in advanced CTF competitions that primarily focus on experienced players largely, because they overlap with open Jeopardy-style CTF competitions, with an additional emphasis on the competitive nature. Additionally, Jeopardy has been the dominant CTF format over the past decade, showing its resilience. As a result, by focusing on this style of competition, this research can address the majority of competitions while still being able to provide specific, actionable insight.

The Jeopardy CTF ecosystem has been steadily developed over the past several years, becoming the standard in leading hosting platforms. One of these platforms, CTFd [8], provides organizers with a user-friendly toolkit for developing a robust platform for CTF competitions.

Later in the paper, the four primary stakeholders in CTF competitions will be discussed. However, due to the inherently complex strategies involved in the domain, this paper will limit the comprehensive analysis of the sponsor stakeholder. This was motivated by the understanding that different sponsors will have different requirements depending on the current organizational needs, and most of the discussion would be moot.

The remainder of this paper is as follows. Section two will provide a literature review of the current field. Sections three and four will utilize the stakeholder requirements to develop a heuristic framework for challenge and competition development. Section five will apply the heuristic frameworks to a case study, conducted by the authors. Section six will conclude with a discussion on future work.

II. Research Background and Related Works

Before going into the design of the framework, this section sheds some light on the previously published work. Our review begins with an overview of the various CTF competition styles, followed by a brief introduction to design science research.

A. CTF Competitions

CTF competitions have primarily been organized through two formats: Attack-Defend and Jeopardy. Attack-Defend originated from DEF CON, the computer security conference [1]. This structure of competition provides each team with a server that is connected to a network. Once the competition is started, a frenzy ensues as teams compete with campaigns of offensive cybersecurity operations on other teams, while concurrently trying to secure their server. Consequently, this competition aims to target both sides of computer security: attack and defense, hence the term attack-defend. However, this format is only truly effective when there is a small number of teams. For instance, Davis et al. [3] argue that attack-defend competitions tend to prioritize exploit automation instead of the previously discussed objectives, as the number of participating teams grows. As a result, competitions with a higher number of teams typically utilize the Jeopardy-style format. Specifically, this format provides participants with several different categories all with several challenges within them. As a result, teams can solve different challenges with varying amounts of points, which contribute to a running score. Through each team's score, a leaderboard can be constructed to form a competitive environment.

As of recently, there has been an increase in publications about CTF competitions. The field has been dominated by two core concepts: adapting CTF competitions for education [2], [5], [9] and analyzing previously hosted contests [1], [3], [10]–[12]. This research focused on applying CTF competitions in pedagogical environments leveraging DSR. This type of research is commonly used as a problem-solving paradigm, especially in dynamic environments (e.g., information systems). On the other hand, research focused on analyzing previously hosted contests utilizes behavioral science research. This adjacent field seeks to derive theories that explain phenomena. Unfortunately, these research methodologies oftentimes struggle to provide meaningful insight, when used in isolation. Specifically, Hevener et al. [13] argued the fields of behavioral science and design science utilize synergism to promote a complementary research cycle. Consequently, current research fails to address how to conceptualize CTF competition success and provide organizers with principles to guide competition and challenge development.

Over the years, researchers have been able to identify the utility of CTF competitions within the classroom. For instance, Leune and Petrilli [5] utilized CTFs to provide students with a gamified environment to learn computer security. In their study, students were provided access to several vulnerable targets to learn concepts ranging from cyber hygiene to privilege escalation. Additionally, other researchers have attempted to adapt CTF competitions for a digital audience. Several

students at the University of Arizona have utilized a lab-based approach to help decouple difficult topics allowing students to accelerate through difficult topics while avoiding the pons asinorum when approaching both topics concurrently [9].

Shifting away from academia, a slew of best practices can be found within previous competitions and their associated publications. For instance, MIT's 2014 CTF competition provided readers with an infrastructure diagram to spark discussion [3]. Other researchers have proposed novel methods for measuring integrity within the competition, by implementing anti-flag sharing methodologies [10], [11]. In addition to best practices, the University of Maryland proposed a challenging development life cycle: conceptualization, creation, and distribution [12]. However, these fail to address the bigger picture of developing a robust heuristic framework for analyzing competition.

