JAIST Repository
https://dspace.jaist.ac.jp/

Title	Usability Evaluation of Open Source and Online Capture the Flag Platforms
Author(s)	Noor Azam, Muhammad Harith bin; Beuran, Razvan
Citation	Research report (School of Information Science, Graduate School of Advanced Science and Technology, Japan Advanced Institute of Science and Technology), IS-RR-2018-001: 1-28
Issue Date	2018-07-09
Type	Technical Report
Text version	publisher
URL	http://hdl.handle.net/10119/15342
Rights	
Description	リサーチレポート（北陸先端科学技術大学院大学先端科学技術研究科情報科学系）

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

Usability Evaluation of Open Source and Online Capture the Flag Platforms

Muhammad Harith bin Noor Azam, Razvan Beuran

School of Information Science
Japan Advanced Institute of Science and Technology

July 9, 2018
IS-RR-2018-001

Usability Evaluation of Open Source and Online Capture the Flag Platforms

Muhammad Harith bin Noor Azam, Razvan Beuran

Japan Advanced Institute of Science and Technology

1. INTRODUCTION

Nowadays, cyber-attacks happen daily in every part of the world and their number is rising every year, making it a huge challenge for organizations that want to ensure their data is protected. Cyber-attacks should be dealt with in the most appropriate manner, such as prevention, mitigation, avoidance or acceptance, hence security professionals and IT personnel must be trained accordingly. One way of security training is through cybersecurity Capture the Flag (CTF) events, which are competitions between security professionals or students learning about cybersecurity that are intended as a learning tool to help sharpen their skills (Harmon, 2018).

There are two main types of CTF, Attack-Defend and Jeopardy. In Attack-Defend CTFs, a team attacks the other team's system, as well as defend their own system (Harmon, 2018). Usually there are two rounds, with one team attacking and the other team defending in the first round, then switching sides for the second round. The attacking team attempts to find flags (text files, images, etc.) in the defending machines as they compromise them. Various hacking tools can be used in order to compromise the defending machines, but there are rules in place to ensure that the teams are not at an advantage with respect to each other. The defending team can do anything within the rules to defend their machines against the attacking team, but they are not allowed to disable any network connections or turn off the machines. If there is any rule violation, the team incurs a penalty or is disqualified.

The Jeopardy-style CTF is similar to an actual Jeopardy game, as the competition board looks like a Jeopardy board with different categories and point values (Harmon, 2018). Categories include web security, cryptography, steganography, and so on; the goal of the Jeopardy-style CTF is again discovering a flag, which needs to be entered into the scoreboard in order to get the corresponding points. Flags are obtained from files provided by the organizers (e.g., by decrypting or processing them in some manner), or by accessing servers set up especially for the CTF (for instance, via SQL injection, privilege escalation, etc.). A timer is used to start and stop the CTF, and once the allocated time elapses, the game is over. The team/participant with the most points at the end of the competition wins. Note that there can be more than two teams in a Jeopardy-style CTF, as the participants are not trying to attack each other.

CTFs are widely used today, either conducted with all the participants at the same location, or online. They serve both as security competitions and for educational purposes. CTFs can be used to enhance the cybersecurity training process, both through their gamification elements, and due to the possibility to update the content on a timely basis in order to ensure that it is relevant to current requirements and needs.

In this research report we shall evaluate the most used CTFs within two categories, open-source CTFs, which can be installed locally, and online CTFs, which can be accessed via the Internet. Our main goal is to provide a guideline for choosing the most appropriate CTF for a given purpose, thus assisting anyone who may want to set up one. Section 2 discusses the open-source CTF platforms, while online CTF platforms are discussed in Section 3. Then, in Section 4, a usability evaluation of these platforms is presented. The report ends with conclusions, acknowledgments, and appendices.

2. OPEN-SOURCE CTF PLATFORMS

Many people and organizations share their work regarding the development of CTF platforms, mostly through GitHub, a well-known site for open-source software. This section presents a survey of four popular CTF platforms that have been published on GitHub: FBCTF, CTFd, Night Shade, and Mellivora.

2.1. FBCTF

The Facebook CTF (FBCTF, 2018) is a platform for hosting Jeopardy, but also "King of the Hill" style CTF competitions, in which teams must plant their tags at locations where scoring bots can find them and award points to teams accordingly. FBCTF can be used for organizing competitions with as few as two participants, all the way up to several hundreds. The platform was designed with flexibility in mind, allowing for different types of installations depending on the needs of the end user. The software can be installed either in Development Mode, or in Production Mode. It is developed by using PHP as programming language and MySQL for its database; the developers recommend to use Ubuntu 16.04 LTS operating system, for which a quick setup is possible. FBCTF has many features such as a timer, a variety of supported languages, detailed configuration pages and game logs. It also enables the administrator to import and export various files that are needed during events, which simplifies the process of creating challenges for the participants. Its strength is the flexibility of the platform that allows different types of installation, and a variety of configuration items that can be set up via the administration panel.

2.2. CTFd

CTFd is a Jeopardy-style CTF framework mainly focusing on ease of use and customizability (CTFd, 2018). It comes with everything that is needed to run a CTF and it's easy to modify with plugins and themes. CTFd is built using the Python language and uses MySQL as its database. The visualization of this platform is user friendly, with graphs, pie charts and other infographics that let users know

everything that happened during an event at a glance. The provided administration panel allows easy control without database queries. CTFd also has its very own HTML editor that can be used to create a FAQ, contact page or any other page needed for the events. CTFd supports SMTP and Mailgun protocols for messaging. It also enables the automation of starting and ending the competition. The strength of this platform lies in the flexibility of content creation, as the administrator has full control over how the platform should be depending on the training purpose.

2.3. MELLIVORA

Mellivora is another CTF platform for which the source code is shared publicly. It hosts Jeopardy-style CTFs, and its engine is written in PHP, using MySQL for the database, with PHP 5.5.9+, MySQL 5.5+ and Apache 2.4+ being recommended (Mellivora, 2018). Mellivora supports local storage or Amazon S3 for challenge file upload, and also provides a reCAPTCHA system to ensure that a computer user is human, so as to protect the website from bots. It also provides SMTP email support for sending both bulk and single emails. Mellivora has internal logs that can be used to catch any exceptions that may happen during an event. The administrator can also publish news from time to time to provide additional information to all the participants. The event organizer can also create signup restrictions based on regular expressions for email addresses. Overall, the Mellivora platform is very lightweight and fast, however it does not provide infographics to ease the task of monitoring participant progress.

