

B.Sc. (Hons) in Software Development



Ollscoil
Teicneolaíochta
an Atlantaigh

Atlantic
Technological
University

Biometric Data Analysis in Digital Game Scenario

By
Rodrigo Almeida

March 20, 2024

Minor Dissertation

**Department of Computer Science & Applied Physics,
School of Science & Computing,
Atlantic Technological University (ATU), Galway.**

Contents

Contents	2
List of Figures	4
List of Tables	1
1 Abstract	2
2 Introduction	3
3 Methodology	6
3.1 Recruitment of Volunteers	6
3.2 Data Collection	7
3.2.1 Data Collection Procedure	10
3.3 Prediction Models	10
3.4 Software Development	11
3.5 Meetings	12
3.5.1 Supervisor Meetings	12
3.5.2 Client Meetings	13
3.6 Development Tools	13
3.7 Version Control	13
4 Technology Review	14
4.1 Executive Dashboard	14
4.1.1 Angular	14
4.1.2 Angular architecture	15
4.1.3 Advantages of using Angular	15
4.2 Data Visualization	17
4.2.1 Chart.js	17
4.2.2 D3.js	18
4.3 AI module - Data analysis	19
4.3.1 Hadoop	19
4.3.2 Apache Spark	20
4.3.3 TensorFlow	23
4.3.4 Node JS	24
4.3.5 PM2	26
4.3.6 Docker	27
4.3.7 NGINX	28
4.3.8 MySQL	29
4.3.9 Express JS	30

6	System Evaluation	36
6.1	Working with Tables	36
7	Conclusion	38
	Appendix	39
	Bibliography	61

List of Figures

3.1	Volunteers Recruitment Poster	7
3.2	Extreme programming methodology [1]	12
4.1	Angular application architecture	15
4.2	Apache Spark Ecosystem	21
4.3	Apache Spark Architecture	22
4.4	TensorFlow Architecture	24
4.5	Development workflow with Docker	28
5.1	System Architecture	33
5.2	Web Application Architecture	34

List of Tables

4.1	Chart.js Features	18
4.2	D3.js Features	19
4.3	Advantages and Disadvantages of Node.js	26
4.4	NGINX Features	29
4.5	MySQL	30
4.6	Express.js Features	31
6.1	Conversion from Hexadecimal to Binary	37

Chapter 1

Abstract

Chapter 2

Introduction

The collection and analysis of biometric data play a significant role in understanding the relationship between physical fitness and gaming performance. This research aims to investigate how various physical parameters, such as HRV (Heart Rate Variability), Heart Rate Max, Heart Rate average, Active Sleep and Quality of Sleep, influence a user's performance in a digital game, in this case PUBG: Battlegrounds. The research will also explore the potential of using biometric data to suggest the most suitable settings for different game scenarios. By using wearable technology, which can monitor metrics such as heart rate, sleep, and activity, this study seeks to establish a correlation between the player's physical condition and their performance in a digital game scenario.

The importance of fitness data in this context cannot be overstated. Heart Rate Variability (HRV), maximum heart rate, active sleep, and sleep quality are all essential indicators of a person's physical condition. These metrics offer a comprehensive insight into the user's physical condition, which can be used to improve their gaming performance.

Heart Rate Variability (HRV): is a measure of the variation in time between each heartbeat, and it is closely linked to the body's stress levels. HRV is widely used as a measure of the body's autonomic nervous system, which controls the body's stress response. A higher HRV is associated with a lower stress level, while a lower HRV is associated with a higher stress level. A research from E. Ortega[2] highlights the importance of HRV in sports science, as a way of understanding the psychological state of athletes before competitions. It found a positive correlation between HRV, self-efficacy, and performance among sport shooters. Advanced shooters demonstrated lower average heart rates and employed mental skills more effectively than less experienced shooters. This suggests HRV as a valuable asset, when transferring the psychological state of the athlete to the performance in a digital game. It allows for a better understanding of the player's stress levels, and how it affects their performance. This personalized approach can enhance

player experience, and potentially improve their performance in the game.

The Heart Rate Variability measurements typically requires a chest strap, and measurements typically span from 5 minutes to 24 hours, it is commonly used in clinical settings to evaluate cardiac conditions.[3] Short-term HRV analysis lasting less than 5 minutes have been proven to provide more accurate estimations compared to longer measurements.[4] Based on this, the research will focus on short-term HRV measurements, as it is more practical for the user, and it provides accurate estimations. It will be more detailed in the methodology section.