B. Design Science Research

Design science research (DSR) has been commonly described as a problem-solving paradigm [13]. For instance, DSR aims to develop artifacts to help enhance the field for researchers and practitioners. Typically, DSR is coupled alongside behavior science research to help guide observations and theories through a symbiotic relationship. In other words, the application of DSR might greatly impact the ecosystem, causing subsequent behavioral science research to observe new phenomena supporting the application of DSR.

This paper utilizes DSR to help systematically build two frameworks geared toward open Jeopardy CTFs. Specifically, this research proposes a model to describe to problems of building successful CTFs, which provides two frameworks, or methods to navigate the problem space and follows up with an example instantiation to evaluate their effectiveness. In turn, the primary purpose of the frameworks is to provide practitioners with implementable suggestions to improve future CTFs while starting a discussion among researchers focused on developing CTF competitions.

CTF competitions place organizers and educators in a great position for cybersecurity education. Specifically, previous attempts to follow DSR when incorporating games within educational endeavors have been restricted to broad theory that encompasses ad hoc games developed by instructors [14]. However, CTFs provide a well-defined outline allowing for organizers and participants to both enjoy a familiar experience at each competition. The closest example of DSR being applied to the CTF deomain was by Carlisle et al. [12], in which they provided a three-stage framework for developing CTF challenges.

III. Foundational Principles of Competitions

Organizing successful CTF competitions is an inherently multi-faceted problem, allowing for a sprawling ecosystem of CTF competitions to offer diverse experiences to competitors. However, several CTF competitions regularly dominate the landscape (i.e., DEF CON CTF [1], CSAW CTF [15], etc.), while others struggle to garner attention among the community. This polarizing gradient of success begs the following question: What makes a competition successful? The first step towards providing a comprehensive answer to this question begins with rigorously defining success in CTF competitions. In this paper, we argue that there are four primary perspectives, or stakeholders, that help define success within CTF competitions: competitors, organizers, educators, and sponsors. By implementing a stakeholder approach, competitions can implement processes that satisfy stakeholders' needs, leading to long-term success.

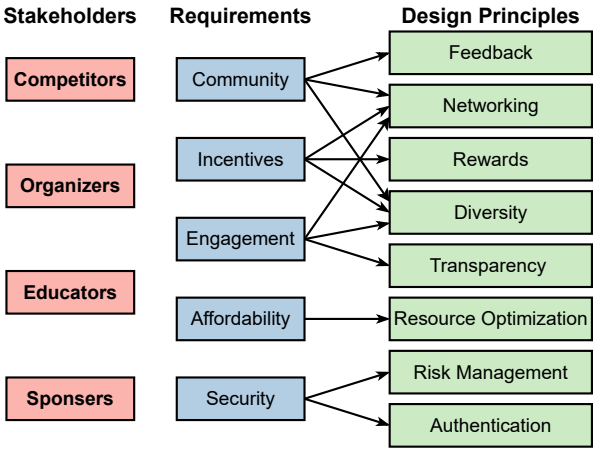


Figure 2: Framework for identifying common components of successful CTF competitions, based on the Design Science Research Methodology (DSRM) process model by Peffers et al. [16]. Stakeholder requirements are first identified and translated into abstract design principles.

Heuristic frameworks can be developed through two methodologies by systematically analyzing CTF competitions. The first is commonly seen through behavioral science, in which they identify key features found within successful CTFs. The second methodology, DSR, would recognize that successful CTFs must satisfy key design components. As a result, this can be seen as the reverse of behavioral science, by defining a problem and devising a solution. In this case, the frameworks would begin by identifying stakeholder motivations. These motivations provide crucial insight into their requirements. After distilling the stakeholder requirements, they can be categorized into abstract design principles, which will help the curation of an implementable feature set of successful competitions. This feature set can be further analyzed and refined through evaluation and case studies. Consequently, this method of framework development is more systematic and will typically provide a more robust and thorough solution.

A. Stakeholders

There are four key types of stakeholders in CTF competitions: competitors, organizers, educators, and sponsors. Each group plays a critical role in ensuring the success of these competitions. For instance, without competitors, there would be no participants to engage with the event, and without organizers, there would be no one to coordinate and host it.

