

The Forensic Challenger - A Digital Forensic E-Learning Platform

Øyvind Aaen Nordhaug



Master's Thesis Project Description
Master of Science in Information Security
5 ECTS
Department of Computer Science and Media Technology
Gjøvik University College, 2013

Avdeling for
informatikk og medieteknikk
Høgskolen i Gjøvik
Postboks 191
2802 Gjøvik

Department of Computer Science
and Media Technology
Gjøvik University College
Box 191
N-2802 Gjøvik
Norway

Abstract

E-learning is becoming an increasingly popular form of education. The ability to teach many students at once, and the students possibility to learn material at their own pace have been very successful. Knowing how digital forensics investigations are performed is becoming a necessity in many fields. The society's dependency on computers, and the massive increase in computing power and data storage, causes the society to require more expertise and efficient methods to handle digital forensics investigations. The use of e-learning can greatly benefit this situation, as more people can be aware of the correct procedures and key elements to digital forensics. The challenge is to efficiently teach these methods through an e-learning platform. By using the Hyper Interactive Presentation (HIP) platform, this thesis provided a proof of concept platform of an e-learning solution for digital forensics investigations, "The Forensic Challenger"(TFC), to improve teaching digital forensics. This thesis outlines the technical development of the platform. The results of the trial experiment with two groups of first year information security bachelor students. One group of 5 students used TFC platform, and the other group of 4 students used a simulation version of the existing platform used in the forensics lab. When we compared the groups in terms of motivation, and time spent, the TFC group indicated higher motivation and spent less time solving the task in the assignment.

Sammendrag

E-læring har blitt en stadig mer populær form for utdanning. Muligheten til å lære flere studenter på en gang, og elevene sin mulighet til å lære materialet i sitt eget tempo har vært svært vellykket. Å vite hvordan digitale etterforskninger er utført blir stadig viktigere i mange felt. Samfunnets avhengighet av datamaskiner, og den enorme økningen i datakraft og lagring av data, fører til at samfunnet trenger mer kompetanse og effektive metoder for å håndtere digitale etterforskninger. Bruken av e-læring kan være veldig nyttig i denne situasjonen, fordi flere personer kan lære seg de riktige prosedyrene og viktige elementer i digital etterforskning. Utfordringen er å effektivt lære disse metodene gjennom en e-læringsplattform. Ved hjelp av Hyper Interactive Presentasjon (HIP) platformen, gir denne oppgaven et proof of concept rammeverk for en e-læring løsning for digital etterforskning, "The Forensic Challenger" (TFC), for å forbedre opplæringen til å utføre digital etterforskning. Denne oppgaven skisserer den tekniske utviklingen av platformen. Resultatene av prøve eksperimentet med to grupper av første-år informasjons-sikkerhet bachelor studenter. En gruppe med 5 studenter brukte TFC platformen, og den andre gruppen med 4 studenter brukte en simulert versjon av den eksisterende platformen brukt i etterforsknings laboratoriet. Når vi sammenlignet de to gruppene ut i fra motivasjon og tid brukt, TFC gruppen indikerte høyere motivasjon og brukte mindre tid på å løse oppgavene.

Acknowledgements

I would like to thank my supervisor Prof. Stewart Kowalski, for providing guidance, assistance and ideas throughout the project. Secondly i would like to thank Ali Shariq Imran for technical assistance and feedback.

Additionally i would like to thank NISLab forensic lab for hosting the project on their server, and Andrii Shalaginov for excellent work with administrating the server.

Finally i would like to thank André Nordbø for providing the video recordings.

Contents

| | |
|--|------------|
| Sammendrag | iii |
| Acknowledgements | v |
| Contents | vii |
| List of Figures | ix |
| 1 Introduction | 1 |
| 1.1 Topic | 1 |
| 1.2 Keywords | 1 |
| 1.3 Problem description | 1 |
| 1.4 Justification, motivation and benefits | 2 |
| 1.5 Research questions | 2 |
| 1.6 Planned contributions | 2 |
| 2 State of The Art | 3 |
| 2.1 Related Work | 3 |
| 2.1.1 E-Learning | 3 |
| 2.1.2 Digital Forensics E-Learning | 3 |
| 2.2 Theoretical Background | 5 |
| 2.2.1 Learning Styles | 5 |
| 2.2.2 Motivation in e-learning | 6 |
| 2.2.3 Learning With Games | 7 |
| 2.3 Technical Background | 9 |
| 2.3.1 Hyper Interactive Presentation (HIP) | 9 |
| 2.3.2 Django | 10 |
| 2.3.3 AIML | 12 |
| 2.3.4 PyAIML | 13 |
| 3 Choice of methods | 15 |
| 3.1 Action Research | 15 |
| 3.2 Action Research Development | 16 |
| 3.3 Experimental Design | 16 |
| 3.3.1 Target Group | 16 |
| 4 Implementation And Results | 19 |
| 4.1 Scope and Limitations | 19 |
| 4.2 Findings - The Forensic Challenger | 19 |
| 4.2.1 Overview | 19 |
| 4.2.2 Development | 20 |
| 4.3 Experimental Trial - Fronter vs TFC | 28 |
| 4.3.1 Questionnaire Design | 29 |
| 4.3.2 Results of Survey | 30 |
| 5 Conclusions | 39 |
| 5.1 Research Questions | 39 |
| 5.1.1 Can digital forensics investigation lab sessions be implemented on a web based e-learning platform? | 39 |

| | | |
|-------|---|-----------|
| 5.1.2 | Does the e-learning platforms interactive elements help to improve motivation to learn digital forensics investigations | 39 |
| 5.2 | Future Work | 40 |
| | Bibliography | 43 |
| A | Questionnaires | 47 |
| B | Specialization Course II - Digital Forensics Game | 49 |

List of Figures

| | | |
|----|---|----|
| 1 | Classroom vs E-learning[3] | 3 |
| 2 | SANS OnDemand[6] | 4 |
| 3 | Django App Files | 11 |
| 4 | Action Research Model[37] | 15 |
| 5 | The Forensic Challenger Web Interface | 20 |
| 6 | The Forensic Challenger Video Frame | 21 |
| 7 | The Forensic Challenger Presentation Frame | 22 |
| 8 | Content Structure for TFC | 24 |
| 9 | End of Challenge Score | 25 |
| 10 | The Forensic Challenger Challenge Frame | 25 |
| 11 | Questions Layout | 26 |
| 12 | The Forensic Challenger Wiki Frame | 27 |
| 13 | The Forensic Challenger Chatbot Frame | 28 |
| 14 | Question 1 Distribution with likert scale for group A | 30 |
| 15 | Question 1 Distribution with likert scale for group B | 30 |
| 16 | Question 2 Group A | 31 |
| 17 | Question 2 Group B | 31 |
| 18 | Question 3 Group A | 32 |
| 19 | Question 3 Group B | 32 |
| 20 | Question 4 Group A | 33 |
| 21 | Question 4 Group B | 33 |
| 22 | Question 5 Distribution with likert scale for group B | 34 |
| 23 | Question 8 Distribution with likert scale for group B | 34 |
| 24 | Question 5 Distribution with likert scale for group A | 35 |
| 25 | Question 6 Distribution with likert scale for group B | 35 |
| 26 | Question 6 Distribution with likert scale for group A | 36 |
| 27 | Question 9 Distribution with likert scale for group B | 36 |

1 Introduction

The purpose of this chapter is to give an introduction to the subject and challenges, as well as justification and motivation. This chapter also proposes research questions to structure the thesis, and a discussion of the planned contribution.

1.1 Topic

Data is stored on not only computer hard drives or CD's, but also portable data and storage devices, such as smart phones. The information stored on these devices is a crucial source of evidence in civil, corporate or criminal investigations.

The need for proper e-learning platforms is of great importance, as the challenges in the digital forensics field is increasing in complexity. E-learning grants the opportunity to teach many people that comes from different fields, at the same time. As the complexity of the topics increases, e-learning provides the ability to learn the material at your own pace. One person may understand everything on the first attempt, and others can go back and look up a topic of interest.

The way a person learns is either through visual, auditory or kinesthetic/tactile (Hands-on experience) learning, or a combination of them. Learning styles refers to the view that different people learn information in different ways[1].

The Hyper Interactive Presentation (HIP) platform, provides a web interface that is constructed out of several modules that interact with each other. This project created a digital forensics e-learning platform, called "The Forensic Challenger"(TFC). TFC uses game elements to further motivate the student to learn digital forensics investigations. Throughout the game it is possible to interact with the HIP platform modules to find information about a particular topic. By combining the HIP platforms auditory and visual learning styles with the kinesthetic learning style the game will offer, we hope to reach a much broader audience and get more effect from the learning.

1.2 Keywords

Digital Forensics, E-Learning

1.3 Problem description

Performing a digital forensic investigation have become more complex as society is more dependant on computers. This creates a gap between the society's capability to investigate the crime, and the criminals likelihood to be caught and convicted. To fill this gap we need more efficient digital forensics investigations.

These investigations are conducted by following various steps to ensure that the process is properly handled. Proper collection and examination of evidence is critical to avoid corruption and to preserve the evidence. Failing to do this, the gathered data may lose its admissibility as evidence. Understanding what must be done, and how it should be done is essential in a digital forensics investigation. This would also be useful for not only law enforcement, but judges, lawyers and prosecutors as well as corporations.

Many digital forensics courses are taught online, but society also struggle to make these online courses more efficient. The success of using online courses is dependant on many factors, including learning styles[2]

The use of video recording to capture lectures and lab sessions performed by professionals in the field, can be a great help for students. However, this is only true if the material is presented in a way that makes it efficient and accurate to use.

1.4 Justification, motivation and benefits

Digital forensics integrates the fields of computer science and law to investigate crime. For any digital evidence to be used in court, investigators must follow a proper set of procedures when collecting and analysing data from computer systems. Many of these laws were written before the era of computer forensics, and are often outdated. The inability of law to keep pace with technological advancements may in the end limit the use of forensics evidence in court.

The increased complexity of the communication and networking infrastructure is making investigation of cybercrimes difficult, clues of illegal activities are often buried in large volumes of data that needs to be sifted through in order to detect crimes and collect evidence. Technologies like encryption and anonymity may be abused by helping criminals hide their actions. Ultimately, the role of technology in computer forensics may not reach its full potential due to legal boundaries and potential malicious intent.

Because of this gap between the criminals and the law enforcement, not only renewal of laws is necessary, but more efficient digital forensics investigations are also needed. The way we chose to approach this problem is by using interactive elements with the HIP platform to provide a learning method where knowledge of how to conduct a digital forensics investigation is in focus.

1.5 Research questions

1. Can digital forensics investigation lab sessions be implemented on a web based e-learning platform?
2. Does the e-learning platforms interactive elements help to improve motivation to learn digital forensics investigations?

1.6 Planned contributions

In this project we are developing a digital forensics investigation e-learning platform, called The Forensic Challenger (TFC), using the Hyper Interactive Presentation (HIP) platform as a test platform. By using the HIP platform we combine the use of visual, auditory and kinesthetic learning. This investigation will go through a fictional scenario and the player have to answer a series of questions regarding this scenario. The video recording provided is captured during a lab session with Jeffery David Hamm for Digital Forensics II, conducted at Gjøvik University College in 2013. This satisfies on a very basic level the kinesthetic learning, as the user get to experience recordings of live examples. Furthermore during this challenge, the ability to access a specifically crafted Wiki as well as presentation of the topics in video. The expected result of this method is increased learning performance and motivation for the students. Furthermore, we conducted an experiment to measure the effectiveness of the forensic challenger platform, opposed to the traditional way of capturing and sharing a video as a raw and unedited recording.

2 State of The Art

This chapter begins with the related literature and current state of research on digital forensics e-learning, then continues with a background overview of the technologies used; Hyper Interactive Presentation (HIP), Django, AIML and PyAIML.

2.1 Related Work

2.1.1 E-Learning

In an article by Zhang et.al 2004[3], the authors claim that "In an E-Learning environment that emphasizes learner-centered activity and system interactivity, remote learners can outperform traditional classroom students". In this article they create a Virtual Mentor (VM) which consists of the following principles:

1. Multimedia Integration - Video with corresponding presentation slides
2. Just-in-time knowledge acquisition - Access any time, anywhere
3. Interactivity - Interact with the VM
4. Self-Directivity - Student choose personal learning strategies
5. Flexibility - Give the student control over the learning process
6. Intelligence - Monitor each students learning progress and give personal tutoring

This solution, much like the HIP platform, gives the students a platform which offers more control over the learning.

| | Traditional Classroom Learning | E-Learning |
|----------------------|---|---|
| Advantages | <ul style="list-style-type: none"> • Immediate feedback • Being familiar to both instructors and students • Motivating students • Cultivation of a social community | <ul style="list-style-type: none"> • Learner-centered and self-paced • Time and location flexibility • Cost-effective for learners • Potentially available to global audience • Unlimited access to knowledge • Archival capability for knowledge reuse and sharing |
| Disadvantages | <ul style="list-style-type: none"> • Instructor-centered • Time and location constraints • More expensive to deliver | <ul style="list-style-type: none"> • Lack of immediate feedback in asynchronous e-learning • Increased preparation time for the instructor • Not comfortable to some people • Potentially more frustration, anxiety, and confusion |

Figure 1: Classroom vs E-learning[3]

Figure 1 shows a comparison between classroom learning and e-learning

2.1.2 Digital Forensics E-Learning

There exists several various solutions for e-learning on the web. Many of these offer courses in digital forensics. The Forensics Wiki [4] provides links to several distance learning solutions for digital forensics, as well as commercial and non commercial train-

ing programmes and tool training.

In a paper published by Yin Pan et.al. "Game-based Forensics Course For First Year Students"[5] The authors have developed a forensics course using the game-based learning (GLB) approach. This system has direct access to forensics tools, and evidence from a suspects machine. This digital forensics game focus on visualizations to illustrate fundamental concepts in computer forensics, as well as interactive lab-based investigation modules to allow the students to practice in gathering, preserving, analysing and reporting digital evidence. This thesis will differ from this solution by being entirely web based. Additionally, by using the HIP platform we gave the users the ability of using the learning style that is most suitable for them.

The SANS Institute offers a multitude of opportunities for security training in various fields, including digital forensics. SANS provides different types of e-learning platforms, the most relevant one is the OnDemand platform[6]. SANS offers 3 different courses in forensics through this platform, these include a course in forensics investigations on Windows, one about malware reverse engineering, and a advanced course in forensics analysis and incident response. Due to the pricing on these courses, we have not been able to test the full version of OnDemand platform at this time.