Heart Rate Max: is the maximum number of times the heart can beat in a minute, and it is a measure of the body's cardiovascular fitness. It is an important indicator of the body's physical condition, and it is used to evaluate the body's ability to perform physical activities, reflecting their cardiovascular fitness. In this research, understanding HRmax is vital for assessing participants fitness and endurance. It could indicate how well players can handle stress and maintain their concentration over extended periods of time. By analysing alongside other biometric data this research could uncover valuable insight into optimizing players performance in a digital game scenario.

Heart Rate Average: is the average number of times the heart beats in a minute, and it is a measure of the body's physical activity. It is also an important indicator of the body's physical condition, and it is another measure of the body's cardiovascular fitness. It is used to evaluate the body's ability to perform physical activities, reflecting their cardiovascular fitness.

Active Step: Active steps, as a measure of physical activity, is an important indicator of the body's physical condition. It is used to evaluate the body's ability to perform physical activities, reflecting their cardiovascular fitness. Counting the numbers of steps taken daily can help track overall physical activity, and it is a good indicator of the body physical condition. It is also used to evaluate the body's ability to perform physical activities, reflecting their cardiovascular fitness.

Quality of Sleep: are important indicators of the body's recovery and readiness for physical activity. Active sleep is a measure of the body's physical activity during sleep, and it is an important indicator of the body's recovery and readiness for physical activity. High-quality sleep, marked by sufficient duration and minimal disruptions, is fundamental for overall health impacting mood, cognitive function. A research that evaluated Sleep and performance in Eathletes [5] underscores the critical role of sleep in Eathletes performance in Esports, showing its impact on cognitive functions as a crucial factor for competitiveness success. Adequate sleep improves information processing, visual motor functioning, attention, working memory, and other functions essential for decision-making and reaction time. On the other hand, sleep deprivation can significantly impair these cognitive abilities, potentially to poor performance. For this research it proves the

importance of sleep in the context of gaming performance, and how it can be used to predict the player's readiness and tune their performance in a digital game scenario.

Chapter 3

Methodology

This research is geared towards finding a correlation between users' biometric data to their performance in a first-person shooter gaming scenario. The quality of the user data collected was critical in ensuring that the research aims are achieved. For Quality Assurance purposes, a systematic step-by-step approach has been designed for the data-capturing process. An activity monitoring device was distributed to the volunteers in the form of a Polar Watch. The individuals were instructed to use these devices at pre-designated times prior to undertaking a test in the form of a First-Person Shooter Game. Results from the test, was than paired with their biometric data for further analysis.

3.1 Recruitment of Volunteers

The process of recruiting the volunteers was carried out in a way that ensured they satisfied the conditions of the research. The volunteers were required to be students of the Atlantic Technological University (ATU) Galway campus. The volunteers were also required to be over the age of 18 and have no known medical conditions that could affect their performance in the test. A microsoft form was created and distributed to the students of the university to register their interest in the research. A poster 3.1 was created wich contained information about the research and a QR code that linked to the microsoft form.

The form can be found in the appendix 7. For a volunteer to be considered, they were required to fill out the form and submit it. The form was designed to capture the following information from the volunteers:

- How often the participant practices exercises.
- The average duration of the exercises sessions.



Figure 3.1: Volunteers Recruitment Poster

- If the participant owns a activity monitoring device, if so which brand (Polar, Garmin, Apple, Fitbit, or other).
- If the participant plays video games, if so which platform (PC, Xbox, Playstation, or other).
- Which game genre the participant enjoy the most. (First Person Shooter, Third Person Shooter, Soccer, Car Racing, or other).
- Participant should also sign the form as a consent to participate in the research.

3.2 Data Collection

For the purpose of the research, different categories of data was collected from the volunteers through the monitoring device to help achieve the stated goals. As personal data was being collected, an application to the Ethicas Research and Ethics Committee was made to ensure that the data collection process was in compliance with the General Data Protection.

The Taught Programme Research Ethics Approval Application Form can be found on appendix 7. It was an extensive and iterative process with the Ethicas Research and Ethics office to ensure that the data collection process was in compliance with the General Data Protection Regulation (GDPR). Unfortunately, it took

longer than expected to get the approval from the Ethics Research and Ethics Committee, which lead to an unexpected delay in the data collection process. The data collection process was divided into two main categories:

Biometric Data

For the very first time processing a volunteer, some user information was needed to have them registered with the activity monitoring device and their data was saved to the manufacturer's repository. The Activity Monitoring Device used on this research was the Polar Vantage Smart Watch. The watch is capable of monitoring and capturing user's biometric data and saving the data to a repository where it can be accessed online via an API. The device is commercially available and accessible to the public. The rationale for using the Polar Vantage Watch is because it has been widely used in both academic and industrial research projects and validation. Most importantly, these devices were available to us in such quantity that could satisfy our research needs.