Educators can be considered a specialized subset of organizers, whose primary focus is participants' educational advancement. Sponsors, on the other hand, participate selectively, contributing occasionally to support these competitions.

1) Competitors: Competitors are the foundational stakeholders in any CTF competition. From a business perspective, they can be seen as analogous to customers. Without their involvement, the event could not occur, as there would be no participants. As a result, competitors are regarded as the primary stakeholders that must be satisfied. Consequently, most design principles and competition features should cater primarily to this group due to the central role they play in the event's success.

2) Organizers: they are another essential group of stakeholders. Without them, CTF competitions would not materialize, as they are responsible for developing the challenges and managing the systems that facilitate these events. Prior research has identified several motivations for organizers, with the three most prominent being participant education, organizer education, and organizational development [12]. Organizers often aim to use CTFs as a pedagogical tool to supplement traditional learning methods. Another important motivation is the organizer's learning, as developing challenges allows them to deepen their understanding of a topic while creating engaging exercises. Finally, many organizers use CTFs to enhance their organization's brand, aiming to host recurring events as part of a long-term strategy or vision.

3) Educators: while similar to organizers, educators focus more narrowly on using CTFs in pedagogical settings. Their primary motivation is participant education, with less emphasis on the other objectives organizers may pursue. Educators typically design CTFs to maximize educational benefits, prioritizing learning outcomes over fostering competition among participants.

4) Sponsors: Sponsors are the final group of stakeholders. They provide financial support or services in exchange for two main benefits. The first is marketing and brand visibility, as they seek to achieve a return on their investment by promoting their brand through the competition. The second is talent identification and recruitment, where sponsors use CTFs as an alternative to technical interviews, identifying highly skilled participants for potential future employment opportunities.

B. Requirements

In a successful CTF competition, several key stakeholder requirements must be met. Basic requirements include fostering a positive user experience by emphasizing community, incentives, and engagement. Furthermore, with the recent rise in CTF popularity, affordability and security have become central to stakeholder priorities. Each of these requirements serves as a basis for the development of abstract design principles, as illustrated in Fig. 2.

1) Community: In successful CTF competitions, the development of a robust community plays a pivotal role in fostering meaningful interactions among stakeholders, yielding numerous advantages. For instance, a primary benefit of a vibrant community could be the networking opportunities provided for each stakeholder. Fostering a community is thus a requirement for a successful CTF competition.

2) Incentives: Another foundational requirement is seen in the incentives offered by the competition. These incentives can either be explicitly provided by the organizers such as monetary rewards for placements or implicit rewards earned through participation, such as skill development. In the first case, physical rewards can help draw a larger audience, especially with more skilled competitors. For the latter case, the competition can offer pedagogical benefits to the competitors through the utilization of well-developed challenges and is aimed primarily toward the less experienced player. In both cases, successful competitions must provide competitors with an incentive to participate in the competition.

3) Engagement: Through the distillation of community and incentives, the engagement requirement can be derived. Specifically, engagement aims to target how well the competition environment can entertain the competitors throughout the competition. As a result, this oftentimes requires a delicate balance between mindfulness when developing challenges and helping drive a thriving community (i.e., helping player engagement throughout the competition).

4) Affordability: In addition to the previous requirements, organizers should also focus on the affordability of the competition. Without sustainable budgeting, the competition may not come to fruition, due to the deficit in funding. As a result, competition organization needs to balance between the different expenses when hosting a competition, whether server infrastructure, development costs, or event expenses. Without adhering to the budget, the competition will be unable to provide the required resources to the participants, preventing the competition's success.

5) Security: An increasingly important requirement is security because the integrity of the competition is dependent on robust security policies. Specifically, without security consideration, the competition platform's security vulnerabilities might be trivially discovered, leading to compromised authentication and services. These compromised services might prevent other competitors from competing within the competitor causing other stakeholder requirements to not be fulfilled.

C. Design Principles

Design principles serve as a set of guidelines for competition organizers to help drive competition development through previously established values and considerations. Consequently, design principles must be directly derived from the stakeholder requirements, as they are being used as a guiding light for organizers in platform and challenge curation. For CTF competitions, there are typically eight dominant principles of design: feedback, networking, rewards, diversity, transparency, resource optimization, risk management, and.