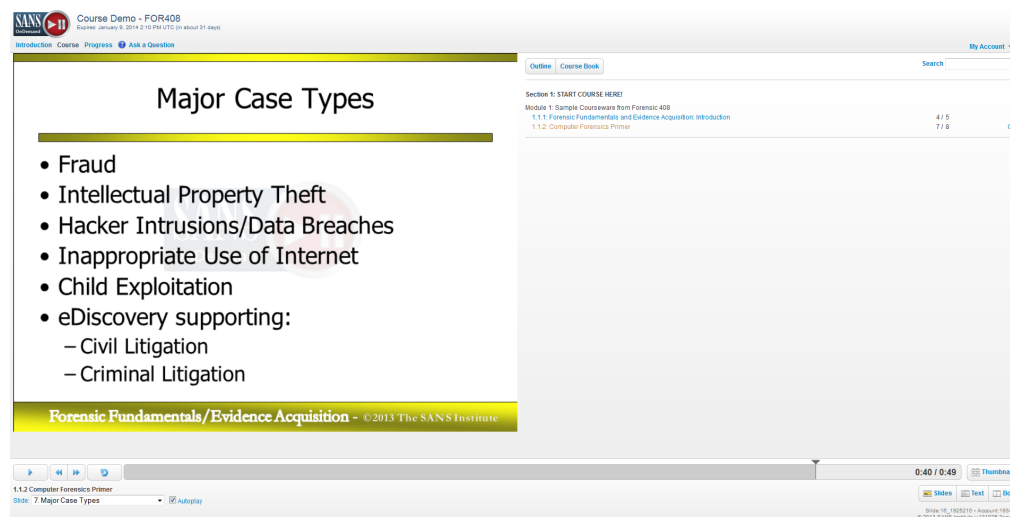


Figure 2: SANS OnDemand[6]

A demo of the courses are available for free. Figure 2 shows the website of the demo. SANS OnDemand is divided into a few modules. The course material is presented with presentation slides and audio from the course lecturer. In addition to these presentations, you have the option of referring to the course book during the presentation, get a progress report and ask questions. You can view the course material as much as you want. To pass the SANS course, you have to participate in several quizzes, and pass the quizzes with a score of 80 percent or higher to be able to access the certificate of completion.

Computer Forensics Training Center Online (CFTCO)[7] offers a distance course in digital forensics. The CFTCO is an authorized training partner of the International Society of Forensics Computer Examiners (ISFCE)[8], which provides the Certified Computer Ex-

aminer (CCE) certification. This course is called CCE BootCamp, and focuses more on the core forensics procedures, and do not cover the forensics tools at all. Their focus is not only on forensics recovery techniques, but what you have to do in order to make sure that the evidence located will be admissible in court. On their webpage, the CCE BootCamp teaches you how to conduct thorough examinations, identifying where and how data is stored, recovering and interpreting data and drawing appropriate conclusions based on the data. The CFTCO course is taught by using reading material, lab exercises, and interaction with an instructor.

A paper published by Gary C. Kessler called "Online Education in Computer and Digital Forensics: A Case Study"[2] describes some of the course design aspects of teaching computer forensics in an online environment. In this paper, the author is writing about Champlain College's online course in digital forensics using the Learning Management System (LMS), which is an asynchronous, virtual classroom. The LMS provides communication tools, allowing online students to interact with each other on many levels. This solution is e-learning combined with traditional classroom learning, where interaction and discussion with your peers is helping to create a richer learning environment. In 2006 the Champlain College performed a study to determine if the learning objectives were met equally well in online and on-campus courses[2]. With no significant difference between the two delivery modes, the average grade of the online sections were slightly higher.

In May 2011, three students at the College of Information Technology of Zayed University, launched a digital forensics e-learning platform called "Jenaei". In a news article on gulfnews.com [9]. Jenaei, translated to "forensics" in english, is a project that seeks to educate lawyers, judges and prosecutors in the United Arab Emirates, to give them a better understanding of the technical aspects that cyber crime investigations involves. The author have not been able to find a version of this project to analyse.

Learning the tools used in a digital forensics investigations are a time-consuming process. Many tool providers have training in using their tools. AccessData[10], the developer of Forensic Toolkit (FTK) provides a Learning Management System (LMS) for their products. The LMS is a online video-based delivery system for all AccessData's training course material. The program can be organized by role, skill level or by course content.

2.2 Theoretical Background

This section will give an overview of the theoretical background of the thesis

2.2.1 Learning Styles

There are different ways to transfer knowledge to humans. Every learner has his/her character or learning style to receive that knowledge[1]. The student learning process has two basic factors that affect the range of its success: user learning style and the way to present the knowledge to learners. To support the learning process, we need to deliver the material adhering to users' preferences based on the various learning styles[11][12]

The student learning style is affected by several factors, like sex, age, culture, com-

munity, subject and environment. In general, learning style should have at least four basic dimensions[13]: cognitive, affective, physiological and psychological. Students can use the cognitive sense such as seeing, focusing, linking to another concept and remembering activities to aid the learning process. Affective dimension is affected by attributes associated with a personality such as emotion, motivation, incentive, curiosity, boredom, concern and frustration. Physiological dimensions is linked with visual, auditory and kinesthetic sense, which students normally uses in the learning process.[12]

The VARK model is one of the most utilized learning style models. VARK is an acronym for Visual(V), Aural/Auditory (A), Read/write (R), Kinesthetic (K). Below is an overview of the various styles[12][14].

1. The visual preference includes the depiction of information in maps, spider diagrams, charts, graphs, flow charts etc. that people use to represent what could have been presented in words. It does not include still pictures or photographs of reality, movies, videos or powerpoint.
2. The auditory mode describes a preference for information that is heard or spoken. Learners who have this as their main preference report that they learn best from lectures, group-discussion, speaking, and talking things through.
3. Read/write is a preference for information from written or printed text materials. Many teachers and students have a strong preference for this mode, and they learn best from text documents, textbooks, lecture notes, lists and handouts etc.
4. Kinesthetic learners prefer to feel and live the learning knowledge, they learn by doing. People in this group learn best from practical sessions, field trips, experiments, role playing or simulation, and other similar exercises.

2.2.2 Motivation in e-learning

In a featured article in ACM's e-learning magazine[15], Matt Guyan says that "motivation has been and continues to be a widely studied area across many of life's domains. Motivation is said to be the energizing force that initiates and sustains behaviour and ultimately produce results"

In this article he talks about the different forms of motivation, and specifies that motivation should not be generated by the educators, but to create conditions that facilitate the internalization of motivation from within the learners.

Motivation can be divided into two types, intrinsic (internal) and extrinsic (external). Intrinsic motivation refers to motivation that is driven by an interest or enjoyment in the task itself. Extrinsic motivation refers to the performance of an activity in order to attain an outcome, whether or not that activity is also intrinsically motivated.[16]

The various forms of motivation includes self-determination theory (SDT) by Deci and Ryan. This theory focuses on control versus autonomy as the differentiating factor between various forms of extrinsic motivation.

Organismic integration theory (OIT), which is a sub-theory of SDT. This theory states extrinsic motivation is not a single construct; it exists in four distinct forms according to the extent to which the motivation for behaviour originates from one's self.

External regulation is the most controlling form of extrinsic motivation, which is performing a task to receive an external reward or avoid a punishment.

A less controlling form of extrinsic motivation is introjected regulation, which is per-

forming a task to avoid feelings of guilt or to affect one's self-esteem or sense of self-worth.

Finally, integrated regulation is the most autonomous form of extrinsic motivation. In this case, motivation is connected with a person's values or beliefs.

Guyana lists a number of strategies that can be used to improve motivation.

1. Give learners some level of control as they work through the module or course
2. Provide regular, meaningful feedback throughout the learning experience
3. Incorporate social elements.
4. Provide opportunities for collaboration between learners
5. Keep the stakes low and allow learners to practice
6. Allow learners to make meaningful choices and pursue challenging goals.

2.2.3 Learning With Games

Rick Raymer describes the term "gamification" as the use of gameplay mechanics in non-game applications. "Games are generally structured so that players have various "layers" of goals. That is, they have the long-term goal of completing the game, the medium-term goal of completing the levels in the game, and the short-term goal of completing the missions in the levels. (Sometimes these missions are even broken up further into additional tasks.) Generally, the requirements of each goal "layer" in a game get increasingly harder as you move from short-term to long-term goals. That is, the final challenge in a game (sometimes called the "boss battles") will always be harder than the short-term missions. This allows players in games to learn and practice skills, prior to having to demonstrate mastery of those skills in the most challenging parts of the game." [17] When utilizing this principle when designing e-learning material, the designers should break up the products into short-term, medium-term and long term goals. For example, to complete an e-learning course, the users must complete several modules. To complete a module, several topics must be completed. In order to complete a topic, several objectives must be finished, and each objective requires several goals to be completed. Structuring the learning process this way, allows users to learn new skills incrementally, and then practice those skills before demonstrating mastery of those skills in assessment exercises.[17]

It is important that a user doesn't feel lost or confused during an e-learning exercise. There should be mechanisms in place so the user knows what they need to do next, or what options are available during any given moment. There should be links provided to essential information the user need to progress. Measuring progress is also important, this progress should be measured at the short-term, medium-term and long term goals. For example you have 100 percent completion of one module, this corresponds to 20 percent completion of the entire course. Showing the progress only when progress is made can feel more like a reward for the user. Showing progress can be done in several ways. Utilizing a progress bar is simple and effective. Another approach can be to use character upgrades. Character upgrades gives the opportunity to reward several different accomplishments, as well as rewards for making an effort.

Peer motivation is an effective approach to enhance motivation of the users, there are several ways you can use the learner's peers to motivate. Using high scores or achieve-

ments gives people the ability to use them as bragging rights and will enhance the effectiveness of the e-learning.[17]

So what makes learning games successful? In a paper by Marc Prensky [18], he talks about the element of "fun", and calls it "the great motivator". He further discuss the relationship between fun and learning and thinks the principal roles of fun in the learning process is to create relaxation and motivation. Relaxation enables the learner to take things in more easily, and motivation enables them to put forth effort without resentment.

There are a number of ways to discuss the forms and roles which online learning materials can take. Oliver and Herrington suggests a possible way of classifying and comparing online learning materials[19]

| Form | Description of Materials |
|-----------------------|---|
| Information Access | Convey information alone to the learner, for example, a course syllabus, a calendar, assignment descriptions, lecture notes, workshop descriptions etc. |
| Interactive Learning | Involve instructional elements that engage the learner, encourage reflection and decision making and provide feedback in response to learner actions |
| Networked Learning | Provide a means for the organisation, communication and exchange of ideas and information among learners and teachers and other parties in the learning process |
| Materials Development | When the online setting is used as a means for learners to create and publish materials. For example, learners publishing web pages, stories etc. |

Table 1: Forms of online learning materials[19]

Learning games fits under the Interactive Learning category, and are characterised by forms of learner control and active engagement where learners can take and make decisions and learn through the consequences of their actions. Some of the elements of interactive learning is[19]:

1. Activities which require students to search and review documents and links in a structured and organised fashion to discover particular information.
2. Questions/activities that are posed which cause students to reflect and to select from among various outcomes, with feedback particular to the students choices being provided and able to influence subsequent directions and activities presented.
3. Through the use of forms, where students can select options to effect particular processing outcomes, for example, database searching, selection processes, simulation control.
4. Through the use of program modules which enables the learners input to be processed and acted on, for example, shockwave movies, virtual reality environments etc.

Interactive elements supports learning in online settings through means such as the possibility of immediate feedback to responses given to tasks and problems and a capacity to respond to the user in ways which recognise the discrete needs of individuals. Additionally you have means of presenting tutorial type activities, which support knowledge acquisition through such tasks as drill and practice, as well as simulations of real like events, which learners can learn to manage and control in a safe and structured

setting[19].

The popularity of the Internet has led to several advancements in educational technology. Still most e-learning solutions utilizes traditional instruction contents, which are placed on the Internet without modifications that take advantage of the full potential of either the Internet or online technology[20][21].

Cognitive Learning Outcomes

Declarative Knowledge

Declarative knowledge refers to knowledge of the facts and data required for task performance. For this type of learning outcome, the learner is typically required to reproduce or recognize some item of information. White published a study in 1984 that[22] showed how students who played a computer game focusing on Newtonian principles were able to answer questions on force and motion problems more accurately than those who did not play the game.[23]

Procedural Knowledge

Procedural knowledge refers to knowledge about how to perform a task. This learning outcome is dependant on a demonstration of the ability to apply knowledge, general rules or skills to a specific case. Whitehall and McDonald performed a study on students who used a variable-payoff electronics game[24]. The students who used the game, achieved higher score on electronics troubleshooting tasks than students who received standard drill and practice.[23]

Strategic Knowledge

Strategic knowledge is the result of applying learned principles to different contexts, or deriving new principles for general or novel situations. This implies the development and application of cognitive strategies and understanding when and why principles apply. Wood and Stewart did a study about using a computer game to improve practical reasoning skills of students, this led to improvements in critical thinking[25][23].

Affective Learning Outcomes

Affective learning outcomes include feelings of self-efficacy, confidence, attitudes, preferences and dispositions. Affective reactions may be viewed as a specific type of learning outcome to the extent that attitude change is a training objective of an instructional program. A study done by Tomas, Cahill and Santilli[26] reported success in using an adventure game format to enhance students' confidence regarding safe sex negotiations[23]

2.3 Technical Background

This section will give an overview of the technology used throughout the thesis. This includes the Hyper Interactive Presentation (HIP) platform, Django and AIML/PyAIML

2.3.1 Hyper Interactive Presentation (HIP)

The work with HIP functionality in this thesis, is based on the work of Ali Shariq Imran for the Ericsson Response Training Programme. Ali's HIP platform as it was when the author received it, did not handle dynamic content. Each presentation would need their own JavaScript file, which included hard-coded timestamps for synchronization, as well as image names.

The core functionality comes from the usage of the HTML5 property "currentTime". Where you can get the current time value the video is at, or you can pass a value to set

the time.