Physical Information

The following physical information was collected from the volunteers, and stored with the manufacturer (Polar) which is GDPR compliant. The data was then accessed through the Polar API.

- **Sleep Data:** According to the manufacturer's manual [6], the watch is capable of recording the quality and quantity of sleep and also showing how long spent on each stage of sleep. Light Sleep, Deep Sleep, and REM Sleep with the duration of each type of sleep. For the purpose of the research, an aggregated total of the various categories of sleep will be recorded and used. Volunteers were instructed to wear the watch to sleep on the previous night before their scheduled test. Unit of measure used for the sleep data is minutes.
- **Daily Activities:** According to the Activity monitor manufacturers' manual [6], "Polar device uses an internal 3D accelerometer to record your wrist movements. It analyses the frequency, intensity and regularity of your movements together with your physical information." calories, active calories, active steps and their respective durations are collected from through the device. For the purpose of the research, the active steps was used. Volunteers were expected to wear the device on the previous day for this data to be available. Unit of measure used for the daily activities is count of steps.
- **Nightly Recharge:** From the Activity monitor manufacturers' manual [6], the Nightly Recharge is recorded as follows: " is an overnight recovery measurement that shows how well your body has coped with overall stress you

have experienced lately." The parameters, measured during roughly the first four hours of your sleep are heart rate, heart rate variability and breathing rate. For the purpose of the research, the following parameters and derived parameters will be used:

- Heart Rate Average (bpm)
- Heart Rate Maximum (bpm)
- Heart Rate Variability (HRV)

Test Data Collection

The test data is generated at the completion of the test session in the Unit Test Application. A test session consists of three categories of tests, which is undertaken sequentially in different stages. For quality assurance, the test was performed in a controlled environment using the same hardware in similar conditions over the course of the trials. For the purpose of the trial, a special designated room has been reserved for the data collection effort. The following metrics for the different tests was then captured and subsequently used for the research effort.

- Audio Test: designed to test users' audio reflexes.
- Average Response Time: this is a measure of how quickly a user is able to react and identify the source of a sound with varying audibility (in decibels) in a 3-Dimensional space.
- Visual Test: designed to test users' visual perception. At the end of this category of test, the following metrics were designed to measure users' performance.
- Average Response Time: a measure of how quickly a user can identify and engage targets.
- Shot Accuracy: a measure of the percentage of successful shots to the number of targets spawned.
- Target Accuracy: a measure of the percentage of successful shots taken to a number of targets hit.
- Fine Motor Test: designed to test users' perception of depth and eye to hand coordination.
- Average Tracking Time: measure of average time it took for a player to successfully track, engage, and eliminate a target.

- Accuracy: measure of the percentage of shots fired to the number of targets hit.

3.2.1 Data Collection Procedure

The data collection process followed the below-listed steps to maximize throughput and minimize the average time spent processing each participant. The whole procedure can be found in the appendix 7.

For this study, participants were required to wear the Polar watch which is referred sometimes to “Activity Monitoring Device” in this text, a day prior to undertaking a gagging test.

Location:

The data collection effort was carried out at the Gym on the Atlantic Technological University (ATU) Galway campus.

Time:

The following weekly schedule was available for the volunteers at various times that best suit their personal schedules.

- Tuesdays: 12:00 - 13:00, 15:00 - 16:00
- Wednesdays: 12:00 - 13:00, 15:00 - 16:00
- Thursdays: 12:00 - 13:00, 15:00 - 16:00
- Fridays: 13:00 - 15:00

Results of the test were then paired with their biometric data for further analysis.

3.3 Prediction Models

- inherited a infrastructure - adapted the infra to our needs - enrolled volunteers - collected data - designed a statistical model to analyze the data -

3.4 Software Development

As a software development methodology for this project, many factors were considered. Initially, a comparison of the different software development methodologies was carried out to determine which would be the most suitable for the project. The comparison was based on the following factors:

- **Continuous Stakeholder Involvement:** The methodology should allow for continuous stakeholder involvement in the development process. Maintaining a continuous feedback loop with the stakeholders was important to ensure that the project was on track and that the stakeholders were satisfied with the progress.
- **Flexibility in Requirements:** The methodology should allow for flexibility in the requirements. The requirements for the project were not fully defined at the start of the project. The methodology should allow for changes to be made to the requirements as the project progresses.
- **Iterative Development:** The methodology should allow for iterative development. The project was divided into different phases and each phase was developed iteratively. This allowed for the project to be developed in a series of small, manageable steps.