1) Feedback: The first design principle is feedback, which enables organizers and educators to identify and address inefficiencies within CTF competitions. Feedback is most commonly gathered from post-competition surveys completed

by participants, offering a structured channel for communication between competitors and organizers. These surveys strengthen the surrounding community and provide organizers with valuable insights into participant concerns, guiding improvements for future events. Without feedback, organizers may risk misaligning their efforts with competitors' needs. These surveys are typically anonymous and administered on digital platforms, such as Google Forms, which encourage candid responses by protecting participants' identities. Furthermore, using established online tools allows organizers to efficiently analyze trends and categorize feedback, enhancing their capacity to make data-driven adjustments.

2) Networking: Another common design principle is networking, allowing for communication between the various stakeholders within the CTF competition. This design principle helps promote various interactions, whether between competitors, organizers, educators, or sponsors. For instance, without providing a platform for networking, discussing challenge solutions after the competition becomes less common, which can proportionately affect the effectiveness of fulfilling other design principles. This collaboration can be inferred to promote engagement among competitors due to the shift from a competition to a team environment. Another example of networking interaction could be between competitors and competition sponsors for professional development, allowing for talent acquisition within competitions, which can be seen as an incentive. Consequently, many of the competitions utilize various text and chat platforms, such as Discord, that provide support for instant messaging in a collaborative setting. Organizations can foster communication and a correlated community by utilizing these collaborative platforms.

3) Rewards: Beyond networking, competitions should offer rewards to competitors, whether explicitly or inherently. For example, many popular competitions might offer rewards through goods, services, or money. This may help attract more proficient individuals and teams to participate in a competition, due to the incentives offered by the final placement of teams. Likewise, some competitions might utilize placement rewards for exclusive CTF competitions, such as in the case of DEF CON CTF. However, rewards can take on a more subtle approach through skill development. For instance, well-developed challenges might provide pedagogical benefits to competitors who complete them, allowing for a more intrinsic reward. By incorporating a variety of rewards in the competition, organizers can help drive the user experience through this form of incentive.

4) Diversity: Although rewards may motivate participants for initial engagement with the CTF competition, diversity within challenges sustains user engagement. Within challenge development, there are two vectors of diversification: intra-category and inter-category. The intra-category diversification refers to the skill gradient present within a category. In other words, by providing tiers of challenges, various levels of users can engage with the competition. In many cases, this gradient ranges from trivial challenges to tricky technical problems. This helps drive engagement by providing competitors with a feedback mechanism to gauge progress through a given category. In some cases, it may be preferable to restrict the gradient to a certain end, such as in exclusive CTF finals. In this example, the challenges might favor technical rigor, due to the average audience's experience. On the other hand, if the competition is developed for beginners, it may make more sense to provide more beginner-friendly challenges.

In addition to intra-category diversification, organizers should also prioritize inter-category diversification. Specifically, by providing several different categories focusing on distinct topics within offensive cybersecurity, competitions can appeal to a larger audience. In most modern CTF competitions, four categories dominate the landscape: reverse engineering, binary exploitation, cryptography, and web exploitation. Additionally, many competitions add other categories depending on the challenge author's interests. Diversifying challenge categories allows for specialization to occur within the competition. This can be seen in larger teams, in which each person specializes in only a subset of the categories, allowing for a more rigorous understanding of the material in each of their specializations. Consequently, this can help foster collaboration among competitors within a team.

Diversification may also refer to the community, regarding how to incorporate accessibility and inclusivity into competition design. This type of diversification is beneficial, especially if the organizer's motivation includes education. This practice is common within similar competitions, such as competitive programming competitions, offering specialized events for field minorities, to draw more attention to the field.

5) Transparency: Transparency is another important design principle that should be considered while developing CTF competitions. By providing competitors with clear, concise communication and insights into the competition, a level of accountability is formed among the organizers. Specifically, by establishing rules and norms of the competition, a more engaging environment is fostered, because competitors will have certain expectations before competition within the competition. Transparency can also extend past the rules and involve any controversies or issues that arise dynamically within the competition. A common instance of issues occurs if there are bugs within a challenge or the competition infrastructure. By informing competitors of the ongoing issues, they can have greater insight into the hard work put forth by the organization hosting the competition.