2.4. NIGHTSHADE

NightShade is another CTF platform that can host Jeopardy-style CTFs. It is developed using Python as programming language and MySQL for the database (NightShade, 2018). It is a simple platform that combines a leaderboard with challenge questions. In addition to the normal Jeopardy style, NightShade also provides “traditional style” and “blind style” CTFs. For normal Jeopardy style, challenges are organized into categories such as cryptography, web, networking, and the number of points for each challenge is shown. In contrast, traditional style only provides the challenge names with their points, hence participants must guess the type of the challenge. The blind style only provides the challenge names, without showing the number of points, hence participants need to also guess the difficulty of each challenge. NightShade also includes user profiles that show for each user which contests he/she has joined, and the list of already solved questions. Another feature provided by this platform is displaying a list of all the competitions hosted by the administrator. NightShade’s strength lies in the variety of contest styles that can be chosen from, but the platform interface is too simple, without any infographic, which makes it hard for the administrator and participants to analyze the current status of competitions.

In Table 1 on the next page we show an overview of the four open-source CTF platforms discussed so far.

Table 1: Open-source CTF Platforms

No.	CTF Platform	CTF Style	Setup	Features	Advantages	Disadvantages
1.	FBCTF	Jeopardy and "King of the Hill"	PHP + MySQL	• Configuration page • Language selection • Import and export various elements of a CTF event • Game logs and timer	• Flexible installation	• Heavy platform with massive graphics
2.	CTFd	Jeopardy	Python + MySQL	• Team management • Customize everything using plugin and theme interfaces • Import and export CTF data for archival • Create your own challenges, categories, hints, and flags from the admin interface	• Easy to deploy • Customization features	• May be difficult to master all of the customization mechanisms
3.	NightShade	Jeopardy (normal, traditional, blind)	Python + MySQL	• List of contests • User profile page • Variety of styles of competition to choose from	• Variety of supported Jeopardy CTF styles	• No infographics allowing the administrator and participants to analyze the current status
4.	Mellivora	Jeopardy	PHP + MySQL	• Local storage or Amazon S3 for challenge file upload • Optional total number and time-based submission throttling • Internal logs for catching exceptions • reCAPTCHA and SMTP email support	• Lightweight and very fast	• No infographic for monitoring participant progress

2.5. OPEN-SOURCE CTF PLATFORM OF CHOICE

There are two main groups of learners that need to be considered as target for CTF events, namely learners in a specific organization and learners from various demographical areas. A suitable CTF platform needs to be chosen depending on the respective needs of these users.

2.5.1. CTF for Specific Organizations

Based on our investigation so far, CTFd seems to be the most suited open-source platform for cybersecurity education and training within a given organization. It is easy to deploy and customize, and can be used to host Jeopardy-style events, which are the fundamental type of CTF. CTFd provides everything that is needed to run a CTF, with a wide range of customizations compared to other platforms, including by using plugins and themes. From an organization perspective, it is important to use a platform the learners are comfortable with; therefore, the ease of use and customizability of CTFd are important aspects to be considered.

The visualization features of CTFd are also good, with graphs, pie charts and other infographics that let users know everything that happened during the event. On the other hand, Mellivora and NightShade do not provide any infographic to help the user of the platform, while FBCTF uses heavy graphics that can burden the server infrastructure. By taking advantages of this platform visualization, the administrator can easily analyze the current situation at a glance.

Furthermore, CTFd provides an administration panel that allows controlling the database without using database queries. It also has its own HTML editor that can be used to create a FAQ, contact page or any other necessary pages. Thus, the administrator can handle an unexpected situation by publishing a new page without stopping the competition. None of the other platforms provide such features.

In conclusion, the strength of CTFd lies in the flexibility of the interface and content creation flow, where the administrator has full control over how the platform should be depending on the organization needs and user types (teaching cybersecurity to university students, improving the skills of company employees, etc.).

2.5.2. CTF for Various Demographic Areas

Our study has revealed that FBCTF is the most appropriate open-source platform for conducting CTFs with people that come from various demographic areas, especially in case they do not master well the *de facto* language of open-source cybersecurity training platforms, English.

FBCTF is a nicely polished and versatile platform that can be used to host both Jeopardy-style and “King of the Hill” style CTFs. The main strength of this platform is that it offers different types of installations depending on the needs of the end users and features of the server infrastructure used by the organizer, while the other platforms only offer one type of installation. FBCTF also offers multi-language support, which benefits the end users with insufficient English skills. Such a feature

is important in order to ensure a smooth learning process for trainees with various backgrounds. Administrators of other platforms need to modify the user interface manually if there is a need to use other languages, for instance, by adding new translations and resources to the platform.

In conclusion, by providing multi-language support and various types of installation and setup choices, FBCTF is the best platform for cybersecurity training with learners that come from various demographical areas.

2.6. OTHER OPEN-SOURCE CTF PLATFORMS

There are many other open-source platforms for hosting CTF competitions, such as iCTF, OpenCTF, picoCTF-Platform 2, mkctf, and RootTheBox (CTF frameworks, libraries, resources and softwares, 2018).

2.6.1. iCTF

The iCTF platform is used by UC Santa Barbara Seclab to host their CTF event. The framework creates several virtual machines (VM), one for the organizers and one for every team. The iCTF framework contains several components, such as central database, score bot, router, dashboard, VM creator and a standard format for creating services. This platform can host Attack-Defend type of CTFs.

2.6.2. HBCTF

The HBCTF platform can be used to host hybrid-style CTFs that combines a DevOps service hack and patch process, Jeopardy-style flags, and an explorable battlefield where players go head-to-head to control strategic network nodes. This platform is written in Python and was developed by HackBama, which is a group of information security professionals that are said to have many years of experience with large corporations and governments.

2.6.3. picoCTF

The picoCTF platform can be used to host Jeopardy-style CTFs. This platform was designed so as to be easily adaptable to other cybersecurity or programming competitions. The development team targets Ubuntu 14.04 LTS as the main target operating system, but according to them it should work on other Linux distributions, and even on Windows.