HIP is an e-learning platform that provides technology-rich pedagogical media for continuous education and connected learning. It combines four media modalities to suit a student learning styles. These include text documents such as wiki pages and pdf documents, PowerPoint presentations, lecture videos, and an interactive dialogue (intelligent pedagogical agent) along with navigational links, tagged keywords, and frequently asked questions(FAQ). HIP supports nano-learning by creating smaller chunks of video learning objects (VLOs), and hyperlinking similar LOs across different media.[27]

HIP is comprised of four main components. The components are designed to present their knowledge in a number of ways, utilizing all the available media modalities. These include hyperlink-video, PowerPoint presentation slides, documents, and a pedagogical agent. These different media modalities in HIP are used to map the VARK model, to support multiple learning styles.

2.3.2 Django

Django is a high-level Python web framework, that offers security features to protect against Cross Site Scripting attacks, Cross site request forgery attacks, SQL injection attacks, clickjacking and more. A complete description of the Django coding can be found in appendix B. In Django you define data models entirely in Python, and it gives you a dynamic database-access API. When starting a new Django project, by issuing the command "django-admin.py startproject mysite", it sets up the basic structure of the web site for you. That could look something like below[28].

```
mysite/
  manage.py
  mysite/
    __init__.py
    settings.py
    urls.py
    wsgi.py
```

The manage.py file is a command line utility that lets you interact with the Django project in different ways. In addition to provide a help function, it gives you the opportunity to create new "Apps", run a shell to your database, run the server on a specified port and much more.

In figure 1, we have created an app called presentations, for use in the HIP platform, django have created the necessary files, and after adding the new app in the list of apps to load, and add some url expressions, we can start populating the app with javascript, html files, views and modules to create interaction on the site.

JavaScript

Since the javascript files used in the HIP platform have hardcoded inputs of images and timestamps etc. We would need to make significant changes to the script, and make it include all these things from the database, this is still in development. Currently the site only loads the original javascript files from the HIP platform on each of the presentations provided. To call a javascript file in a html document, you must modify the code by adding

```
{% load staticfiles %}
```

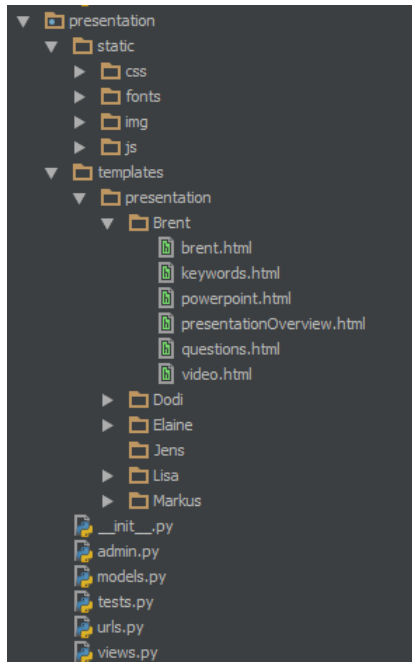


Figure 3: Django App Files

And then refer to the static file you wish to use

```
<link rel="javascript" type="text/javascript"
href="{% static 'js/script.js' %}">
```

This fetches the javascript file located in the static directory. This directory have to be specified in the settings.py file by adding the filepath where the directory is located. In this case its in the project directory under the folder "static"

```
STATIC_ROOT = os.path.join(PROJECT_DIR, 'static')
STATIC_URL = '/static/'
```

Templates

Django have a template engine that uses a mini-language for defining the user-facing layer of the application. Using templates offers the possibility of separating application and presentation logic. "A Django template is a string of text that is intended to separate the presentation of a document from its data. A template defines placeholders and various bits of basic logic (template tags) that regulate how the document should be displayed. Usually, templates are used for producing HTML, but Django templates are equally capable of generating any text-based format".[29]

The use of templates is particularly useful for pages that is changing a lot, but still requires some elements to be present at every page. For example a menu bar. The challenge frame in TFC uses templates, we have defined a template for the menu, and every sub-page of the challenge will inherit this menu.

First we created a block

```
{% block menubar %}
...
```

```
...  
{% endblock %}
```

Then insert the HTML code for the menu into this block. We also have a separate block for content, where we define with CSS the position and appearance. To use the template, the sub-page needs to have the following code at the start.

```
{% extends "base.html" %}
```

This command links the sub-page to the base page, where the HTML code is located. Only the base page need to have a full HTML code structure. If nothing is specified in the sub-page, the HTML code from the base page will appear on the sub-page. This can be changed by adding

```
{% blockname %}  
...  
...  
{% endblock %}
```

In the sub-page as well. The template engine is really helpful for developing dynamic content, and saves the developer a lot of time and reduces redundancy in code. Although the developer doesn't have to write the code, Django will generate the full HTML code before displaying the page.

2.3.3 AIML

AIML consists of data objects called AIML objects. These objects are made up of units called topics and categories. Topics is an optional element, and has a name attribute and a set of categories related to that topic. Categories are the basic unit of knowledge in AIML. Each category is a rule for matching an input and converting to an output, and consists of a pattern, which matches against the user input, and a template, which is used in generating the chatbots answer. The AIML format is as shown below.[30]

```
<aiml version ="1.0">  
<topic name="the topic">  
<category>  
<pattern>PATTERN</pattern>  
<that>THAT</that>  
<template>Template</template>  
</category>  
..  
..  
</topic>  
</aiml>
```

The tags used in the example is explained below. In addition to these tags, there is the <this>, <random>, <star> and <think> tags

1. The <aiml> tag is the parent tag for the entire AIML document.
2. The <topic> tag is used if there has been a topic <set> matching the query.
3. The <category> tag should contain the "matching" part of the expression, and the "returning" part of the expression
4. The <pattern> tag is the matching part, and specifies which pattern to match.

5. The <that>tag is used when the current pattern is depending on an earlier output.
6. The <template>tag specifies how the bot responds when the pattern within the category is matched.

Categories

AIML have three types of categories, atomic categories, default categories and recursive categories. These three types overlap, because atomic and default refer to the <pattern> and recursive refers to a property of the <template>.

Atomic

Atomic categories are those with atomic pattern, for example the pattern contains no wild card "*" or "_" symbol. Atomic categories are the easiest categories to add in AIML.[30]

```
<category>
<pattern>What is a hard disk drive?</pattern>
<template><set_it>hard disk drive</set_it>Hard disk drives (HDD)
  is a data storage device for digital information
</template>
</category>
```

The category above matches the client input of "what is a hard disk drive?" and sets the IT variable to the value of "hard disk drive" and sends the response in the template

Default

The default category derives from the fact that its pattern uses a wildcard "*" or "_". These default responses are often called "pickup lines" because they generally consists of leading questions designed to focus the client on known topics. The more common default categories have patterns combining a few words and a wild card[31], for example the category:

```
<category>
<pattern>I NEED HELP *</pattern>
<template>Can you ask for help in the form of a question?</template>
</category>
```

Recursive

Recursive categories are those that "map" inputs to other inputs, either to simplify the language or to identify synonymous patterns. Many synonymous inputs have the same response. This is accomplished with the recursive <srai>tag. For example the input "GOODBYE" have dozens of synonyms: "BYE", "BYE BYE", "CYA", "GOOD BYE", and so on. To map these inputs to the same output for GOODBYE, we use categories like[31]:

```
<category>
<pattern>BYE BYE</pattern>
<template><srai>GOODBYE</srai></template>
</category>
```

2.3.4 PyAIML

PyAIML is an interpreter for AIML, the Artificial Intelligence Markup Language, implemented as a standard python package. Using the PyAIML package is pretty simple[32]:

1. First you have to import the AIML package

```
import aiml
```

2. Then you have to instantiate a kernel object, the kernel is the only class you need to use when dealing with PyAIML

```
kernel = aiml.Kernel()
```

3. The kernel is now ready to respond to user input, so the next step is to load the AIML files, for this we will use the kernels learn() method

```
kernel.learn("std-startup.xml")
```

4. In this example, we have loaded the startup file from the standard AIML set, it defines a single AIML pattern, "LOAD AIML B", which causes the rest of the set to load. To trigger this process, we pass "load aiml b" as input to the kernel.

```
kernel.respond("load aiml b")
```

5. The respond method() returns a string containing the Kernel's response to the input. We are now ready to start responding intelligently to user input.

```
while True: print kernel.respond(raw_input("> "))
```

6. To reduce startup time, we can use the kernels loadBrain() and saveBrain() methods, these functions lets you dump the contents of the bot's brain to a file. So next time you will only have to use the loadBrain() function.

3 Choice of methods

This chapter discusses the methodology used to guide and analyse the research questions and the project in its entirety. It will discuss the methods used and why these methods are appropriate to answer the questions, as well as the expected result.

3.1 Action Research

This thesis used action research[33] as the main approach to research the problem, then used an experimental trial of the platform to validate the results. Action research is one of the contemporary approaches to educational research and creation of new knowledge[34][35][36]. Action research is based on the Plan -> Act -> Observe -> Reflect approach as shown in Figure 4



Figure 4: Action Research Model[37]

Action research in education has gained increasing attention in the past 20 years. It is viewed as a practical yet systematic research method that enables teachers to investigate their own teaching and their students' learning. There are a few different types of action research, in this thesis we will use the "practical action research" type, which is intended to address a specific problem within a classroom, school or other communities. The primary purpose of practical action research is to improve practice in the short term and to inform of larger issues.

The methodology of action research is a self-reflective process where one is constantly evaluating what one is doing, evaluating solutions and self evaluation with a goal to improve the quality through taking time to reflect on what is happening and change it.

3.2 Action Research Development

At Gjøvik University College, in the Digital Forensics Courses, there is a lab session spanning approximately 4 days, where the basics of how to conduct a digital forensics investigation is in focus. These lab sessions can be improved by porting it to an e-learning platform where the students can learn the topics in their own pace. The author have participated in these lab sessions, and have knowledge of which topics and methods should be focused.

To address the first research question, "Can digital forensics investigations lab sessions be implemented on a web based e-learning platform?".

Action research is appropriate to find a solution to this question. The action research approach gives the ability to change or improve something that is a problem area. The lab sessions in digital forensics investigations is a valuable opportunity for students to get insight in how investigations are conducted. The topics covered and tools explained might be too much to learn in just a few days.

To improve the learning experience, we will create an e-learning platform, and simulate the lab session by using video/audio recording and the presentation slides from the lab session.

Through the development, the author is solely responsible to implement appropriate features for the e-learning platform. These features were analysed and selected throughout the development, based on what the author experienced in the course, and ideas found in previous literature reviews.

To interpret the findings from the action research approach, an experimental trial was conducted to compare the difference between using the TFC platform and the current solution available to students at GUC, Fronter.

The second research question "Does the e-learning platforms interactive elements help to improve motivation to learn digital forensics investigations?".

To answer this question, the author used the action research approach, the author used his own knowledge and experience gained from developing the e-learning platform, to see what is working and what is not. The results derived from the action research will be interpreted by the experiment proposed in question 1.

3.3 Experimental Design

We created an experimental trial to interpret the findings from the action research process. During the action research process, we have made several choices when it comes to design and technical implementation. This experiment will highlight the difference between using The Forensic Challenger platform and the learning platform Fronter. After the experiment, the participants answer a survey which will hopefully validate our findings from the action research.

3.3.1 Target Group

The target group of the survey are first year bachelor students in Information Security. It is assumed that there are participants within the same age group and similar educational background. The participants were introduced to the experiment by the author, with permission from their class supervisor, and they could choose to participate or not. The participants were asked to provide their email address if they were interested in joining.

During the presentation of the experiment, 16 people were present in class, 14 of those agreed to participate in the experiment.

4 Implementation And Results

In the early planning, TFC was supposed to be a standalone program which utilized the original HIP platform through an integrated browser plug-in. After some estimations into how much time it would take to build such a system, we decided that it would be infeasible to finish it in four months.

Other options were discussed, but we decided to instead build a web interface using the Django framework as back-end running on an Apache web server. Using a framework like Django saves a great deal of time when developing web services, this will be discussed more thoroughly in the development section. The HIP functionality is based on work done by Ali Shariq Imran (described in section 2.3.1).

Deploying a Django website is rather complicated, and many web service providers do not run python code. After talkin with the head of forensic laboratory at NISlab, they were happy to host and administer the website at their server. The website is hosted at <http://tigger.tfl.nislab.no/the.forensic.challenger/>.

4.1 Scope and Limitations

Adding support for multiple browsers and file format were not prioritized when developing TFC. Currently the only supported browser is Google Chrome version 34. The web interface and video playback can behave unexpectedly on other browsers. The reason for this is mostly that the media codecs are different, and .ogg / .webM videos are not supported. The website appears different on various resolutions, the intended resolution is 1920x1080, if anything else is used, the user needs to zoom out to fit the four frames in the browser window.

To limit the size of the project, we focus on developing the functionality of the platform. Creating content and graphic design is a very time consuming process and is out of scope for this thesis. During the project work, there was another digital forensics II lab session with Jeffery Hamm, where the author participated to record the lecture so it could be used as content for the forensic challenger. However; the recording turned out to be of poor quality, and could not be used. The recording used in TFC is from the previous year, where the quality of the video and audio were better.

4.2 Findings - The Forensic Challenger

This section will explain the forensic challenger platform in detail. This includes decisions made regarding design and technical features based on the knowledge the author gained during the development.

4.2.1 Overview

The Forensic Challenger is hosted on a web server program called *Apache* and is built up using Django (section 2.3.2) and the corresponding modification "mod-wsgi" to handle the interactions with Apache. The conversational agent back-end is a python module called PyAIML, which is based on the A.L.I.C.E project. HyperText Markup Language (HTML5) and Cascaded Style Sheets (CSS3) as well as JavaScript handles the front-end

of the platform. TFC consists of four iframes. Each iframe holds one or more model that is included in TFC. The web interface is shown in figure 5.