6) Resource Optimization: Although challenge design has commonly been at the forefront of most organizers, logistics is another concern that should be held by organizers. Specifically, organizers should work on resource optimization, because by establishing sustainable practices with competition finances, competitions can run on a lower budget satisfying multiple stakeholders. As a result, this design principle can help guide organizers by reminding them of the investment each CTF competition requires, both monetarily, skill-wise, and time-wise.

7) Risk Management: Imagine a CTF competition's resources get compromised; how would that impact the com-

petition's reputation? Without proper risk management, a competition will struggle to satisfy the security requirements. There has been extensive research on risk management, along with several technical standards. A popular standard that helps guide practitioners through the software development lifecycle is IEEE 1540-2001. This technical standard proposes the risk management process model, identifying a systematic approach to identifying and managing a project's risk profile. While this standard may initially appear overly formal for CTF competitions, its underlying principles offer valuable insights. Applying these principles can help organizers proactively address potential risks associated with challenge design and platform deployment, thereby enhancing the competition's resilience and reliability.

8) Authentication: Due to the digital presence of CTF competitions and the importance of security, authentication is another foundational design principle. Authentication is important because several different teams are participating on the same platform. Consequently, without authentication, teams would be able to falsely claim their identity, and possibly negatively impact another team. This isn't in the spirit of CTFs, which causes the authentication design principle to be established.

IV. Challenge Framework

While the previous section discussed foundational principles outlining strategic goals for designing a successful CTF competition, we now turn our attention to the challenge framework, which provides a tactical approach to creating effective challenges that support a CTF competition.

The challenge development life cycle consists of six unique phases: 1) challenge ideation, 2) design specification, 3) implementation, 4) testing, 5) iteration, and 6) deployment (Fig. 3. *Challenge ideation* is where the foundations of the problem are theorized. The *design specification* phase utilizes the concepts theorized in the ideation phase to translate them into diagrams that help the developer implement the challenge. Once the challenge has been designed the *implementation phase* begins in which the initial proof of concept is developed. Once a proof of concept has been developed it is *iteratively evaluated* and revised until the final challenge has been produced. Once the challenge has been produced the challenge can be deployed onto the CTF platform.

The ideation phase is the most important. This phase consists of several steps that the author must thoroughly address before continuing to challenge development. The first and most fundamental issue that the author must address is which themes and categories the challenge is aligned with. For instance, jeopardy-style CTF competitions generally have several different categories. Each of these categories boasts a plethora of material to challenge authors to build from. For instance, binary exploitation focuses on the exploitation of vulnerabilities in compiled binaries. Similarly, reverse engineering shifts the focus from exploitation to understanding, with a greater emphasis on logic puzzles and obfuscated programs. Cryptography takes inspiration from both of the previous categories but applies them to vulnerabilities within cryptosystems. Finally, web exploitation shifts the attack surface from compiled binaries to the internet, setting up vulnerable websites and infrastructure, for competitors to wrangle with.

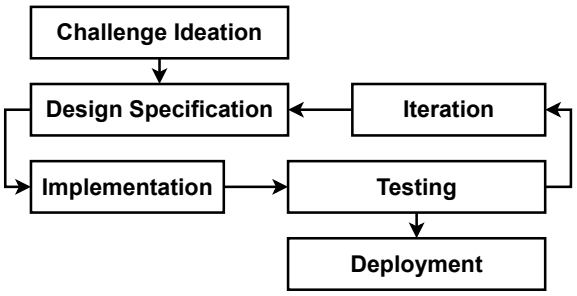


Figure 3: Challenge Development Life Cycle (CDLC) for Creating Robust and Engaging CTF Challenges. The CDLC comprises six phases: ideation, design specification, implementation, testing, iteration, and deployment. Once deployed, challenges are rarely reused, concluding the CDLC.