2.6.4. Root the Box

Root the Box is written in Python and can host Jeopardy-style CTF games. Different from other platforms, in Root the Box, teams can also create “botnets” by uploading a small bot program to target machines. The teams are periodically rewarded with (in-game) money for each bot in their botnet. Such money can be used to unlock new training levels, to buy hints to flags, to download a target’s source code, or even to “swat” other players by bribing the (in-game) police. Encrypted bank account passwords are publicly displayed, allowing players to crack each other’s passwords and steal each other's money. This platform is the one using the largest number of gamification elements to motivate competitors and make training more fun.

2.6.5. HackTheArch

HackTheArch was developed using the web-application framework Ruby on Rails by the Military Cyber Professionals Association (MCPA). This platform is based on picoCTF, which was extended with features such as offering competitors hints at a cost, and the ability to create/modify problems from the web interface.

3. ONLINE CTF PLATFORMS

Cybersecurity skills are in very high demand given that even the most basic devices start to be connected to the Internet. Consequently, several websites provide online cybersecurity training using CTF style, and some of the most popular such resources are Hack The Box, WTHack, BackDoor, and Hack This Site.

3.1. HACK THE BOX

Hack The Box is an online platform which allows a user to test his/her penetration testing skills and exchange ideas and methodologies with other members of similar interests (Hack The Box , 2018). The platform contains several challenges that are constantly updated, some of them simulating real-world scenarios, and others leaning more towards an Attack-Defend style of challenge. This website gives users opportunities to complete challenges and prove their skills, as well as connect to a private network, called HTB Labs. This network consists of a number of virtual machines, currently 54 in total, which are set up as targets to be hacked. By hacking these machines or each other's VM, users get points that help them to advance in the Hall of Fame. Note that in order to be able to join the Attack-Defend style CTFs, new users need to solve entry-level challenges in order to demonstrate their skills. Hack The Box also takes into account the need for communication between users, as this website provides a forum for competitors to discuss problems and solutions among themselves. Although the platform provides a large selection of penetration testing challenges, it does lack challenges related to other domains.

3.2. BACKDOOR

BackDoor is an online CTF hosting Jeopardy-style CTF (BackDoor, 2018) conceived as a platform for hackers to demonstrate their talent in a competitive environment. Initially launched only within the Indian Institute of Technology Roorkee campus, it has been made available recently for anyone over the Internet. This platform hosts many competitions for its users, such as "n00bctf2018" and "BackdoorCTF 2018". "BackdoorCTF" is the annual flagship CTF competition conducted by SDSLabs and InfoSecIITR, while "n00bctf2018" is a beginner-level CTF, targeting mainly first year students; the top two participants at the end of "n00bctf2018" are automatically entered into the final round of the SDSLabs competition. Although BackDoor provides a practice arena for advanced users, as well as a beginner-level challenge to prepare novice learners for real CTF competitions, it lacks a wide range of challenges as provided by other platforms, such as Hack The Box.

3.3. HACK THIS SITE

Hack This Site is a free, safe and legal training ground for users to test and expand their hacking skills (Hack This Site, 2018). This platform provides an open learning environment via a series of hacking challenges, articles, resources, and discussions of the latest events in hacker culture. There are two types of challenges provided, namely "basic challenges" for beginners, and "realistic missions" that employ an Attack-Defend style for more advanced users. Basic challenges are relatively straight-forward and are designed to outline the fundamentals of a hacker's first steps into the world of web hacking, while realistic missions provide hacking objectives that target websites with built-in security flaws, which are made available via the platform.

Hack This Site uses a role-playing game approach to motivate players and make the competition more entertaining. Thus, each user plays the role of a freelance hacker who is contracted by several individuals and organizations to hack for social justice causes. The user is given objectives and is left to explore the site on his/her own, trying to discover and exploit the vulnerabilities. The web hacking skills learned in this series of challenges can be directly applied to systems in the real world. Hack This Site also provides different ways of communication, from private messaging to a forum for all users. The main advantages of using this platform are the relevant information provided to users, and the different ways of communication available, which can enhance the learning process. However, the web interface is not very well designed, with the small font size making it difficult to use.

3.4. WTHack

WTHack is a Jeopardy-style CTF platform that allows users to answer the challenges provided, as well as enables them to add challenges for other users (WTHack, 2018). Thanks to this feature, the number of challenges on this platform keeps growing, thus enabling learners to always discover fresh challenges. WTHack also provides an instant messaging platform for its users, which is based on the Telegram application, so as to ensure that discussions among them can be done privately. The challenges of WTHack are grouped into 5 main categories, namely: Web application security, Cryptography, Forensics, Reverse engineering, and Scavenger hunt. All the challenges that are not related to these categories are placed into the Miscellaneous group. Although the WTHack website provides some basic features such as a leaderboard section, it is not very comprehensive, as it lacks statistical data for users' activity.

Table 2 on the next page shows an overview of the four online CTF platforms that we discussed up to this point.

Table 2: Online CTF Platforms

No.	CTF Platform	CTF Style	Features	Advantages	Disadvantages
1.	Hack The Box	Attack-Defend	• 54 virtual machines to be hacked • Members can hack each other's VMs • Offline challenges such as steganography, reverse engineering and cryptography • User forum	• Big selection of penetration testing challenges	• Lacks challenges related other domains
2.	BackDoor	Jeopardy	• Challenges are classified by tags • Includes easy challenges for welcoming beginners to CTF events	• Used in actual official competitions • Beginners can learn the concept of Jeopardy-style CTF	• Lacks a large number of challenges
3.	Hack This Site	Jeopardy and Attack-Defend	• Different range of vulnerabilities to be exploited, including realistic cases • Several ways of communicating, from private messaging to forum • Includes latest information about current vulnerabilities	• Good background information is provided for novice learners • Different communication ways among users can enhance the learning process	• Online interface is difficult to use due to small font size
4.	WTHack	Jeopardy	• Users can add challenges for other users • Telegram group for private communication	• Users can communicate via a secure instant messaging system	• Lacks statistical data for user activity

3.5. OTHER ONLINE CTF PLATFORMS

There are many other online platforms that can be used as cybersecurity education and training tools. Among those which employ a CTF style, we mention here Exploit Exercises, pwnable.kr, Smash The Stack, and W3Challs (CTF frameworks, libraries, resources and softwares, 2018).