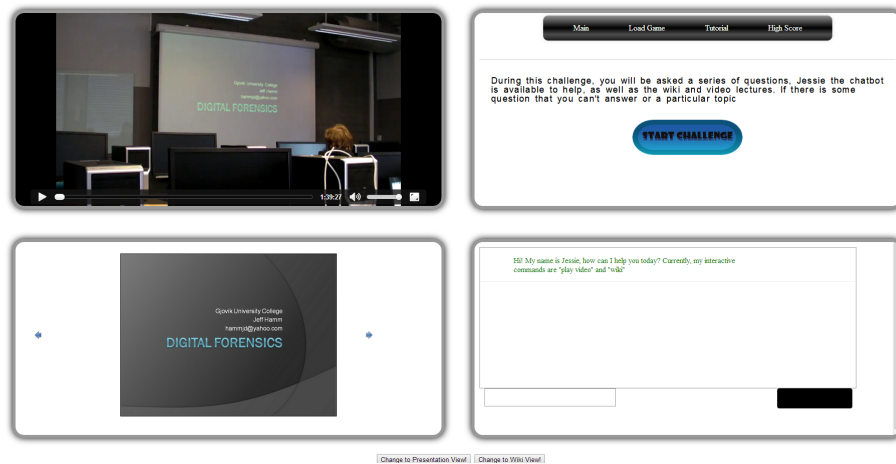


Figure 5: The Forensic Challenger Web Interface

The use of these four frames was the result of early testing of which features should be present in the main interface. They are mostly static except for the presentation/wiki frame. The reason for this being changeable is that the user won't need to look at the presentation slides, and the wiki at the same time.

The top-left frame is reserved for video playback, and is a JavaScript media player using the videojs[38] library. The bottom-left frame is used for presentation slides, as well as the Wiki page. The top-right frame is used for the challenges, this frame consists of a menu bar, to navigate through the different challenges, look up high scores or other parts related to the challenge. Most of the interaction with TFC happens from this frame. The bottom-right frame is used exclusively for the chatbot. The chatbot, named Jessie, uses the standard A.L.I.C.E AIML set, in addition to a set that is specific to TFC. Jessie can perform some operations similar to what is found in the challenge frame, play video and commands to open the Wiki.

4.2.2 Development

This section will describe the core concepts used in the development of the TFC platform. The experimentation and testing of features were mostly performed on the local Django development server, and later integrated to the deployment server.

During the development and action research process, these four frames were the subject of investigation. There is a great deal of changing between frames in TFC, most of these changes originate from the challenge frame. Due to a security policy in Javascript called the "Same Origin" security policy, which job is to limit the access of one window to another. That means you cannot change the content of a frame, which does not originate from the same protocol://domain:port, or in short, from the same origin. So we decided to limit the content to what is exclusively on the domain.

To change between the content the specific frames will display, we created a function to easy change the target of the frames.

```
function changeframe(frame , url)
{
```



```

    if (typeof(frame)=='string'){//find iframe (will not find a frame):
    frame=parent.document.getElementById(frame);
}
if (frame.tagName=='IFRAME'){//iframe :
    if (url==undefined){ return( frame.contentWindow.document.location );}
    frame.contentWindow.document.location.href=url;
} else{//frame:
    if (url==undefined){ return( frame.document.location );}
    frame.document.location.href=url;
}
}
}

```

This function requires two inputs, the first is the id of the frame will be changed, the second is the url of the content that will be displayed there. So if we want to change the wiki page to display the content of the presentation template, we can use it like this

```
changeFrame( 'wiki' , '/presentation' );
```

Back-end

There is a variety of back-end programming languages and frameworks available. The main programming languages used are PHP, Java, Python and Ruby. Under these programming languages, there is several frameworks that are continuously being developed, and have very good tutorials and community. Choosing between these frameworks was a matter of personal preference. The author have previously used python. To limit the time spent learning a new framework, the choice fell on Django.

Video and Presentation Playback

The core Hyper Interactive Presentation (HIP) functionality, comes from the HTML5 property *currentTime*. The *currentTime* property sets or returns the current position (in seconds) of the audio/video playback. When the property is set, the playback will jump to the specified position. With some additional code, you are able to synchronize presentation slides with the video, or jump to a position of interest. The video frame can be seen in figure 6 To be able to synchronize the presentation slides, you need a list of timestamps



Figure 6: The Forensic Challenger Video Frame

where the slides appear in the video. These timestamps can then be loaded into an array,

then use the `currentTime` property like below.

```
function next slide ()
{
    if (counter > 0)
    {
        counter += 1;
        video = innerdoc.getElementsByTagName('video')[0];
        video.currentTime = timeArray[counter];
    }
}
```



Figure 7: The Forensic Challenger Presentation Frame

The next slide function will be called from a HTML button used to go forward in the presentation. In addition, another function is called, this function changes the slide image itself.

```
function switchImage ()
{
    if (imageNum < imageArray.length - 1)
    {
        imageNum = imageNum + 1;
        if (imageNum < imageArray.length)
            document.getElementById('slideImg').src =
                imageArray[imageNum].src;
    }
}
```

This function checks whether there are any images left to change. The image URLs for the presentation is loaded into an array and the source of the image in the presentation frame is changed. Both of the previous functions have their negative counterpart, which handles going to the previous slide.

To be able to show the correct slide image if you are skipping in the video, or using `currentTime` with other functionality, we need some additional code to synchronize it. First we need to define a function that runs a loop to update the synchronization function. The synchronization function looks like this.

```
function process()  
{  
    video=innerdoc.getElementsByTagName('video')[0];  
  
    for(var g=0; g<imageArray.length; g++)  
    {  
        if(video.currentTime<timeArray[g]){  
            document.getElementById('slideImg').src=imageArray[g].src;  
            break;  
        }  
    }  
}
```

This function will keep running to always show the correct slide image based on what time the video is currently at.

One of the biggest challenges when writing the HIP functionality, was to make it as dynamic as possible. The original solution from the HIP platform [27] was to hard-code all timestamps and image names in the Javascript file, resulting in a similar but different Javascript file for each presentation. To be able to utilize our code more efficiently, we decided to import the timestamps and keywords from CSV files, and the image url's from folders. These values were placed in an array using a library called CsvToArray, published under the MIT license.

We also experimented with making a drag and drop Content Management System(CMS), where it was possible to drag the CSV files and images directly onto the website through a Java plug-in, and they were placed in the correct folders on the server. This solution was scrapped due to authentication problems. We looked into other various CMS for Django, but due to time restrictions we never got a working one in place.

Instead we made it as simple as possible to add new content, by structuring the folders on the server in a way that the name of the video file, was the name of the folders with corresponding material. The static files used with the HIP functionality is structured as shown in figure 8

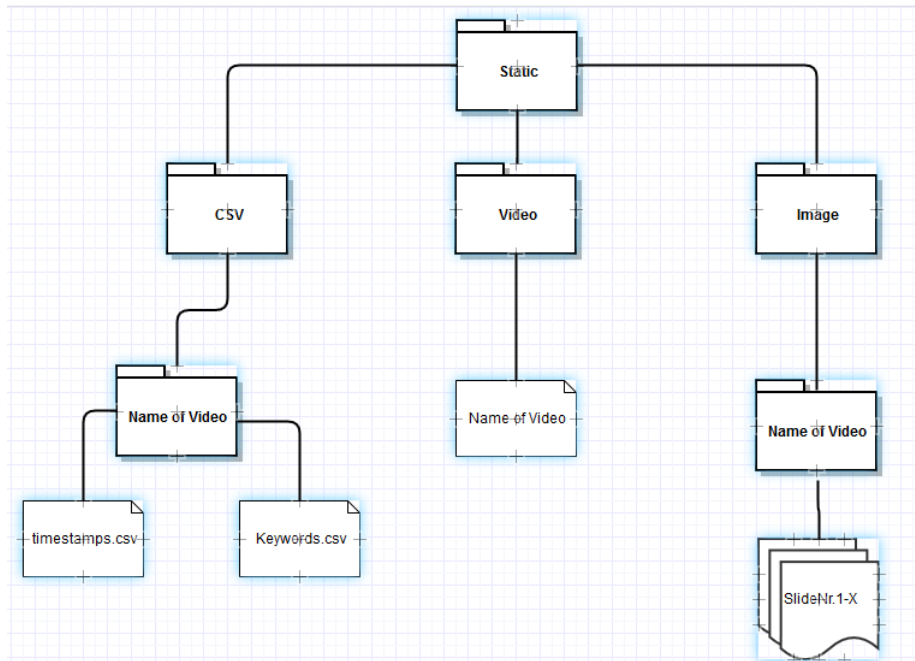


Figure 8: Content Structure for TFC

The name of the videos was placed in an array "videoid", and the functions to import images and csv files is based on this variable, below is part of the code for importing the images.

```
img.src = 'static/img/' + videoid[0] + '/Slide' + i + '.PNG';
```

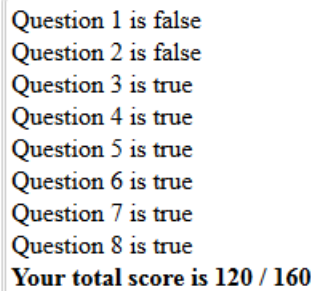
These images are pushed into an image array until all the images in the folder is loaded.

Placing the files in the correct folders with the right name is all that needs to be done to include a new presentation. Even without a working CMS, this process is quick if the system is familiar.

The Challenge

Designing the challenge frame itself was a simple process. It uses an out of the box CSS3 navigation bar to traverse between the different challenges, look at tutorials or high-scores. In the early literature review, some factors that may have a significant impact on motivation were discovered (see more in section 2.2.2). Multiple of these features were meant to be implemented. This included a character progression system. Some experimentation with this type of feature showed that it was too complex to implement in a short time. A character progression system would require a login and authentication service with a database to handle the users and save progression. Additionally it would require a lot of graphical interface design to have much effect.

Instead of the character progression system, we started looking at more primitive solutions. The score system that have been implemented is one of these features. After the challenge is complete, the game gives you the results of your answers in points. The questions answered will be either true or false, but it will not show what the right answer is. This is intentional, as it encourages the user to spend more time finding the correct answer. The score page is shown in figure 9.

A screenshot of a challenge score display. It lists the results of eight questions: Question 1 is false, Question 2 is false, Question 3 is true, Question 4 is true, Question 5 is true, Question 6 is true, Question 7 is true, and Question 8 is true. At the bottom, it states 'Your total score is 120 / 160' in bold.

Question 1 is false
Question 2 is false
Question 3 is true
Question 4 is true
Question 5 is true
Question 6 is true
Question 7 is true
Question 8 is true
Your total score is 120 / 160

Figure 9: End of Challenge Score

This score is only shown locally and not stored on the server in the current version. The idea was that the users can submit their score by clicking a button on the last page of the challenge. This score would then be transmitted to the database and shown on the high score list. The high score list have not been implemented in the last version.

The challenge itself is started after a brief introduction to the challenge, by clicking the "Start Challenge" button, as seen in figure 10. In case of the user have already completed the challenge, the next challenge can be loaded from the menu.

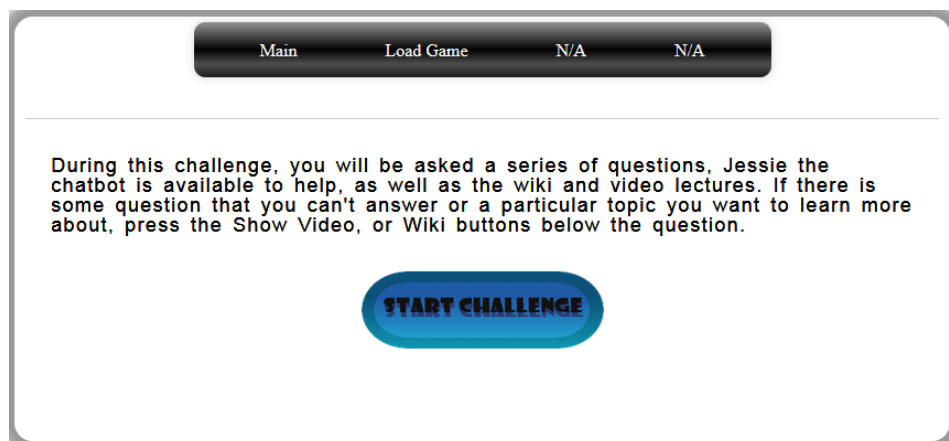


Figure 10: The Forensic Challenger Challenge Frame

The questions in this challenge have all the same format, a multiple choice question with varying number of answers. There are a few questions that show a picture which relate to the question, that will be shown below the question. We had some discussion regarding using video as a question as well, and not just for a source of knowledge. One of the things we thought about implementing was to show a short clip of digital forensics performed in a popular TV series or a Film, and ask the user why the approach they used in the video clip would not work in real life. No such questions were ever implemented, because there is a limited amount of clips where you get enough useful information about the process to have a solid idea of why it is wrong.

Figure 11 shows an example of one of the questions. In this example, the challenge gives the user a choice between 4 answers, if the user knows the answer, he can select his choice and click "Next". If the user wish to watch the specific position in the video where this is mentioned, he can click the "Show Video" button, and the video and presentation

Question 1:What is the difference between using the commands "parted" and "fdisk"?

- ☐ Finding partition offsets with fdisk is more suitable for mounting
- ☐ Fdisk lists the partition offsets in bytes
- ☐ None, the result is the same
- ☐ Parted lists the partitions offset in bytes

[Next >>](#) [Show Video](#) [Wiki](#)

Status Bar

Figure 11: Questions Layout

slides (if any) is loaded in the other frames at the correct position. If he desires to read on the wiki page, he will click the "Wiki" button, and the wiki page is loaded at the correct topic. When the user is comfortable with answering the question, he can select his choice and click next, and the next question is loaded. The progress bar will keep track of how many questions are answered in the challenge. After answering the last question, the user will get a score, each question is equally weighted at 20 points per right answer, and 0 points for a wrong answer. Additionally, it will list what question were right or wrong, but not what the correct answer is.

Wiki Page

To give the users access to information specifically for the challenge, we decided to create a knowledge database in form of a wiki. There are a few out of the box solutions for wiki's, but to be able to have more control over content was processed, we decided it was best to make one from scratch. Tutorials for the Django framework have a few very simple solutions for writing a wiki and specify it for the planned usage. These are not dependant on heavy plug-ins or libraries and work very well for small projects. We had a very simple wiki to begin with, but started to run into problems when using it on the deployment server. Wiki pages uses a syntax called Markdown, to link between related pages and make the text "pretty". The wiki we created was unable to use these markdown commands due to Django's Cross Site Request Forgery (CSRF) protection, this can be explicitly deactivated for the wiki, but we decided that it was a too big risk when putting it on the NISlab server. Instead we searched for other solutions, and found a out of the box wiki that was simple to integrate with Apache

This wiki page was a open source wiki called "django-Wiki"[39] designed specifically for Django. The framework was developed by Ben Jao Ming, with help from the development community. Figure 12 shows the wiki. Adding and modifying pages to this wiki is rather easy, but it is required that you log in. Currently, not much content have been added here, mostly pages that will help the user answering the questions in the challenge.