Typically challenges align with one of these categories, however, some challenges might align with multiple categories. Once a category has been selected, the author must derive a learning objective that they wish to achieve by developing their challenge. In other words, this learning objective can be seen as the foundation of the challenge. While deriving this learning objective, authors must be conscious of the audience participating in the CTF competition.

Once the challenge has been conceptualized the design specification phase is where the author transitions the concepts and learning objective into an intermediate representation of the final challenge. For instance, this intermediate representation might be provided through diagrams, documentation, or even through a simple computational mindset.

After designing a challenge the organizer should aim to produce a minimal viable product in which the core foundations of the learning objective are portrayed through interactive challenge. In this phase, organizers should not worry about deploying it on competition infrastructure, because of its unstable nature. During initial prototyping, the organizer should be more concerned about restricting unintended solutions and overall security.

After the initial prototypes have been developed the organizer should begin the testing and evaluation phase. Specifically, this is where a bulk of auxiliary development will occur, in which the organizer adds additional components to the challenge. After initial evaluation, the organizer should utilize a peer review to further bolster the challenge's quality.

After the challenge has been conceptualized, designed, prototyped, and evaluated it is ready to be deployed to the CTF platform. This will typically require the organizer to consider infrastructure requirements and integration within the platform, as well as designing and developing a challenge description, hints, and solutions.

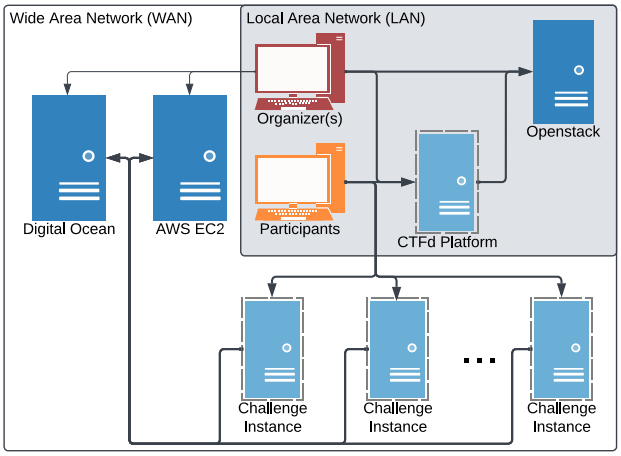


Figure 4: CTF competition system architecture: The organizers deployed a CTFd instance on a local OpenStack environment, while challenge servers were hosted externally using DigitalOcean and AWS EC2 instances.

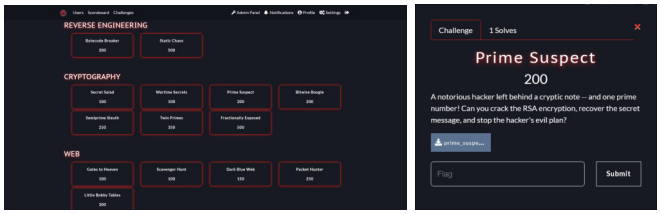
A. Best Practices in Challenge Development

Developing CTF challenges is similar to writing an entry in an encyclopedia. Challenge organizers must work together and coordinate throughout the challenge development life cycle to create a cohesive repository for the competition. In light of this, it may be beneficial for organizers to be vocal with other organizers throughout the process, especially after the ideation and deployment of the challenge. Maintaining effective communication regarding the current and future shape of the challenge repository can help satisfy the diversity design principle. Additionally, many CTF competitions publish their challenge repository after the competition has concluded. For instance, Google CTF publishes competition challenge sources on its GitHub [17] repository [18].

In addition to coordination among organizers, software development should also prioritize well-established software engineering practices [19]. Although following standard practices might not seem critically important, due to the small size of the challenges, the importance is exemplified in failures during deployment. Specifically, many competitions will inevitably face an issue with a given challenge during the CTF competition, whether an unintended solution, resource throttling, etc. By following standard practices during challenge development, authors can, in theory, identify the root cause of the issue promptly. This practice, in turn, can be correlated with the design principle of risk management.