3.5.1. Exploit Exercises

The Exploit Exercises platform provides several virtual machines, as well as documentation and challenges that can be used to learn about a variety of computer security issues, such as privilege escalation, vulnerability analysis, exploit development, debugging, reverse engineering, and general cybersecurity issues. The platform doesn't require any specific experience level to conduct the included challenges, but their content needs to be downloaded from the website.

3.5.2. pwnable.kr

The pwnable.kr platform is a non-commercial wargame site which provides various "pwn" type of challenges regarding system exploitation. The main purpose of pwnable.kr is to make the experience enjoyable, therefore it contains many graphics to make it more fun to be engaged. The challenges are divided into four categories: Toddler's Bottle, Rookiss, Grotesque and Hacker's Secret, with increasing difficulty levels. For each challenge one can display the author's solution, however, often it is possible for users to find alternative solutions as well.

3.5.3. Smash The Stack

The Smash The Stack platform hosts several wargames, i.e., ethical hacking environments that support the simulation of real-world software vulnerability theories or concepts, and allow for the legal execution of exploitation techniques. The term software in this context can represent operating systems, network protocols, or user applications. To access the wargames in Smash The Stack, one only needs an ssh client, as each challenge has its own set of connection details which are made available on the platform's webpage.

3.5.4. W3Challs

The W3Challs online CTF platform hosts penetration testing sessions, offering various computer challenges in categories related to cybersecurity, such as hacking, cracking, wargame, forensic, cryptography, steganography and programming. According to the platform rules, users are not limited to testing themselves against the challenges that are provided but can even try to hack the site itself. Nevertheless, brute-forcing or Denial of Service attacks are forbidden.

4. USABILITY EVALUATION

Usability refers to the quality of a user's experience when interacting with products or systems, including websites, software, devices, or applications. Hence, usability is about effectiveness, efficiency and the overall satisfaction of the user. In its turn, *usability evaluation* refers to assessing how well users can learn and use a product to achieve their goals, and how satisfied they are with that process (Usability Evaluation Basics, 2018). A variety of methods can be used to gather this information; in this research, we use two of the most well-known evaluation methods, as follows:

- *Criteria-based assessment* is a methodology that assesses usability aspects based on the documentation that is provided by the system developer. We have selected it for evaluating the open-source CTF platforms because hardware limitations and time constraints prevented us from directly installing and running all the software. This evaluation is discussed in Section 4.1.
- *System Usability Scale* (SUS) is an assessment method that uses a series of 10 statements covering various usability aspects to provide an objective score that quantifies a system's usability. Since online CTF platforms are readily usable over the Internet, we have selected SUS for evaluating them, as it will be detailed in Section 4.2.

4.1. USABILITY ASSESSMENT OF OPEN-SOURCE CTF PLATFORMS

The criteria-based assessment method makes it possible to evaluate the usability of a system based on its documentation (Software Evaluation: Criteria-based Assessment, 2018). This assessment involves checking whether the software, and the project that develops it, conforms to the various characteristics and exhibits the properties that are expected of sustainable software; the more characteristics are satisfied, the more sustainable the software is considered to be. The assessment criteria are grouped into four categories: capability to understand the system, the documentation itself, capability for installation, and capability to learn the system. The assessment sheets for each of the open-source CTF platforms that we evaluated are included in Appendices A to D of this report.

4.1.1. FBCTF Results

Regarding FBCTF (see Appendix A), we conclude that in terms of capability to understand the platform, FBCTF does not provide high-level information, such as what the platform does and what it is for, but only brief information about it. No case studies, no intended use cases and no architectural overview are included.

The documentation, although not very thorough, is partitioned into sections for the developer and administrator of the platform. The only section that is not provided is for the end user. The documentation seems to be written mostly from the developer's point of view and does not provide further information regarding the platform. On the positive side, the documentation is task-oriented and consists of a set of clear, step-by-step instructions for platform installation by developers and for configuration by administrators.

From the capability for installation point of view, we noticed that the website of FBCTF provides step-by-step instructions for installing the platform, but it only provides a short list of third-party dependencies for quick installation, even though there are multiple ways to install it. Moreover, there is no method provided in order to verify the success of the installation.

Lastly, in terms of capability for the user to learn how to use the platform, FBCTF only provides a "Getting Started" guide for administrators who need to set up the CTF competition, and does not provide basic use cases for the end user of the platform (i.e., the trainee).

4.1.2. CTFd Results

For CTFd (see Appendix B), we observe that in terms of capability to understand the platform, no high-level information about the platform is available, such as what it does and what it is for, but only brief information is provided. No case studies, no intended use cases and no architectural overview are presented.

In terms of documentation, CTFd does not provide a good overview of the platform, and the documentation is not partitioned into clear sections; moreover, no specific information is available for the end user. Although the documentation targets mainly the developers, it includes a list of resources with further information suitable for general users. The CTFd documentation is task-oriented and consists of step-by-step instructions for the system configuration by the administrator. On the other hand, troubleshooting information in case of problems and error messages, such as symptoms and step-by-step solutions, is not provided.

From the capability for installation point of view, the CTFd website provides clear step-by-step instructions for installing the platform, including with commands that can be copied and pasted, as well as a list of links for third-party dependencies, such as Docker. However, although the list of mandatory third-party dependencies is provided, the website does not provide information about the optional third-party dependencies that one may wish to install. In addition, no method is provided to verify the success of the installation.

Lastly, in term of capability for the user to learn, the CTFd provided "Getting Started" guide is intended for administrators who need to set up the CTF competition, and it does not provide end user information.

4.1.3. Mellivora Results

Regarding Mellivora (see Appendix C), again we conclude that, in terms of capability to understand the platform, the documentation does not include high-level information about it, such as what the platform does and what it is for, but only provides brief information. No case studies, no intended use cases, and no architectural overview are presented.

In terms of documentation, Mellivora does not provide a good overview of the platform, and the documentation is also not partitioned into sections, such as for the end user and administrator. The documentation is written mostly from a developer's

point of view and does not provides further information regarding the platform. Error symptoms and step-by-step solutions are also not included in the documentation. Additional information includes a series of screenshots of the Mellivora interface pages, provided however without any explanation. On the positive side, the documentation is task-oriented and consists of clear, step-by-step instructions for the configuration by the administrator.