Populating a wiki with content is very time consuming and consequently only had a limited scope in the thesis. Discussions with the Forensic Laboratory at NISlab, we agreed there was a possibility to use "The Forensic Wiki", which is a rich source of knowledge for digital forensic investigations. However; as explained earlier with the Javascript "Same



Figure 12: The Forensic Challenger Wiki Frame

Origin" security policy, this solution proved difficult. Additionally, making changes at the Forensic Wiki to support the content of the challenge is equally time consuming.

There are different ways of reaching the wiki, if the button in the challenge is clicked, a keyword for the specific question is sent, and added to the url. For example:

```
changeiframe('wiki','wiki/'+KeyArray[counter]);
```

This will send you to the specific wiki page, based on which question you are currently answering.

Django-wiki uses a simple markdown format for editing text. For example, the way of linking from one topic to another is done by writing e.g

```
[Imaging](Imaging)
```

The first bracket shows what the displayed name of the link is, and the second is the direct link based on what is called a "slug". Which is basically a shortened name that removes spaces and other symbols.

The Django wiki have it's own CMS to be able to handle image uploads on the wiki, and will store these files directly on the server.

Chatbot

The chatbot is based on the work of Encorehu's "django-buddy"[40], which uses PyAIML (section 2.3.4), and is modified to work with the django framework. Django runs all the code, while a javascript handles the display of the conversation.

The chatbot of TFC have very limited functionality, the reason for this is that PyAIML in the current version (0.8.6) does not support javascript code. For useful features a chatbot can bring, see the future work section. However; the chatbot have the same functionality as the buttons in the challenge frame. You can tell it to "play video" and it will offer the same functionality as the Show video button, likewise with the wiki. Additionally you can say "I need help with (something)" and it will open the wiki and search for this specific keyword.

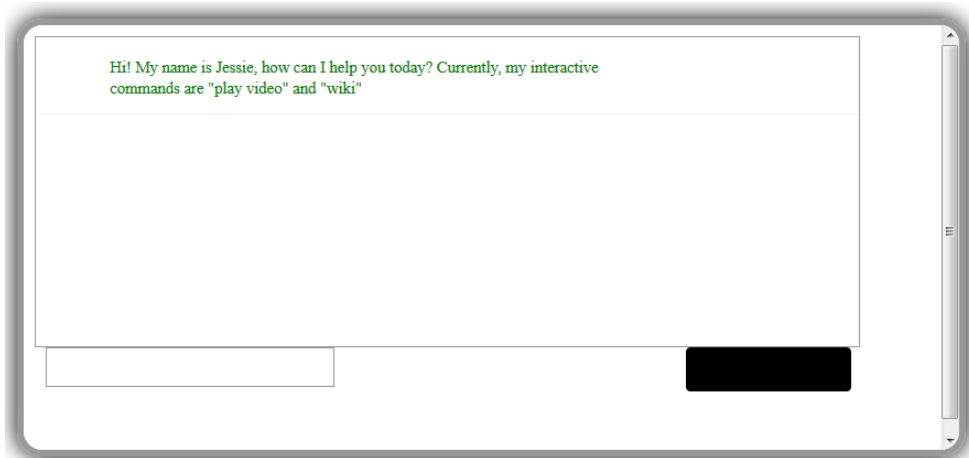


Figure 13: The Forensic Challenger Chatbot Frame

4.3 Experimental Trial - Fronter vs TFC

We have designed an experiment to test the effectiveness of the web interface compared to the alternative students at Gjøvik University College currently have.

This experiment was designed to interpret our findings during the action research process to see if the use of game elements will increase motivation and participation when learning digital forensics investigations. The experiment highlights the differences when using The Forensic Challenger versus the Fronter learning platform. Due to Fronter rooms being restricted to participants who are signed up for a certain course, our target group was unreachable in fronter. After discussing this with the head of the class, we decided it would be easier to provide the material needed for the experiment with a different approach, that is very similar to how fronter works.

The experiment requires between 10-15 participants, which will be divided into two groups.

Group A consisted of people gaining access to the recorded lab session, as well as the presentation slides in Digital Forensics II from 2013, held by Jeff Hamm. Additionally they gained access to a set of questions that needed to be answered, where the answers must be found in the video recordings. Due to restrictions in gaining access to the participants fronter room, we decided to simulate the usage of fronter by giving access to download the files directly over the "uninett filesender" feature.[41]. The additional presentation file and questions were sent to the participants over e-mail.

Group B gained access TFC platform, where the same video recordings and presentation slides are HiP'ed. The same set of questions are available for this group, but were presented as a game challenge on the TFC platform. Group B was asked to answer the questions using only the tools that are available on the TFC website, this includes the wiki page, and the video recordings/presentations. The questions are of varying difficulty, and the answers are found directly or indirectly by watching the video and look in the wiki.

The participants got about seven days to complete the experiment. Both groups was asked to participate in a short survey that determines if TFC have an impact on motivation for digital forensics. The full set of questions can be found in appendix A.

4.3.1 Questionnaire Design

The questionnaire was created using a web survey tool. Group A got a list of 7 questions, group B got 10 questions. They can both be found in the appendix A.

To be able to see what differences TFC have on motivation and participation, we made two separate questionnaires to be handed out to the participants in the experiment. They are very similar, except for some questions specific to the TFC platform for group B. Trying to measure motivation through questionnaires are not straight forward. Most questionnaires involving motivation usually base itself in the form of Likert-scale[42] agreement statements. You can't ask a participant a direct question about something they are not used to thinking about, as they would be forced to take a position on an issue they do not feel strongly about, or they may act on an attitude that they are unaware of or cannot express[43]. This is why it's suitable to use an experiment. If we manipulate one factor and then look at some measure of contributions or their characteristics as evidence of motivation, we get a view of motivation without directly asking about it.

To be able to establish if bringing game mechanics into e-learning will increase motivation to learn digital forensics, the questionnaire first establishes how much the subject is interested in digital forensics "*On a scale from 0 to 5, how interested are you in digital forensics?*". A motivated student can learn material regardless of how it is presented. To get further information about the motivation of the subject, we ask if the participant would recommend digital forensics and The Forensic Challenger platform to fellow students interested in digital forensics.

Further we asked *How much time did you use to find all the answers in the challenge?* to get an estimation of the effectiveness of the TFC platform to find the material they need, with this information, we have to compare the subjects of similar motivation level. Additionally we got an idea of how much time they were willing to spend trying to find the answers, in regards to the previously established motivation level.

An important question was asked for both groups, *On a scale from 0 to 5, how much did you learn from this challenge?*, this is not an easy question to answer, as the subjects might have learned something without being aware of it. Nevertheless, it will give an indication of the difference between the two approaches regarding learning outcome.

4.3.2 Results of Survey

This section shows the results of the survey, comparing group A, the fronter group to group B, the TFC group. Group A is the group which performed the experiment using only the provided raw material, while group B got access to the TFC platform. The number of participants were not as high as we had hoped, with a total of four for group A, and five for group B.

On a scale from 0 to 5, how interested are you in digital forensics?

Figures 14 and 15 shows how interested the participants are in digital forensics. The results show that the answers are similar for both groups, with a slightly higher interest in group B. These results were expected, and shows that the participants were evenly distributed to the two groups.

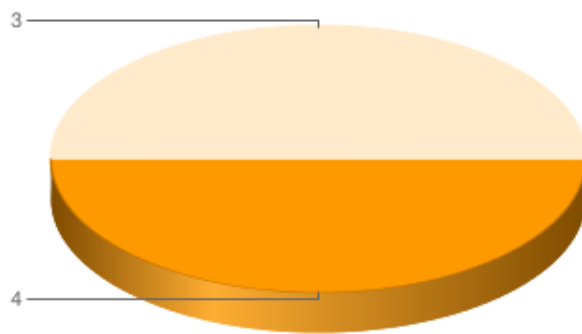


Figure 14: Question 1 Distribution with likert scale for group A

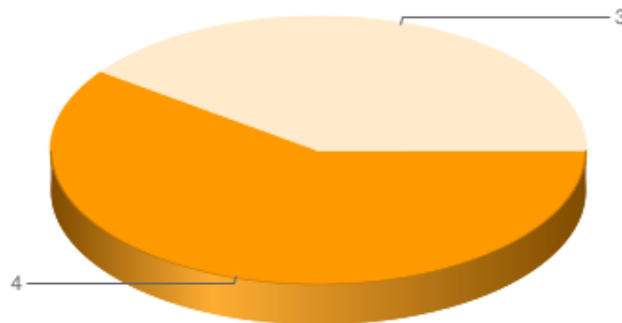


Figure 15: Question 1 Distribution with likert scale for group B

Would you recommend The Forensic Challenger / File-share approach to a fellow student who is interested in digital forensics?

These results showed clearly that group A was not particularly satisfied with the solution of giving the students the raw video footage and presentation slides from the lab session. In group B, all the participants answered that they would recommend TFC to a fellow student interested in digital forensics.



Figure 16: Question 2 Group A

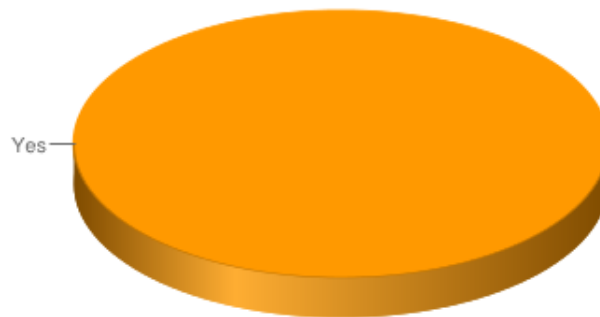


Figure 17: Question 2 Group B

How much time did you use to find all the answers in the challenge?

In this question we tried to estimate the time consumption of locating the correct information. As seen in figure 18 and 19, group A spent a lot longer time to answer the same set of questions. Most of the participants in group A spent 30 to 40 minutes, while on the contrary most of group B spent less than 10 minutes. Since the motivation level of the two groups were almost identical, the result from this question shows that group B had a much easier time answering the questions.

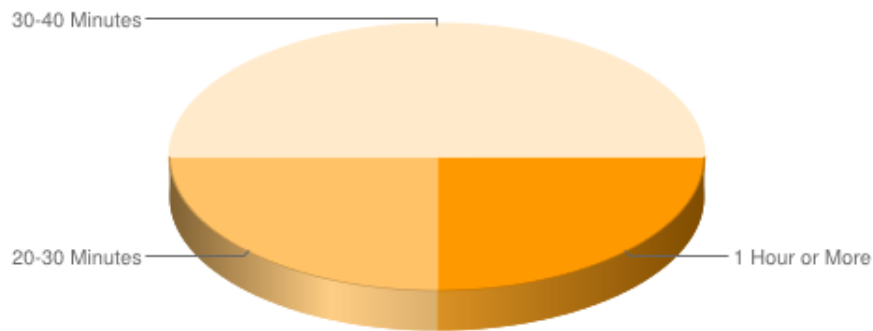


Figure 18: Question 3 Group A

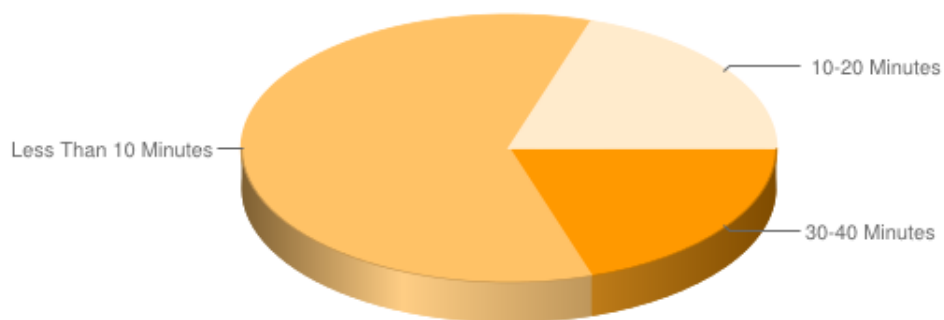


Figure 19: Question 3 Group B

Group A exclusive - How many questions were you able to complete?

For group A we asked a question about how many questions they managed to complete. Although group A does not get any feedback whether the answers are correct or not. Completion of the questions mean that they had obtained enough information to be confident in their answer.

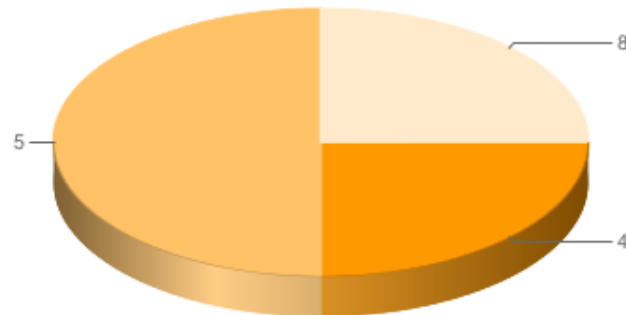


Figure 20: Question 4 Group A

Group B exclusive - How many of the answers did you find in the video recording?

Question 4,5 and 8 tried to establish what elements in the TFC platform which were most useful. For question 4, which is how many answers were found in the video recording, most participants selected 20 percent. This could have various reasons; the video recording itself was not very good, and the participants chose to look for information elsewhere. It could also mean that they are not visual learners and is subconsciously drawn to other types of information.

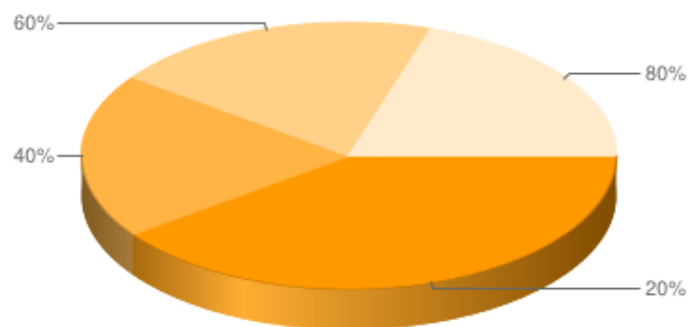


Figure 21: Question 4 Group B

Group B exclusive - How many of the answers did you find in the wiki page?

As with question 4, the number of answers found in the wiki page. The reasons for this could be the same as we discussed for the previous question. If these results are based on the learning style of the students or not is hard to determine without a learning style test. However; it may be that the wiki page shows the answers too obviously and makes it the most viable resource for answering the questions quickly.