V. Case Study with Intra-College Competition

We conducted a case study on an intra-collegiate CTF competition for framework evaluation. Specifically, the CTF competition was developed for an introductory college class to help students get introduced to the field of computer security through a gamified environment. This competition was structured to have all four stakeholders represented, allowing for framework evaluation. For instance, the enrolled students



(a) Competition challenge board (b) Crypto Challenge

Figure 5: Example CTF competition platform featuring multiple categories. Figure 5b depicts a specific challenge, including a description, Python script, and flag submission field.

can be seen as the participants, while the students developing the challenges would be the organizers, the class's professor would be the educator, and the cybersecurity department was the competition's sponsor.

A. Application of Challenge Framework

We utilized the agile software development methodology through a privately hosted GitLab [20] instance to fully utilize the framework. This infrastructure allowed the organizers to develop a challenge repository, utilizing version control software. The repository is structured hierarchically, with the base directory containing sub-directories for each category. Within each of these directories, each challenge was assigned a directory. This hierarchical approach, optimizes version control, by preventing any merge conflicts from arising.

In addition to traditional git functionality, GitLab offers a myriad of auxiliary features, helping facilitate agile development. For instance, the organizers utilized the kanban board in conjunction with issues to help actively manage challenge development. By creating issue tags for each phase of the development lifecycle, organizers can visualize the current state of each challenge. Likewise, issues were also tagged with the corresponding categories further helping to visualize category distributions.

Shifting from development to deployment, the organizers utilized CTFd [8] as the competition platform. Specifically, CTFd is an open-source project, offering a user-friendly framework for building CTF competition. As a result, the organizers utilized CTFd's control panel to register the pre-developed challenges, alongside managing auxiliary competition details. For dynamic challenges, that required a backend service for participants to interact with, the organizers utilized the GitLab container registry to manage challenge images (i.e., binary exploration, web exploitation, and cryptography). As a result, challenge deployment to the competition's infrastructure was streamlined due to the modularity of the challenge base.

In regards to infrastructure, the team utilized cloud-based hosting which enabled them to have a scalable infrastructure at a relatively low cost. Specifically, the team utilized Digital Ocean to host the CTFd platform alongside any interactive challenges within the competition. Since this was an introductory CTF competition they prioritized web exploitation due to the low tooling requirements required for participants, and straight away from binary exploitation due to the steep

learning curve involved with that category. To ensure that the challenge servers add adequate security they utilized docker and other virtualization technologies to sandbox each different challenge.

During the event, they utilized Discord as the primary networking platform. This allowed interested students to inquire more about cybersecurity and provided a centralized communication source between participants, organizers, and cybersecurity faculty.

B. Framework Evaluation

Through the development of the CTF competition challenges and infrastructure, both frameworks were able to be evaluated in a critical light. The challenge development framework, in practice, appears to augment the traditional agile development lifecycle. However, in agile development, the software development lifecycle is localized to the design, implementation, testing, and iteration of the challenge, which also terminates with deployment. Additionally, due to the lifecycle similarities, several of GitLab's agile-oriented features were able to be used seamlessly, allowing the organizers to visualize their progress on each challenge.

Regarding the competition level framework, the design principles helped guide the organization of infrastructure and challenges systematically. For instance, the diversity design principle helped prioritize the value in a diverse challenge repository. Similarly, resource optimization helped guide the selection of cloud infrastructure providers and services utilized by the organizers. In the system architecture diagram, the team utilized digital ocean AWS EC2 and a locally hosted open stack instance to host challenges and the CTF platform. Additionally, all the challenge instances were hosted externally, outside the local area network, this design choice helped the organizers manage the risk involved with creating vulnerable infrastructure. This application of the competition framework not only ensured a structure and secure environment but also demonstrated the framework's adaptability in addressing both logistical and strategic challenges, in a more comprehensive CTF competition.

VI. Conclusion

The rise in popularity of CTF competitions has led to a demand for relevant research to be conducted on the CTF ecosystem. However, previous research has struggled to provide robust and systematic frameworks for general competition and challenge development. Our paper analyzed the myriad of components within successful CTFs and developed two frameworks that have been evaluated in an applicable case study.

The first framework focused on the competition-level organization by first identifying the primary stakeholders in CTF competitions. Once identified, the stakeholder's needs and requirements were identified, which allowed for abstract design principles to be derived. Utilizing the requirements and design principles, a concise and robust set of features was identified. Each of these tangible features helps foster an environment where CTF competitions can thrive.