From the capability for installation point of view, we conclude that the website of Mellivora provides step-by-step instructions for installing the platform, but it only provides a list of mandatory third-party dependencies for installing the tools, without links to those resources. Moreover, there is no provided method for verifying the success of the installation.

Lastly, in terms of capability for the user to learn, Mellivora provides a "Getting Started" Guide only for the administrator who needs to set up the CTF competition and does not provide use cases for regular users of the platform.

4.1.4. NightShade Results

For NightShade (see Appendix D), we observe that in terms of capability to understand the platform, detailed high-level information about it is not available, such as what the platform does and what it is for, but only brief information is provided. No case studies, no intended use cases, and no architectural overview are presented for the users.

In terms of documentation, NightShade does not provide a good overview of the platform, and the documentation is not partitioned into sections such as for the user and administrator. The documentation mostly uses the developer point of view, and does not provide further details regarding the platform. It also does not provide error symptoms and step-by-step solutions. Additional information is provided in the form of web interface screenshots; however, explanations are lacking. The documentation is nevertheless task-oriented and consists of clear, step-by-step instructions for system configuration by the administrator.

Regarding the capability for installation aspect, we notice that the NightShade website provides brief step-by-steps instructions for installing the platform, but it includes no lists of optional or mandatory third-party dependencies. Moreover, there is no method provided to verify the success of the installation.

Lastly, in terms of capability for the user to learn, NightShade provides a "Getting Started" guide only for the administrator who needs to set up the CTF competition, and it does not provide use cases for regular users.

4.1.5. Discussion

Considering the remarks about each of the evaluated open-source CTF platforms, we conclude that FBCTF is the most usable of the platforms, although not by a wide margin; this supports our recommendation in Section 2.5.2 as best CTF for various demographic areas. An evaluation summary per class of assessment criteria follows:

- In terms of capability to understand, all the platforms provide only brief information to the user, although FBCTF seems to be the most polished. None of the platforms provide any case studies or intended use cases, and also no architectural overviews are presented.
- In terms of documentation, all platforms provide only an overview description, and only FBCTF partitions the documentation into sections, namely for developers and administrators. None of the platforms provide error symptoms and step-by-step solutions in the documentation, however the documentation is written in a task-oriented manner and is sufficiently clear, including setup instructions.
- Regarding the capability for installation, all the platforms provide instructions for installing the software, CTFd being somewhat more thorough; NightShade instructions are the least detailed and lack information about dependencies. None of the platforms provide clear details regarding optional third-party dependencies. Furthermore, no platform provides a method to verify the success of the installation process.
- Lastly, in term of capability for the user to learn, all platforms provide a more or less detailed "Getting Started" guide, but these guides are intended for administrators who need to set up the CTF competitions, and no basic use cases are provided for the end users.

4.2. USABILITY EVALUATION OF ONLINE CTF PLATFORMS

As mentioned already, due to the fact that online CTF platforms are readily accessible over the Internet, we have used the System Usability Scale (SUS) methodology to assess them. In what follows we provide an overview of SUS, followed by the results for each of the evaluated platforms.

4.2.1. System Usability Scale

The System Usability Scale is a reliable, low-cost usability scale that has been widely used for global assessments of system usability (Brooke, 2013). SUS includes ten statements (see Table 3), each having a five-point scale that ranges from Strongly Disagree to Strongly Agree. These statements cover various aspects of system usability, such as the need for support, system complexity, etc., thus introducing a high level of validity for measuring the usability of a system.

SUS uses five positive statements and five negative statements, which alternate. Responses to the SUS questions indicate strength of agreement or disagreement, so strongly disagreeing with a negative statement is equivalent to strongly agreeing with a positive one. In the next paragraph we present the manner in which the overall SUS score is computed by taking into account the need to harmonize the positive and negative nature of the statements. Note that the SUS authors considered best to use a scale from 0 to 100 for the overall score, which makes it easier for people to interpret the results and gives the final score a sufficient range of values.

Table 3: System Usability Scale Statements

No.	Statement
1	I think that I would like to use this system frequently.
2	I found the system unnecessarily complex.
3	I thought the system was easy to use.
4	I think that I would need the support of a technical person to be able to use this system.
5	I found the various functions in this system were well integrated.
6	I thought there was too much inconsistency in this system.
7	I would imagine that most people would learn to use this system very quickly.
8	I found the system very cumbersome to use.
9	I felt very confident using the system.
10	I needed to learn a lot of things before I could get going with this system.

Score Calculation: Based on the above considerations, the method for calculating the final SUS score is as follows. For the positively-worded items 1, 3, 5, 7, and 9, the positive score contribution, *P*, is given by:

$$P = \sum_{i=1,3,5,7,9} (sp_i - 1)$$

where sp_i represents the scale position for statement *i*, with values between 1 for "Strong Disagree", and 5 for "Strong Agree". For the negatively-worded items 2, 4, 6, 8, and 10, the negative score contribution, *N*, is computed as:

$$N = \sum_{i=2,4,6,8,10} (5 - sp_i)$$

where sp_i has the same meaning as above. The overall value of SUS, *S*, is calculated from *P* and *N* by using the following formula:

$$S = (P + N) \times 2.5$$

According to (Orfanou, 2015), the SUS score interpretation is: (i) a value over 85 indicates a *highly usable* system; (ii) a score between 70 and 85 represents a *good to excellent* system; (iii) a SUS value from 50 to 70 shows that the system is *acceptable* but has some usability problems and needs further improvement; (iv) finally, a system with a SUS score below 50 is considered *unusable/unacceptable*.

Assessment Methodology: Our SUS assessment of online CTF platforms was conducted via a survey of 5 respondents. All of them were Computer Science undergraduate students from Universiti Sains Islam Malaysia (USIM) majoring in Information Security and Assurance. They all had at least some basic knowledge about cybersecurity, as all of them had attended related courses. Three of the respondents had joined CTF competitions organized in Malaysia at least once,

whereas the other two had never joined any CTF competition before the time they answered the questionnaire. The following subsections provide the detailed survey results for each of the evaluated platforms.