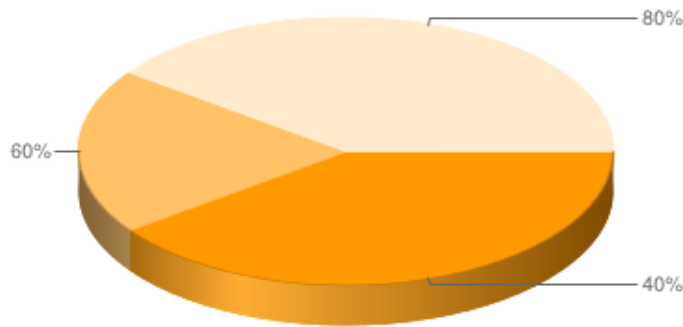


Figure 22: Question 5 Distribution with likert scale for group B

Group B exclusive - On a scale from 0 to 5, did you find the chatbot useful?

The usage of the chatbot was not expected to be high. It is in the latest version not very useful to begin with. All responses for this question was either 0 or 1.

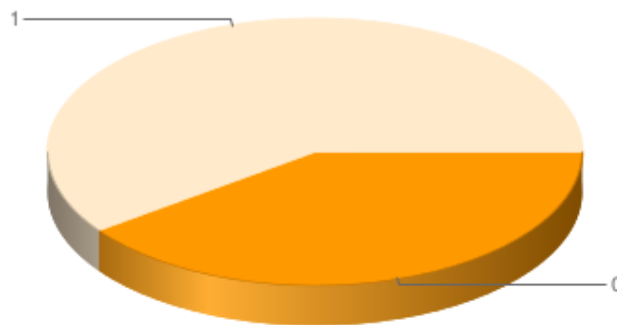


Figure 23: Question 8 Distribution with likert scale for group B

On a scale from 0 to 5, how much did you learn from this challenge?

The result of this question is of high value. The participants were asked how much they thought they learned from participating in the challenge, and the differences were quite obvious. Most of group A did not learn anything from the challenge, but two of the participants learned something. As the first question establish their interest in digital forensics, those answers seem to correspond to the results of this question. The higher motivated participants would learn something regardless of how it is presented to them. With group B, the participants seemed to have learned a lot from doing the challenge, with four of the five participants answering 4 on the scale.

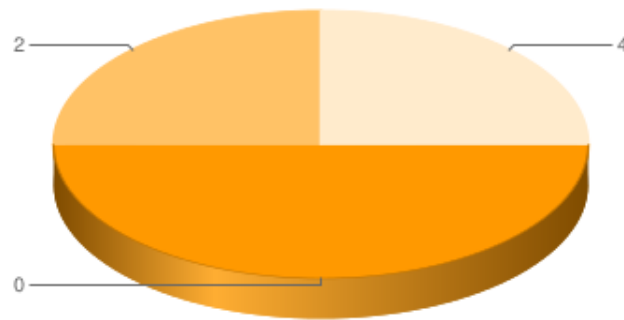


Figure 24: Question 5 Distribution with likert scale for group A



Figure 25: Question 6 Distribution with likert scale for group B

On a scale from 0 to 5, how much would you like to participate in a similar challenge?

Group A mostly did not wish to participate in a similar challenge, in contrary to group B where the interest was much higher.

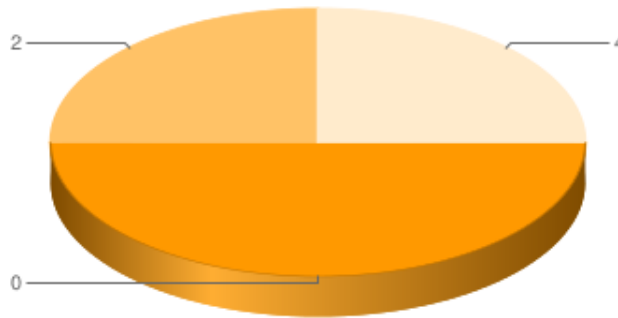


Figure 26: Question 6 Distribution with likert scale for group A

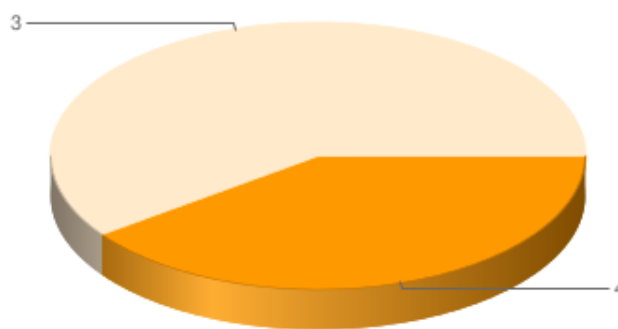


Figure 27: Question 9 Distribution with likert scale for group B

Group B exclusive - How would you improve The Forensic Challenger? As an open question for group B, we were interested in knowing how they would improve the forensic challenger platform. Their responses is shown below

1. Screen capture recording
2. Option to speed up video
3. Improve audio and give ability to boost the sound
4. More information in the wiki
5. Better video recording

The suggestions here are mostly based on the recordings and how they should be improved. This is a fair argument, as the recording itself was not very good and it was sometimes hard to understand the speech in the audio.

Would you be willing to participate in a more in-depth interview?

We hoped someone from group B was willing to participate in an interview were we

could establish in more detail how the platform was used, but unfortunately there was no volunteers.

5 Conclusions

In this thesis, we have researched how game elements improve the learning of digital forensics investigations. We developed an e-learning platform called "The Forensic Challenger", which is a web interface consisting of several modules. TFC is utilizing the four different learning styles proposed in the VARK model (Visual, Auditory, Read/write, Kinesthetic). By using Hyper Interactive Presentation (HIP) functionality, we are able to give each student the ability to learn with their own preferred learning style. The visual and auditory learners would prefer the video module, the students that prefer reading or writing would prefer the chatbot or the wiki. By using game elements, we tried on a basic level to satisfy the kinesthetic learning. This is different from other HIP implementations as the computer is the one asking the questions to the user, and not the other way around. Using game elements in an e-learning platform is a good method to increase motivation to learn the material. Developing the TFC platform was very time consuming, and a lot of features and solutions that would greatly benefit the platform could not be implemented due to time restrictions. The core functionality of the platform have been in focus. Any other functionality must be specified for the learning content, which in this case was limited due to a lack of recordings. Many of the respondents said the video quality was poor, and it was hard to hear what was being said, and that recording the screen of the presenter would be a helpful feature.

5.1 Research Questions

5.1.1 Can digital forensics investigation lab sessions be implemented on a web based e-learning platform?

The results of the survey show that the TFC platform have potential to become a valuable asset when teaching digital forensics investigations. The target group of the experiment should have little to no prior experience in doing digital forensics investigations, but should have a common interest since they are studying information security. All the respondents said that they would recommend TFC to a fellow student who is interested in digital forensics. Additionally they were interested in participating in a similar challenge.

Digital forensics investigation lab sessions can be implemented on an e-learning platform, however; the learning outcome is dependant on the content being appropriate for the platform. When creating the content, the platforms' possibilities and restrictions should be considered.

5.1.2 Does the e-learning platforms interactive elements help to improve motivation to learn digital forensics investigations

The usage of interactive elements to enhance motivation in learning and learning styles is discussed in section 2.2 Several interactive elements were implemented and designed to work with the digital forensics lab session. With the use of several learning styles, a student can choose his or hers preferred method of obtaining knowledge. The game elements is used to push the student further by offering rewards for progressing and completing something. Giving the student more control over the learning process is one

way of improving the motivation.

Looking at the results from the survey (section 4.3.2), we can say that the interactive elements affect the motivation of the students. Both the experiment groups had a similar interest in digital forensics, but a significant difference between the two experiment groups, was that group B who used the TFC platform, said they learned a lot more than group A, who only used the raw material provided to them.

One of the elements implemented was a chatbot, to satisfy the read/write learning style of the VARK model. The idea was to make the chatbot into a pedagogical agent, which could participate in discussions and ask or answer questions related to the content. Due to time limitations and complexity of an intelligent AI, it was not prioritized. The participants in the survey did not find the chatbot useful at all.

Using the TFC platform gives you the opportunity to easily find the correct information, without too much effort. Saving time and effort to find the information the student are looking for, improves the motivation.

5.2 Future Work

This section will reflect on some of the features that the author think can be improved or implemented in The Forensic Challenger.

In order to validate that the platform is appropriate for usage in an educational setting, the platform needs to be evaluated. A validation group would consist of 4-10 people who are active in the digital forensics field. This group would share their opinions of for example; how the platform performs, how it affects teachers and administration and how it would affect students. These validation groups are important in action research because the researchers opinion should not be the only one. It is important to get critical feedback on an action research project, critical feedback can be gained from a colleague or partner whose opinion you value, and gives you the ability to see your work in a new perspective [44].

Below is implementations and improvements the author feel is a starting point for further development.

Chatbot

A chatbot is a very interesting addition to an e-learning environment. However; for the chatbot to be useful, it will need the ability to intellectually communicate with the user. This gives the ability of having an somewhat intellectual discussion, to provide feedback and hints based on the learners performance, which may benefit learners greatly. A.L.I.C.E (AIML) is a pattern-matching program (where the "brain" is the AIML scripting language). It searches a large database for a phrase or term that matches something in the input, and selects a reply from the set designated from the closest match. It does not write to its own files or generates spontaneous output. Any changes or new information must be hard-coded into the AIML files.

Using a chatbot commonly called a "learning bot" is a much better approach to the path of achieving this goal. Natural Language Processing (NLP) is the computerized approach to analysing, attempt to understand or produce one or more human languages. The input might be text, spoken language or keyboard input. NLP can be used in an intelligent tutoring system, which have the ability to act as peers, co-learners, competitors, helpers or instructors. They should be able to ask and respond to questions, give hints

and explanations, monitor students and give feedback.

Kay [45] states that open learner models may enhance the student experience by encouraging effective learning, by creating opportunities for the learner to disagree or negotiate with the system, by asking the student to reflect on their knowledge and compare this with their student model, and by asking the student to use their model to identify areas to revise.[46][45]

Game Mechanics

The game mechanics in TFC is very limited. Implementing elements described in section 2.2.2; short-term, mid-term and long-term goals, as well as character progression and elements to motivate the user through competition and rewarding. Character progression can work very well with a HIP platform. For instance, the user can obtain knowledge points for watching a segment of a video or the whole video, and these points can then be used to unlock hints to answer a question.

Additionally, a system can be implemented to monitor the performance of the user, this can be done by watching which topics the user are struggling with, or what needs to be improved. The system can then select questions from a database of available questions, more suited for the users current knowledge level.

Learning Content

Standard presentations with presentation slides and a lecturer should be enhanced in the TFC platform. Other more interactive lectures like lab sessions, where practical work is performed and discussed in real time offers a bigger challenge. Using content from "outside the box" sessions requires some planning to make it really useful. For example, the use of a screen capture program in addition to video recording would be very helpful, but the work afterwards with editing and synchronization to make the material easily accessible would maybe prove difficult and time consuming. The correct usage of forensic tools is important, learn where the various tools performs well, and where they don't. TFC would benefit from a new module that is specifically designed to teach the student how to use tools.

Bibliography

- [1] Franzoni, A. L., Assar, S., Defude, B., & Rojas, J. 2008. Student learning styles adaptation method based on teaching strategies and electronic media. In *Advanced Learning Technologies, 2008. ICALT'08. Eighth IEEE International Conference on*, 778–782. IEEE.
- [2] Kessler, G. 2007. Online education in computer and digital forensics: A case study. In *System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on*, 264a–264a.
- [3] Zhang, D., Zhao, J. L., Zhou, L., & Nunamaker, Jr., J. F. May 2004. Can e-learning replace classroom learning? *Commun. ACM*, 47(5), 75–79.
- [4] Forensics wiki. http://www.forensicswiki.org/wiki/Training_Courses_and_Providers.
- [5] Pan, Y., Mishra, S., Yuan, B., Stackpole, B., & Schwartz, D. 2012. Game-based forensics course for first year students. In *Proceedings of the 13th Annual Conference on Information Technology Education, SIGITE '12*, 13–18, New York, NY, USA. ACM.
- [6] Sans institute. <http://www.sans.org/ondemand>.
- [7] Computer forensics training center online. <http://www.cftco.com/>.
- [8] International society of forensics computer examiners. <http://www.isfce.com/>.
- [9] Students create e-learning site on forensics. <http://gulfnews.com/news/gulf/uae/students-create-e-learning-site-on-forensics-1.813723>.
- [10] Accessdata - online training center. <http://www.accessdata.com/training/online-learning-center-lms>.
- [11] Felder, R. M. & Silverman, L. K. 1988. Learning and teaching styles in engineering education. *Engineering education*, 78(7), 674–681.
- [12] Alaeddin MH Alawawdeh, Ali Shariq Imran, S. J. K. Norwegian information security lectures as a case study for hyper interactive presenter.
- [13] Drago, W. A. & Wagner, R. J. 2004. Vark preferred learning styles and online education. *Management Research News*, 27(7), 1–13.
- [14] The vark modalities. <http://www.vark-learn.com>.
- [15] Guyan], M. Improving motivation in elearning. <http://elearnmag.acm.org/featured.cfm?aid=2527388>.
- [16] Intrinsic and extrinsic motivation. http://en.wikipedia.org/wiki/Motivation#Intrinsic_and_extrinsic_motivation.