The second framework describes the challenge development lifecycle, in which developers begin with the ideation of a challenge. This is then complemented through the design, implementation, and testing of the challenge. Finally, the lifecycle is concluded through challenge deployment.

Overall, CTF competitions are inherently complex, and design science research can be readily applied. This research should serve as a starting point for fostering a more compressive discussion on how to improve CTF competitions through abstract and applicable science, instead of ad hoc practices seen in the modern landscape.

References

[1] C. Cowan, S. Arnold, S. Beattie, C. Wright, and J. Viega, "Defcon capture the flag: Defending vulnerable code from intense attack," in *Proceedings DARPA information survivability conference and exposition*, vol. 1, pp. 120–129, IEEE, 2003.

[2] E. Trickel, F. Disperati, E. Gustafson, F. Kalantari, M. Mabey, N. Tiwari, Y. Safaei, A. Doupé, and G. Vigna, "Shell we play a game?{CTF-as-a-service} for security education," in *2017 USENIX Workshop on Advances in Security Education (ASE 17)*, 2017.

[3] A. Davis, T. Leek, M. Zhivich, K. Gwinnup, and W. Leonard, "The fun and future of {CTF}," in *2014 USENIX Summit on Gaming, Games, and Gamification in Security Education (3GSE 14)*, 2014.

[4] R. Raman, S. Sunny, V. Pavithran, and K. Achuthan, "Framework for evaluating capture the flag (ctf) security competitions," in *International Conference for Convergence for Technology-2014*, pp. 1–5, IEEE, 2014.

[5] K. Leune and S. J. Petrilli Jr, "Using capture-the-flag to enhance the effectiveness of cybersecurity education," in *Proceedings of the 18th annual conference on information technology education*, pp. 47–52, 2017.

[6] CTFtime team, "CTFtime." https://ctftime.org/. Accessed: 2024-11-01.

[7] C. Taylor, P. Arias, J. Klopchic, C. Matarazzo, and E. Dube, "{CTF}:{State-of-the-Art} and building the next generation," in *2017 USENIX Workshop on Advances in Security Education (ASE 17)*, 2017.

[8] CTFd LLC, "CTFd." https://ctfd.io/. Accessed: 2024-11-01.

[9] C. Nelson and Y. Shoshitaishvili, "Pwn the learning curve: Education-first ctf challenges," in *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*, pp. 937–943, 2024.

[10] P. Chapman, J. Burket, and D. Brumley, "{PicoCTF}: A {Game-Based} computer security competition for high school students," in *2014 USENIX Summit on Gaming, Games, and Gamification in Security Education (3GSE 14)*, 2014.

[11] J.-G. Park, S.-H. Choi, H.-i. Kim, and H. Dowon, "Our experiences on the design, build and run of ctf," in *The 4th International Conference on Next Generation Computing*, 2018.

[12] B. Carlisle, M. Reininger, D. Fox, D. Votipka, and M. L. Mazurek, "On the other side of the table: Hosting capture the flag (ctf) competitions," in *Proceedings of the 6th Workshop on Security Information Workers, ser. WSIW*, vol. 20, 2020.

[13] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design science in information systems research," *MIS quarterly*, pp. 75–105, 2004.

[14] A. Amory, "Game object model version ii: a theoretical framework for educational game development," *Educational Technology Research and Development*, vol. 55, pp. 51–77, 2007.

[15] K. Chung and J. Cohen, "Learning obstacles in the capture the flag model," in *2014 USENIX Summit on Gaming, Games, and Gamification in Security Education (3GSE 14)*, 2014.

[16] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *Journal of management information systems*, vol. 24, no. 3, pp. 45–77, 2007.

[17] GitHub Inc., "GitHub." https://github.com/. Accessed: 2024-11-01.

[18] G. LLC, "Google ctf." https://github.com/google/google-ctf, 2017.

[19] R. C. Martin, *Clean code: a handbook of agile software craftsmanship*. Pearson Education, 2009.

[20] GitLab Inc., "GitLab." https://gitlab.com/. Accessed: 2024-11-01.