4.2.2. WTHack Results

The SUS results for our assessment of WTHack are shown in Table 4.

Table 4: SUS Scores for WTHack

Statement	Respondent Response				
	User 1	User 2	User 3	User 4	User 5
1	4	5	4	5	5
2	2	3	4	3	1
3	4	4	5	5	4
4	3	2	2	5	2
5	3	4	4	1	4
6	2	1	2	1	2
7	3	4	4	4	4
8	2	1	1	1	2
9	3	4	4	5	4
10	3	3	3	1	2
SUS Score	**62.5**	**77.5**	**72.5**	**72.5**	**80.0**
				Average SUS Score	**73.0**

4.2.3. Hack This Site Results

The SUS results for our assessment of Hack This Site are shown in Table 5.

Table 5: SUS Scores for Hack This Site

Statement	Respondent Response				
	User 1	User 2	User 3	User 4	User 5
1	4	4	4	5	4
2	3	1	4	1	2
3	4	4	4	5	4
4	3	2	3	3	3
5	4	4	5	4	4
6	2	2	2	1	2
7	3	5	4	2	3
8	3	1	2	4	3
9	3	4	4	2	4
10	3	2	3	4	2
SUS Score	**60.0**	**82.5**	**67.5**	**62.5**	**67.5**
				Average SUS Score	**68.0**

4.2.4. BackDoor Results

The SUS results for our assessment of BackDoor are shown in Table 6.

Table 6: SUS Scores for BackDoor

Statement	Respondents Response				
	User 1	**User 2**	**User 3**	**User 4**	**User 5**
1	2	5	4	5	5
2	3	3	4	3	1
3	3	4	3	5	4
4	3	1	3	1	2
5	2	5	4	5	4
6	2	2	2	1	1
7	2	4	4	2	4
8	3	1	2	2	1
9	2	3	3	2	4
10	3	2	4	4	2
SUS Score	**57.5**	**80.0**	**57.5**	**70.0**	**85**
				Average SUS Score	**70.0**

4.2.5. Hack The Box Results

The SUS results for our assessment of Hack The Box are shown in Table 7.

Table 7: SUS Scores for Hack The Box

Statement	Respondent Response				
	User 1	**User 2**	**User 3**	**User 4**	**User 5**
1	4	4	4	5	4
2	2	2	2	1	2
3	3	4	3	5	4
4	3	1	3	2	3
5	4	3	4	5	4
6	2	2	3	3	2
7	2	3	3	1	3
8	2	2	3	1	2
9	3	4	3	5	4
10	3	1	2	5	2
SUS Score	**60.0**	**75.0**	**60.0**	**72.5**	**70**
				Average SUS Score	**67.5**

4.2.6. Discussion

Considering the average SUS values shown in Tables 4 through 7, we observe that among the evaluated online CTF platforms WTHack has the highest average score, namely 73.0, followed by BackDoor with 70.0 and Hack This Site with 68.0; the lowest average score is 67.5 for the Hack The Box.

According to the interpretation of SUS mentioned in Section 4.2.1, WTHack and BackDoor are considered to have *good to excellent* usability, as their scores are above the threshold value 70. The other two CTFs, Hack This Site and Hack The Box fall in the category of *acceptable* usability, meaning that they have some usability issues and need further improvement. We note however that the differences

between the lowest and highest scores are less than 10%, hence the usability differences should be considered marginal.

The comparison of the individual scores for each statement shows some significant differences, especially regarding statements 7, 8, 9, and 10. For instance, the scale positions for statement 10 in Table 7 have values between 1 and 5, demonstrating a high variability in the way in which respondents evaluated the amount of knowledge they still need to acquire before using that system (Hack The Box). Such variability could be explained by the different experience and cybersecurity education background among the respondents. In conclusion, we consider that all the evaluated CTF platforms have passed the SUS assessment, with WTHack being the winner in terms of usability, albeit by a slight margin.

5. CONCLUSION

In this report we have conducted a survey of two categories of CTF platforms, namely open-source CTFs and online CTFs. In each category we have studied in detail several representative systems, and also provided an overview of various alternatives. Finally, we have conducted a usability evaluation of the analyzed platforms by following the criteria-based assessment and System Usability Score methodologies.

The open-source CTF platforms studied in detail were FBCTF, CTFd, Mellivora and Nightshade. FBCTF provides a wide range of installation choices, but the infrastructure might suffer from a heavy processing load due to the massive graphical interface; FBCTF was selected as the platform of choice for learners from various demographic areas due to its multiple language support features. CTFd offers easy deployment with a huge ability to customize the platform, although the administrator may need to remove some of the unneeded features in order to make the learning experience smoother; nevertheless, owing to its customization features, including via templates, CTFd was selected as the platform of choice for organizations trying to conduct CTFs. Mellivora's main advantages are related to it being a lightweight platform, however user experience is not so good as no infographics are provided to ease the monitoring the participants' progress. Although NightShade does not provide infographics either, it offers three different CTF styles that can be enjoyed by the learner.

The open-source CTF platforms mentioned so far have been evaluated using the method called criteria-based assessment, for which the evaluation sheets were included at the end of this report (Appendices A to D). Our conclusion is that there is still much work needed on the documentation of these platforms, especially since the currently available resources are more focused on the developer point of view. Use cases are not provided for any of the platforms to illustrate their applications. Test cases of the installation process are also not provided, making it difficult to ensure that the platforms are installed properly. On the positive side, step-by-step installation instructions are included in the documentation in all cases. Among these platforms, we have concluded that FBCTF is the most polished and easy to use.

The online CTF platforms that we have analyzed in detail are WTHack, BackDoor, Hack This Site and Hack The Box. WTHack provides an instant messaging platform to ensure a better communication experience for its users, but it lacks statistical data and infographics of user activity. BackDoor is recommended for beginners to learn more about Jeopardy-style CTFs, but the number of challenges that are offered on the corresponding website is limited. Hack The Box has a big selection of penetration testing challenges, with new ones being added occasionally; however, no challenges related to other domains are provided. Hack This Site supplies a variety of reading information related to cybersecurity and several ways of communication between the users who wish to discuss security-related issues; nevertheless, users of this platform may suffer due to the poor organization of information and the small font size used, which is not suitable for prolonged reading.