- [17] Raymer, R. & Design, E.-L. 2011. Gamification: Using game mechanics to enhance elearning. *Elearn Magazine*, 2011(9), 3.
- [18] Prensky, M. 2001. Fun, play and games: What makes games engaging. *Digital game-based learning*, 1–31.
- [19] Oliver, R. & Herrington, J. 2001. *Teaching and learning online: A beginner's guide to e-learning and e-teaching in higher education*. Edith Cowan University. Centre for Research in Information Technology and Communications.
- [20] Fu, F.-L., Su, R.-C., & Yu, S.-C. 2009. Egameflow: A scale to measure learners' enjoyment of e-learning games. *Computers and Education*, 52(1), 101 – 112.
- [21] Oliver, R. & Australia, W. 2004. Factors impeding instructional design and the choice of learning designs in online courses.
- [22] White, B. Y. 1984. Designing computer games to help physics students understand newton's laws of motion. *Cognition and instruction*, 1(1), 69–108.
- [23] Garriss, R., Ahlers, R., & Driskell, J. E. 2002. Games, motivation, and learning: A research and practice model. *Simulation & gaming*, 33(4), 441–467.
- [24] Whitehill, B. V. & McDonald, B. A. 1993. Improving learning persistence of military personnel by enhancing motivation in a technical training program. *Simulation & Gaming*, 24(3), 294–313.
- [25] Wood, L. E. & Stewart, P. W. 1987. Improvement of practical reasoning skills with a computer game. *Journal of Computer-Based Instruction*.
- [26] Thomas, R., Cahill, J., & Santilli, L. 1997. Using an interactive computer game to increase skill and self-efficacy regarding safer sex negotiation: field test results. *Health Education & Behavior*, 24(1), 71–86.
- [27] Ali Shariq Imran, S. J. K. Hip - a technology-rich and interactive multimedia pedagogical platform.
- [28] The django project - documentation. <https://www.djangoproject.com/>.
- [29] Djangobook. <http://www.djangobook.com/en/2.0/chapter04.html>.
- [30] Shawar, B. A. & Atwell, E. 2007. Chatbots: are they really useful? In *LDV Forum*, volume 22, 29–49.
- [31] Aimpl documentation. <http://www.alicebot.org/.../aiml-primer.html>.
- [32] Pyaiml documentation. <http://pyaiml.sourceforge.net/>.
- [33] Action research in education. <http://www.edu.plymouth.ac.uk/resined/actionresearch/arhome.htm>.
- [34] Salite, I., Mičule, I., Kravale, M., Iliško, D., & Stakle, A. 2007. Toward the sustainability in teacher education: Promise of action research. *Education and sustainable development: First steps toward changes*, 2, 263–292.

- [35] Herington, C. & Weaven, S. 2008. Action research and reflection on student approaches to learning in large first year university classes. *The Australian Educational Researcher*, 35(3), 111–134.
- [36] Kapenieks, J. & Salīte, I. 2012. action research for creating knowledge in an e-learning environment. *Journal of Teacher Education for Sustainability*, 14(2), 111–129.
- [37] Nsw department of education. <http://www.dec.nsw.gov.au>.
- [38] Videojs - open source html5 video player. <http://www.videojs.com/>.
- [39] Django-wiki. <https://github.com/benjaoming/django-wiki>.
- [40] Django-buddy. <https://github.com/encorehu/django-buddy>.
- [41] Uninett filesender. <https://filesender.uninett.no/>.
- [42] Likert-scale. http://en.wikipedia.org/wiki/Likert_scale.
- [43] Measuring motivation. <http://www.technotaste.com/blog/measuring-motivation/>.
- [44] McNiff, J. 1995. *Action research for professional development*. Hyde Publications Bournemouth.
- [45] Kay, J. 1997. Learner know thyself: Student models to give learner control and responsibility. In *Control and Responsibility, International Conference on Computers in Education, AACE*, 17–24.
- [46] Kerly, A., Hall, P., & Bull, S. 2007. Bringing chatbots into education: Towards natural language negotiation of open learner models. *Knowledge-Based Systems*, 20(2), 177 – 185. {AI} 2006 The 26th {SGAI} International Conference on Innovative Techniques and Applications of Artificial Intelligence.

Appendices

A Questionnaires

Webpage Screenshot

The Forensic Challenger Group A

1. On a scale from 0 to 5, how interested are you in digital forensics?

☐ 0
☐ 1
☐ 2
☐ 3
☐ 4
☐ 5

2. Would you recommend this filesender approach to fellow students interested in digital forensics?

☐ Yes
☐ No
 If not, why?

3. How much time did you use to find the answers in the challenge?

☐ Less than 10 minutes
☐ 10 - 20 minutes
☐ 20 - 30 minutes
☐ 30 - 40 minutes
☐ 40 - 50 minutes
☐ 1 hour or more
 Other (please specify)

4. How many questions were you able to complete?

☐ 0
☐ 1
☐ 2
☐ 3
☐ 4
☐ 5
☐ 6
☒ 7
☐ 8

5. On a scale from 0 to 5, how much did you learn from this challenge?

☐ 0
☐ 1
☐ 2
☐ 3
☐ 4
☐ 5

6. On a scale from 0-5, how much would you like a similar link to another subject in digital forensics?

☐ 0
☐ 1
☐ 2
☐ 3
☐ 4
☐ 5

7. Would you be willing to participate in a more in-depth interview?

☐ Yes
☐ No
 If yes, provide email

Fertig

Zurück zu SurveyMonkey
 Öffnen Sie es, um Ihre Antworten zu überprüfen und zu bewerten

The Challenge

1. On a scale from 0 to 5, how interested are you in digital forensics?

- ☐ 0
☐ 1
☐ 2
☐ 3
☐ 4
☐ 5

2. Would you recommend The Forensic Challenger to a fellow student who is interested in digital forensics?

- ☐ Yes
- ☐ No

If not, why?

Page 1 of 1

3. How much time did you use to find all the answers in the challenge?

- ☐ Less than 10 minutes
- ☐ 10 - 20 minutes
- ☐ 20 - 30 minutes
- ☐ 30 - 40 minutes
- ☐ 40 - 50 minutes
- ☐ 1 hour or more

Other (please specify)

4. How many of the answers did you find in the video recording?

- ☐ 0 %
☐ 20 %
☐ 40 %
☐ 60 %
☐ 80 %
☐ 100 %

5. How many answers did you find in the wiki page?

- ☐ 0 %
- ☐ 20 %
- ☐ 40 %
- ☐ 60 %
- ☐ 80 %
- ☐ 100 %

6. On a scale from 0 to 5, how much did you learn from this challenge?

- ☐ 0
☐ 1
☐ 2
☐ 3
☐ 4
☐ 5

7. How would you improve The Forensic Challenger for learning digital forensics?

8. On a scale from 0-5, did you find the chatbot useful?

- ☐ 0
☐ 1
☐ 2
☐ 3
☐ 4
☐ 5

9. On a scale from 0-5, how much would you like to take another forensic challenge?

- ☐ 0
☐ 1
☐ 2
☐ 3
☐ 4
☐ 5

10. Would you be willing to participate in a more in-depth interview?

- ☐ Yes
- ☐ No

If yes, provide email

Done

Powered by **SurveyMonkey**
Check out our [sample surveys](#) and create your own now!

B Specialization Course II - Digital Forensics Game

Work done by the author for specialization course II at Gjøvik University College

The Digital Forensics Game

Øyvind Aaen Nordhaug

December 18, 2013

Abstract

Games for learning is becoming a more and more popular form of education. The ability to teach many more students at once, and introducing elements of fun to further motivate the students. Knowing how digital forensics investigations are performed is becoming need-to-know knowledge in many fields. The society's dependency on computers, and the massive increase in computing power and data storage, causes the society to require more expertise and efficient methods to handle digital forensics investigations. This projects goal is to create a digital forensics learning game, which will utilize course material on the Hyper Interactive Presentation Platform (HIP) as a source of knowledge. By interacting with an AI, the user will have to teach a robot to perform the investigation.

Preface

This project is a continuation of the pre project planning report, and will explain the game details and development of the master thesis project. The pre project planning report is included as appendix.

Contents

| | | |
|-------|--|----|
| 1 | Introduction and Motivation | 3 |
| 2 | High Concept | 3 |
| 3 | Game Description | 3 |
| 3.1 | Educational Value | 4 |
| 3.1.1 | Learning Games | 4 |
| 3.2 | Challenges | 6 |
| 3.3 | Target Audience | 6 |
| 4 | HIP Platform | 6 |
| 4.1 | HIP Improvement | 6 |
| 4.2 | Dynamic HIP with Django - Work in progress | 7 |
| 4.2.1 | Databases | 8 |
| 4.2.2 | Javascript | 8 |
| 4.2.3 | Admin Actions | 9 |
| 5 | Content | 10 |
| 6 | Interface | 10 |
| 7 | Development | 11 |
| 7.1 | PyGame | 11 |
| 7.2 | PyAIML | 11 |
| 7.3 | Development Plan | 12 |
| 8 | Jessie - the AI | 13 |
| 8.1 | AIML | 13 |
| 8.2 | Categories | 14 |
| 8.2.1 | Atomic | 14 |
| 8.2.2 | Default | 14 |
| 8.2.3 | Recursive | 15 |
| | References | 16 |

1 Introduction and Motivation

Digital information grows at an exponential rate. Data is stored on not only computer hard drives or CD's, but also more and more on portable data and storage devices, such as smart phones. The information stored on these devices is a crucial source of evidence in civil, corporate or criminal investigations.

The need for proper e-learning platform is of great importance, as the challenges in the digital forensics field is increasing in complexity. E-learning grants the opportunity to teach many people that comes from different fields, at the same time. As the complexity of the topics increases, e-learning provides the ability to learn the material in your own speed. One person may understand everything on the first attempt, and others can go back and look up what is unclear. See appendix 1.1 for details

In this paper, we will present an idea of a digital forensics learning game, where the student play as a digital forensics investigator and his job is to teach the player's sidekick, Jessie, to perform the investigation.

2 High Concept

In DFGame you play as a digital forensics investigator. You have been assigned to a crime scene where criminal activity has taken place. Your newly assigned partner "Jessie", a robot straight from the factory, have been tasked to accompany you. Your job as the lead investigator is to analyse this crime scene by the help of Jessie. DFGame is a 2D text-based adventure game.

3 Game Description

The goal of DFGame is to teach a simple Artificial Intelligence (AI) how to perform a digital forensics investigation. The game will be supplemented with the Hyper Interactive Presentation (HIP) platform, which will include course material in form of both video presentations and reading material. Viewing the course material on the HIP platform grants the player points, and will cause the AI to unlock that knowledge, and it can be used in the game.

The player will be able to gain points by performing other various actions, for example: Managing to correctly make a hard drive image in a forensically sound manner. These points are needed to gain access to new tasks that have to be solved. These points accumulate throughout the game, and will be used in a leaderboard to indicate the performance of the players. The more course material that is used, and tasks solved, the higher the score will be.

The interaction with the AI will be the key to perform the investigation. By using the course material on the HIP platform, the player can learn the way a digital forensics investigation is performed. When the course material have been accessed, the game will give the option of downloading the new knowledge to the AI, which will then understand the commands. For example: The player learns about how to capture a image of the

hard drive in a forensically sound manner. Then proceeds to download this information to the AI. Then issue the command to the AI to begin the imaging. If the player already know how to capture a HDD image, simply telling it how to perform the task is also possible.

3.1 Educational Value

Most people are familiar with learning through books or in classes, or from teachers and peers, but most people overlook a different approach to learning, which is teaching. The best test of whether or not you really understand a concept is trying to teach it to somebody else. The great learning potential inherent in teaching would appear to be generated as the results of a particular aspect of the teaching process itself: the encounter with diversity, which on the one hand tends to increase reflexivity, while on the other hand tends to break down resistance to change[5]

In a paper published by Yin Pan et.al. "Game-based Forensics Course For First Year Students"[9] The authors have designed a forensics course using the game-based learning(GLB) approach. This system has direct access to forensics tools located on a forensics machine, and evidence from a suspects machine. This is a game that is suitable for first year students in college. In their paper they claim the use of game based learning for forensics is a novel idea, especially in combination with the use of visualization technologies to understand abstract concepts. [9]

Disregarding the use of visualization techniques, the game proposed in this report, and the paper published by Yin Pan et.al, are very similar.

3.1.1 Learning Games

So what makes learning games successful? In a paper by Marc Prensky [10], he talks about the element of "fun", and calls it "the great motivator". He further discuss the relationship between fun and learning and thinks the principal roles of fun in the learning process is to create relaxation and motivation. Relaxation enables the learner to take things in more easily, and motivation enables them to put forth effort without resentment.

There are a number of ways to discuss the forms and roles which online learning materials can take. Oliver and Herrington suggests a possible way of classifying and comparing online learning materials[8]

| Form | Description of Materials |
|-----------------------|---|
| Information Access | Convey information alone to the learner, for example, a course syllabus, a calendar, assignment descriptions, lecture notes, workshop descriptions etc. |
| Interactive Learning | Involve instructional elements that engage the learner, encourage reflection and decision making and provide feedback in response to learner actions |
| Networked Learning | Provide a means for the organisation, communication and exchange of ideas and information among learners and teachers and other parties in the learning process |
| Materials Development | When the online setting is used as a means for learners to create and publish materials. For example, learners publishing web pages, stories etc. |

Table 1: Forms of online learning materials[8]

Learning games fits under the Interactive Learning category, and are characterised by forms of learner control and active engagement where learners can take and make decisions and learn through the consequences of their actions. Some of the elements of interactive learning is[8]:

1. Activities which require students to search and review documents and links in a structured and organised fashion to discover particular information.
2. Questions/activities that are posed which cause students to reflect and to select from among various outcomes, with feedback particular to the students choices being provided and able to influence subsequent directions and activities presented.
3. Through the use of forms, where students can select options to effect particular processing outcomes, for example, database searching, selection processes, simulation control.
4. Through the use of program modules which enables the learners input to be processed and acted on, for example, shockwave movies, virtual reality environments etc.

Interactive elements supports learning in online settings through means such as the possibility of immediate feedback to responses given to tasks and problems and a capacity to respond to the user in ways which recognise the discrete needs of individuals. Additionally you have means of presenting tutorial type activities, which support knowledge acquisition through such tasks as drill and practice, as well as simulations of real like events, which learners can learn to manage and control in a safe and structured setting[8].

The popularity of the Internet has led to several advancements in educational technology. Still most e-learning solutions utilizes traditional instruction contents, which are

placed on the Internet without modifications that take advantage of the full potential of either the Internet or online technology[6][7].

3.2 Challenges

The biggest challenge with developing a game like this, is the difficulty level of the tasks given. The reason for this is that by having difficult tasks to complete, the AI would have to be very flexible. There is usually many ways to solve a task, and many tools that can be used. Therefore by limiting the solutions to solving the tasks to the methods explained in the course content, the solutions may be too easy to find.

3.3 Target Audience

The target audience for this game is students who are new to the digital forensics field or people in other fields, like lawyers or judges, who are interested in learning about digital forensics investigations.

4 HIP Platform

The Hyper Interactive Presentation (HIP) platform, created by Ph.d Ali Shariq Imran for the Ericsson Response Training programme, provides a web interface that is constructed out of several modules that interact with each other. The two main modules are the presentation slides, with the corresponding video capture of the presentation. The other modules are keywords that are often used in the presentation, as well as frequently asked questions. These keywords and questions can be clicked, and you will be taken to the time in the presentation where this question was addressed.

The DFGame will utilize this platform as a source of information for the player (see section 5 for details). During the game it is possible to interact with the HIP platform modules to look up additional information about a particular topic or view a presentation.

4.1 HIP Improvement

The HIP platform as it is now, does not handle dynamic content, and each presentation will have to have their own javascript which includes hardcoded timestamps for the video and presentation synchronization, as well as hardcoded image names.

Additionally, the use of multiple javascript files that does more or less the same work is undesirable, so a main javascript that requests the timestamps and images from the database so all presentations could be handled from the same javascript is the desired outcome.

What should be done with the HIP platform is making a database which includes the timestamp information for each presentation, these timestamps should be imported from a excel sheet in a CSV format. No standard format for these excel sheets have been agreed upon at this time. Additionally the filenames and paths for the presentation

images should be imported from the database. Additionally, the HIP platform should have a feature where administrators can authenticate against the service and perform maintenance and edit and upload the material on the HIP site.

4.2 Dynamic HIP with Django - Work in progress

To make this dynamic approach much easier, I decided to start remaking the whole website using Django as a back-end. JavaScript will still take care of the client side interactions on the HIP platform. As the DFgame will be developed in Python as well, the use of a Django based backbone for the HIP platform will hopefully make this process smoother.

Django is a high-level Python web framework, that offers security features to protect against Cross Site Scripting attacks, Cross site request forgery attacks, sql injection attacks, clickjacking and more. In Django you define data models entirely in Python, and it gives you a dynamic database-access API. When starting a new Django project, by issuing the command "django-admin.py startproject mysite", it sets up the basic structure of the web site for you. That could look something like below[2].

```
mysite/
  manage.py
  mysite/
    __init__.py
    settings.py
    urls.py
    wsgi.py
```

The manage.py file is a command line utility that lets you interact with the Django project in different ways. In addition to provide a help function, it gives you the opportunity to create new "Apps", run a shell to your database, run the server on a specified port and much more.

In figure 1, we have created an app called presentations, for use in the HIP platform, django have created the necessary files, and after adding the new app in the list of apps to load, and add some url expressions, we can start populating the app with javascript, html files, views and modules to create interaction on the site.

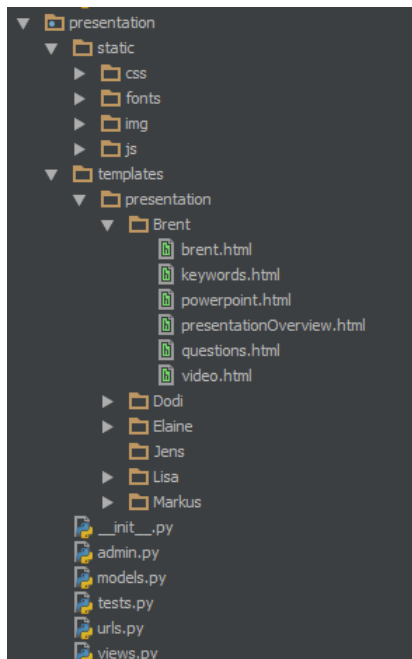


Figure 1: Django App Files

The settings.py file contains all the configuration of the Django installation. In here we can in addition to other things, specify the database we want to use, as shown in Figure 2.

```
DATABASES = {
    'default': {
        'ENGINE': 'django.db.backends.sqlite3',
        'NAME': os.path.join(BASE_DIR, 'hipdatabase.sqlite3'),
    }
}
```

Figure 2: Django settings - Database

4.2.1 Databases

The Django framework gives you an easy way to create and interact with databases. You can use several different types of databases, which includes PostgreSQL, MySQL, SQLite (included in python) and Oracle Database Server. In addition to the officially supported databases, you can also use databases like Sybase SQL Anywhere, IBM DB2, Microsoft SQL Server 2005, Firebird and ODBC

Django lets you interact with databases by the use of data models. A model class represents a database table, and an instance of the class represents a particular record

in the database table. A class like this can look like the following in python code[2]:

```
class Entry(models.Model):
    pres = models.ForeignKey(pres)
    headline = models.CharField(max_length=255)
    body_text = models.TextField()
    pub_date = models.DateField()
```

4.2.2 Javascript

Since the javascript files used in the HIP platform have hardcoded inputs of images and timestamps etc. We would need to make significant changes to the script, and make it include all these things from the database, this is still in development. Currently the site only loads the original javascript files from the HIP platform on each of the presentations provided. To call a javascript file in a html document, you must modify the code by adding

```
{% load staticfiles %}
```

And then refer to the static file you wish to use

```
<link rel="javascript" type="text/javascript" href="{% static
    'js/script.js' %}">
```

This fetches the javascript file located in the static directory. This directory have to be specified in the settings.py file by adding the filepath where the directory is located. In this case its in the project directory under the folder "static"

```
STATIC_ROOT = os.path.join(PROJECT_DIR, 'static')
STATIC_URL = '/static/'
```

4.2.3 Admin Actions

The Django framework comes with an admin interface which is very customizable after the needs of a specific app. It gives you the possibility of writing and register "actions" on the admin page. On the HIP platform this can be the ability to upload material and/or delete/edit it.

5 Content

The course content that will be implemented will be based on the digital forensics lab lectures at Høgskolen i Gjøvik (HIG). This content should have modules like:

1. Incident Response - Where to find the evidence (ram, hdd, cd, cellphones etc.)
2. Forensics Essentials and General Procedure
3. Volatile data - Dump memory, find encryption keys
4. Imaging techniques and tools - how to do imaging
5. Preservation of the image - Validate images with hash functions
6. Unix Filesystem and Unix Forensics Techniques
7. Windows Filesystems and Forensics Techniques
8. Forensics Investigation Reports

These topics are subject to change when the game is being developed as a master thesis project. There will need to be some research conducted into what topics are most important and relevant for a digital forensics investigator today.

6 Interface

The interface for the DFGame will consist of several modules, including a menu bar where the player can choose tasks, access the HIP platform and perform various operations to the AI bot etc. Additionally the interface will have a module that handles the communication with the AI, where you can chat with the AI and tell it to perform an action. Figure 1 show an example of what the interface may look like



Figure 3: Example of interface

7 Development

The game itself will be developed using Python. PyGame offers a set of python modules that is designed for writing games, and adds functionality on top of the SDL library. PyGame is highly portable, and runs on nearly every platform and OS. As this game will be developed in 2D, a solution like this is sufficient. In addition to PyGame, we will be using PyAIML, which is a interpreter for the AIML language.

7.1 PyGame

Using python to make games might not be suitable in all cases, it depends on the game. A game is usually split in two parts; the game engine, which must be as fast as possible, and the game logic, which makes the engine do something. PyGame and SDL (Simple Directmedia Library) serves as an excellent C engine for 2D games. Most of the runtime in games is spent inside SDL handling the graphics. By enabling the use of graphics hardware acceleration, you get a good speed with python. PyGame consists of several modules, below is a list of the most relevant modules[4]:

1. cdrom - manages cdrom devices and audio playback
2. cursors - load cursor images, including standard cursors

3. display - control the display window or screen
4. draw - draw simple shapes onto a surface
5. event - manage events and the event queue
6. font - create and render Truetype fonts
7. image - save and load images
8. key - manage the keyboard
9. mouse - manage the mouse
10. movie - playback of mpeg movies
11. time - control timing
12. transform - scale, rotate and flip images.

7.2 PyAIML

PyAIML is an interpreter for AIML, the Artificial Intelligence Markup Language, implemented as a standard python package. Using the PyAIML package is pretty simple[3]:

1. First you have to import the AIML package

```
import aiml
```

2. Then you have to instantiate a kernel object, the kernel is the only class you need to use when dealing with PyAIML

```
kernel = aiml.Kernel()
```

3. The kernel is now ready to respond to user input, so the next step is to load the AIML files, for this we will use the kernels learn() method

```
kernel.learn("std-startup.xml")
```

4. In this example, we have loaded the startup file from the standard AIML set, it defines a single AIML pattern, "LOAD AIML B", which causes the rest of the set to load. To trigger this process, we pass "load aiml b" as input to the kernel.

```
kernel.respond("load aiml b")
```

5. The respond method() returns a string containing the Kernel's response to the input. We are now ready to start responding intelligently to user input.

```
while True: print kernel.respond(raw_input("> "))
```

6. To reduce startup time, we can use the kernels loadBrain() and saveBrain() methods, these functions lets you dump the contents of the bot's brain to a file. So next time you will only have to use the loadBrain() function.

7.3 Development Plan

The development of the game will be performed in stages, beginning with performing an requirement analysis to determine the needs or conditions for the project.

The requirement analysis will specify all necessary requirements that are required for the project development. This includes what digital forensics content to use. See appendix 1 chapter 3 for details. Additionally we need to specify what changes must be done with the HIP platform to make it easier to use within the game.

When the requirement analysis is complete, we will start designing the software, implementing it and do testing. In accordance with the time estimations proposed in the pre project planning report (see appendix 1 - 4.4 for details). We can roughly estimate the time consumption in the development process.

| Task | Task Description | Theory | Lab | Doc | Supervision |
|-------|--------------------------------|--------|---------|---------|-------------|
| 1 | Requirement Analysis | 15 | 0 | 70 | 3 |
| 2 | Software Design | 10 | 0 | 80 | 2 |
| 3 | Implementation and Integration | 5 | 180 | 10 | 2 |
| 4 | Testing | 0 | 80 | 0 | 0 |
| Total | N/A | 30 hrs | 300 hrs | 120 hrs | 8 hrs |

Table 2: Time Estimation: Theory, Lab Work, Documentation, Supervision

8 Jessie - the AI

The AI, Jessie, will be programmed using Artificial Intelligence Markup Language(AIML). AIML is a derivative of Extensible Mark-up Language(XML) and was developed by Wallace and the Alicebot free software community.

AIML is rule-based, matching a sequence of words to generate either a completely canned response, or a response involving substitution of input words into an output template.

8.1 AIML

AIML consists of data objects called AIML objects. These objects are made up of units called topics and categories. Topics is an optional element, and has a name attribute and a set of categories related to that topic. Categories are the basic unit of knowledge in AIML. Each category is a rule for matching an input and converting to an output, and consists of a pattern, which matches against the user input, and a template, which is used in generating the chatbots answer. The AIML format is as shown below.[11]

```
<aiml version="1.0">
<topic name="the topic">
<category>
<pattern>PATTERN</pattern>
```

```

<that>THAT</that>
<template>Template</template>
</category>
..
..
</topic>
</aiml>

```

The tags used in the example is explained below. In addition to these tags, there is the <this>, <random>, <star>and <think>tags

1. The <aiml>tag is the parent tag for the entire AIML document.
2. The <topic>tag is used if there has been a topic <set>matching the query.
3. The <category>tag should contain the "matching" part of the expression, and the "returning" part of the expression
4. The <pattern>tag is the matching part, and specifies which pattern to match.
5. The <that>tag is used when the current pattern is depending on an earlier output.
6. The <template>tag specifies how the bot responds when the pattern within the category is matched.

8.2 Categories

AIML have three types of categories, atomic categories, default categories and recursive categories. These three types overlap, because atomic and default refer to the <pattern>and recursive refers to a property of the <template>.

8.2.1 Atomic

Atomic categories are those with atomic pattern, for example the pattern contains no wild card "*" or "_" symbol. Atomic categories are the easiest categories to add in AIML.[11]

```

<category>
<pattern>What is a hard disk drive?</pattern>
<template><set_it>hard disk drive</set_it>Hard disk drives(HDD) is a
    data storage device for digital information
</template>
</category>

```

The category above matches the client input of "what is a hard disk drive?" and sets the IT variable to the value of "hard disk drive" and sends the response in the template

8.2.2 Default

The default category derives from the fact that its pattern uses a wildcard "*" or "_". These default responses are often called "pickup lines" because they generally consists of leading questions designed to focus the client on known topics. The more common default categories have patterns combining a few words and a wild card[1], for example the category:

```
<category>
<pattern>I NEED HELP *</pattern>
<template>Can you ask for help in the form of a question?</template>
</category>
```

8.2.3 Recursive

Recursive categories are those that "map" inputs to other inputs, either to simplify the language or to identify synonymous patterns. Many synonymous inputs have the same response. This is accomplished with the recursive <srai>tag. For example the input "GOODBYE" have dozens of synonyms: "BYE", "BYE BYE", "CYA", "GOOD BYE", and so on. To map these inputs to the same output for GOODBYE, we use categories like[1]:

```
<category>
<pattern>BYE BYE</pattern>
<template><srai>GOODBYE</srai></template>
</category>
```

References

- [1] Aimpl documentation. <http://www.alicebot.org/.../aiml-primer.html>.
- [2] The django project - documentation. <https://www.djangoproject.com/>.
- [3] Pyaiml documentation. <http://pyaiml.sourceforge.net/>.
- [4] Pygame documentation. <http://www.pygame.org/>.
- [5] Claudio G. Cortese. Learning through teaching. *Management Learning*, 36(1):87–115, 2005.
- [6] Fong-Ling Fu, Rong-Chang Su, and Sheng-Chin Yu. Egameflow: A scale to measure learners’ enjoyment of e-learning games. *Computers Education*, 52(1):101 – 112, 2009.
- [7] Ron Oliver and Western Australia. Factors impeding instructional design and the choice of learning designs in online courses, 2004.
- [8] Ron Oliver and Jan Herrington. *Teaching and learning online: A beginner’s guide to e-learning and e-teaching in higher education*. Edith Cowan University. Centre for Research in Information Technology and Communications, 2001.
- [9] Yin Pan, Sumita Mishra, Bo Yuan, Bill Stackpole, and David Schwartz. Game-based forensics course for first year students. In *Proceedings of the 13th Annual Conference on Information Technology Education*, SIGITE ’12, pages 13–18, New York, NY, USA, 2012. ACM.
- [10] Marc Prensky. Fun, play and games: What makes games engaging. *Digital game-based learning*, pages 1–31, 2001.
- [11] Bayan Abu Shawar and Eric Atwell. Chatbots: are they really useful? In *LDV Forum*, volume 22, pages 29–49, 2007.