The online CTF platforms were evaluated using the System Usability Scale by five respondents who accessed them online. Although no significant score differences appeared, only two of the platforms, namely WTHack and BackDoor, were above the good or excellent usability threshold. The other two, Hack This Site and Hack The Box, fell into the acceptable category, meaning that they still need to be further refined in order to ensure that learners can fully enjoy using them. We also concluded that the previous experience and knowledge of the respondents have a significant impact on their assessments; hence, studies with a larger number of respondents who have representative backgrounds are needed in order to provide more accurate results.

ACKNOWLEDGMENT

The research presented in this report was conducted during the internship that Muhammad Harith bin Noor Azam, student at Universiti Sains Islam Malaysia (USIM), did at Japan Advanced Institute of Science and Technology (JAIST) during the period February-April 2018.

REFERENCES

BackDoor. (2018, June 21). Retrieved from https://backdoor.sdslabs.co/

Brooke, J. (2013). SUS: A Retrospective. *Journal of Usability Studies*, 29-40.

CTF frameworks, libraries, resources and softwares. (2018, June 21). Retrieved from https://github.com/apsdehal/awesome-ctf

CTFd. (2018, June 21). Retrieved from https://github.com/CTFd/CTFd

FBCTF. (2018, June 21). Retrieved from https://github.com/facebook/fbctf

Hack The Box. (2018, June 21). Retrieved from https://www.hackthebox.eu/

Hack This Site. (2018, June 21). Retrieved from https://www.hackthissite.org/

Harmon, T. D. (2018, June 21). *Cyber Security Capture The Flag (CTF): What Is It?* Retrieved from: https://blogs.cisco.com/perspectives/cyber-security-capture-the-flag-ctf-what-is-it

Orfanou, K., Tselios, N., Katsanos, C. (2015). Perceived Usability Evaluation of Learning Management Systems: Empirical Evaluation of the System Usability Scale. *The International Review of Research in Open and Distributed Learning.*

Mellivora. (2018, June 21). Retrieved from https://github.com/Nakiami/mellivora

NightShade. (2018, June 21). Retrieved from https://github.com/UnrealAkama/NightShade

Software Evaluation: Criteria-based Assessment. (2018, June 21). Retrieved from https://software.ac.uk/sites/default/files/SSI-SoftwareEvaluationCriteria.pdf

Usability Evaluation Basics. (2018, June 21). Retrieved from https://www.usability.gov/what-and-why/usability-evaluation.html

WTHack. (2018, June 21). Retrieved from https://www.onlinectf.com/challenges/

APPENDIX A

USABILITY ASSESMENT FOR FBCTF

Capability to understand

High-level description of what/who the software is for is available.	No, but it explains in brief
High-level description of what the software does is available.	No, but it explains in brief as this platform is used to organize CTF competition or events
Architectural overview, with diagrams, is available.	No, it doesn't provide any architectural overview of this platform
Descriptions of intended use cases are available.	No, it doesn't explain any use cases provided in this platform
Case studies of use are available.	No, it doesn't explain any case study

Documentation

Provides a high-level overview of the software.	No, it explains only in brief
Partitioned into sections for users, user-developers and developers (depending on the software).	Yes, but the only sections are for developer and administrator
States assumed background and expertise of the reader, for each class of user.	No, the content is explained only from a developer point of view
Lists resources for further information.	No, list of resources is not provided
Is task-oriented.	Yes, it explains configuration to be done by the administrator
Consists of clear, step-by-step instructions Gives examples of what the user can see at each step e.g. screen shots or command-line	Yes, it also provides screenshots for each step of the installation and configuration processes
For problems and error messages, the symptoms and step-by-step solutions are provided.	No, there are no solutions provided except forum for people to discuss the issue
Does not use terms like "intuitive", "user friendly", "easy to use", "simple" or "obviously", unless as part of quotes from satisfied users	Yes
Further information is suitable for the level of the reader, for each class of user.	A list of resources is not provided

Capability for Installation

Web site has instructions for installing the software.	Yes, it shows step-by-step instructions with syntax that can be copy-pasted
Web site lists all third-party dependencies that are not bundled, along with web addresses, suitable versions, licences and whether these are mandatory or optional.	No, the only things that are listed is OS that support quick installation which is Ubuntu 16.04 LTS
All mandatory third-party dependencies are currently available.	No, the dependencies are only stated in the installation guideline
All optional third-party dependencies are currently available.	No, the optional dependencies are not shown nor explained even briefly
Tests are provided to verify the install has succeeded.	No test is provided for verification
When an archive (e.g. TAR.GZ or ZIP) is unpacked, it creates a single directory with the files within. It does not spread its contents all over the current directory.	Yes, the archive creates a single directory

Capability to Learn

A getting started guide is provided outlining a basic example of using the software.	Yes, it shows a guideline on how to set up the CTF
Instructions are provided for many basic use cases.	No use case is provided
Instructions are provided supporting all use cases. Reference guides are provided for all command-line, GUI and configuration options.	No use case is provided

APPENDIX B

USABILITY ASSESMENT FOR CTFd

Capability to Understand

High-level description of what/who the software is for is available.	No, but it explains in brief
High-level description of what the software does is available.	No, but it explains in brief as this platform is used to organize CTF competition or events
Architectural overview, with diagrams, is available.	No, it doesn't provide any architectural overview of this platform
Descriptions of intended use cases are available.	No, it doesn't explain any use case
Case studies of use are available.	No, it doesn't explain any case study

Documentation

Provides a high-level overview of the software.	No, it explains only in brief
Partitioned into sections for users, user-developers and developers (depending on the software).	No, there is no partitioning at all, as it is only explained from developer point of view
States assumed background and expertise of the reader, for each class of user.	No, the content is explained only from developer point of view
Lists resources for further information.	Yes, list of resources is provided
Is task-oriented.	Yes, it explains the configuration that will be done by the administrator
Consists of clear, step-by-step instructions Gives examples of what the user can see at each step e.g. screen shots or command-line	Yes, it also provides screenshots for each step of the installation and configuration processes
For problems and error messages, the symptoms and step-by-step solutions are provided.	No, there is no solution provided except for a forum for people to discuss the issue
Does not use terms like "intuitive", "user friendly", "easy to use", "simple" or "obviously", unless as part of quotes from satisfied users	Yes
Further information is suitable for the level of the reader, for each class of user.	Yes, it is suitable for general range of user

Capability for Installation

Web site has instructions for installing the software.	Yes, it shows step-by-step instructions with syntax that can be copy-pasted, but briefly
Web site lists all third-party dependencies that are not bundled, along with web addresses, suitable versions, licences and whether these are mandatory or optional.	Yes, it also provides link to install Docker, which is needed to start the installation process
All mandatory third-party dependencies are currently available.	Yes, the dependencies are available
All optional third-party dependencies are currently available.	No, the optional dependencies are not shown nor explained even briefly
Tests are provided to verify the install has succeeded.	No, no test had been provided for verification
When an archive (e.g. TAR.GZ or ZIP) is unpacked, it creates a single directory with the files within. It does not spread its contents all over the current directory.	Yes, the archive creates a single directory

Capability to Learn

A getting started guide is provided outlining a basic example of using the software.	Yes, it shows a guideline on how to set up the CTF
Instructions are provided for many basic use cases.	No use case is provided
Instructions are provided supporting all use cases. Reference guides are provided for all command-line, GUI and configuration options.	No use case is provided

APPENDIX C

USABILITY ASSESMENT FOR MELLIVORA

Capability to Understand

High-level description of what/who the software is for is available.	No, but it explains in brief
High-level description of what the software does is available.	No, but it explains in brief as this platform is used to organize CTF competition or events
Architectural overview, with diagrams, is available.	No, it doesn't provide any architectural overview of this platform
Descriptions of intended use cases are available.	No, it doesn't explain any use case
Case studies of use are available.	No, it doesn't explain any case study

Documentation

Provides a high-level overview of the software.	No, it explains only in brief
Partitioned into sections for users, user-developers and developers (depending on the software).	No, there is no partitioning at all as it is only explained from developer point of view
States assumed background and expertise of the reader, for each class of user.	No, the content is explained only from developer point of view
Lists resources for further information.	No, it only provides screenshot of the pages without any explanations
Is task-oriented.	Yes, it explains the configuration that will be done by the administrator
Consists of clear, step-by-step instructions Gives examples of what the user can see at each step e.g. screen shots or command-line	Yes, it provides step-by-step instructions for installation and also for configuration
For problems and error messages, the symptoms and step-by-step solutions are provided.	No, there is no solution provided except for a forum for people to discuss the issue
Does not use terms like "intuitive", "user friendly", "easy to use", "simple" or "obviously", unless as part of quotes from satisfied users	Yes, it mentions that it is fast but without any reference regarding the evaluation methodology
Further information is suitable for the level of the reader, for each class of user.	No information is provided except for screenshots of pages

Capability for Installation

Web site has instructions for installing the software.	Yes, it shows step-by-step instructions with syntax that can be copy-pasted, but briefly
Web site lists all third-party dependencies that are not bundled, along with web addresses, suitable versions, licences and whether these are mandatory or optional.	No, it only provides a list of third-party dependencies without any further details
All mandatory third-party dependencies are currently available.	Yes, the dependencies are available
All optional third-party dependencies are currently available.	No, the optional dependencies are not shown nor explained even briefly
Tests are provided to verify the install has succeeded.	No test has been provided for verification
When an archive (e.g. TAR.GZ or ZIP) is unpacked, it creates a single directory with the files within. It does not spread its contents all over the current directory.	Yes, the archive creates a single directory

Capability to Learn

A getting started guide is provided outlining a basic example of using the software.	Yes, it shows a guideline on how to set up the CTF
Instructions are provided for many basic use cases.	No use case is provided
Instructions are provided supporting all use cases. Reference guides are provided for all command-line, GUI and configuration options.	No use case is provided

APPENDIX D

USABILITY ASSESMENT FOR NIGHTSHADE

Capability to Understand

High-level description of what/who the software is for is available.	No, but it explains in brief
High-level description of what the software does is available.	No, but it explains in brief as this platform is used to organize CTF competition or events
Architectural overview, with diagrams, is available.	No, it doesn't provide any architectural overview of this platform
Descriptions of intended use cases are available.	No, it doesn't explain any use case
Case studies of use are available.	No, it doesn't explain any case study

Documentation

Provides a high-level overview of the software.	No, it explains only in brief
Partitioned into sections for users, user-developers and developers (depending on the software).	No, there is no partitioning at all as it is only explained from developer point of view
States assumed background and expertise of the reader, for each class of user.	No, the content is explained only from developer point of view
Lists resources for further information.	No, there is no further explanation
Is task-oriented.	Yes, it explains the configuration that will be done by the administrator
Consists of clear, step-by-step instructions Gives examples of what the user can see at each step e.g. screen shots or command-line	Yes, it provides step-by-step instructions for installation and also for configuration
For problems and error messages, the symptoms and step-by-step solutions are provided.	No, there is no solution provided except for a forum for people to discuss the issue
Does not use terms like "intuitive", "user friendly", "easy to use", "simple" or "obviously", unless as part of quotes from satisfied users	Yes
Further information is suitable for the level of the reader, for each class of user.	No information is provided except for screenshots of pages

Capability for Installation

Web site has instructions for installing the software.	Yes, it shows step-by-step instruction with syntax that can be copy-pasted, but briefly
Web site lists all third-party dependencies that are not bundled, along with web addresses, suitable versions, licences and whether these are mandatory or optional.	No, it does not provide third-party dependencies
All mandatory third-party dependencies are currently available.	No, the mandatory dependencies are not shown nor explained even briefly
All optional third-party dependencies are currently available.	No, the optional dependencies are not shown nor explained even briefly
Tests are provided to verify the install has succeeded.	No test is provided for verification
When an archive (e.g. TAR.GZ or ZIP) is unpacked, it creates a single directory with the files within. It does not spread its contents all over the current directory.	Yes, the archive creates a single directory

Capability to Learn

A getting started guide is provided outlining a basic example of using the software.	Yes, it shows a guideline on how to set up the CTF
Instructions are provided for many basic use cases.	No use case is provided
Instructions are provided supporting all use cases. Reference guides are provided for all command-line, GUI and configuration options.	No use case is provided