After all the considerations, Agile Methodology [1] appeared as the preferred choice over Waterfall, as it allows for continuous stakeholder involvement, flexibility in requirements and iterative development. Unlike the rigid Waterfall methodology [1], Agile allows for changes to be made to the requirements as the project progresses. The Agile methodology also allows for the project to be developed in a series of small, manageable steps. This iterative nature enables continuous refinement and adjustment, ensuring that the project is always on track and that the stakeholders are satisfied with the progress.

Agile - Extreme Programming (XP)

Extreme Programming (XP) is an Agile framework that moves from the traditional software development methodologies by breaking down the development process into smaller and more manageable increments. Instead of extensive planning, analysis, and design upfront, XP focuses on continuous and iterative development. To minimize errors and defects, XP places a strong emphasis on unit testing. This test-driven development approach ensures that the code is always working and that any defects are caught early in the development process. XP also places a strong emphasis on customer satisfaction. The customer is involved in the development process and is able to provide feedback on the project as it progresses.

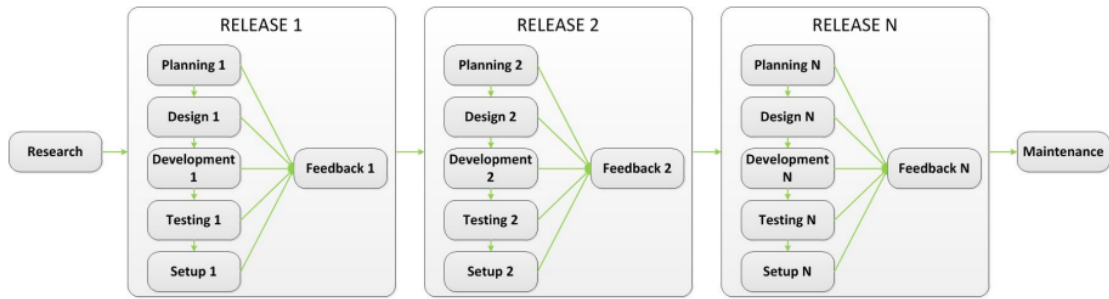


Figure 3.2: Extreme programming methodology [1]

This ensures that the project is always on track and that the customer is satisfied with the progress. [1] Overall, Extreme Programming offers a flexible and iterative approach to software development that is well-suited to the needs of this project.

3.5 Meetings

In this section we will discuss the different types of meetings that were held throughout the project. The meetings were held to ensure that the project was on track and that the stakeholders were satisfied with the progress.

3.5.1 Supervisor Meetings

Weekly meeting were setup with the supervisor to discuss the progress of the project. The meetings were held in the supervisor's office on campus every Tuesday at 12 pm throughout the duration of the project. The meetings were used to discuss:

- The progress of the project.
- Any issues or challenges that were encountered.
- Any changes that needed to be made to the project.
- Any feedback that the supervisor had on the project.

These meetings were important, specially in the early stages of the project, due to various challenges that were encountered. The feedback from the supervisor was invaluable in helping to overcome these challenges and ensure that the project was on track.

3.5.2 Client Meetings

At the beginning of the year we were offered the opportunity to work with a client on a project. In our meeting with the client, we were informed that the project was a research project that aimed to find a correlation between users' biometric data to their performance in a first-person shooter gaming scenario. The aim was to quantify how biometric such as Heart Rate Variation, Heart Rate, active steps taken, quality and quantity of sleep, etc. affect select gaming skills like eye-to-hand coordination, fine-motor skills and reaction time. The client also informed us that the project was a continuation of a previous project and that we would be inheriting an existing infrastructure. For the collection of the biometric data, the client informed us that there was four Polar Vantage Smart Watches available for the project. The meetings were also held to get the devices registered on the Polar API and to get the volunteers enrolled in the project.

The meetings were held in the client's office on the Gym at Unit D, Racecourse Business Park, Ballybrit, Galway.

3.6 Development Tools

3.7 Version Control

Chapter 4

Technology Review

4.1 Executive Dashboard

When it came to developing the Executive Dashboard for data analysis it was crucial to select the tools that could transform data into easily comprehensible information. After conducting research and exploring alternatives Angular and Chart.js were chosen for specific reasons.

4.1.1 Angular

Angular is based on TypeScript and often used for building single-page applications (SPAs) or enterprise-level solutions. It is a powerful tool for building client-side web applications and is known for its component-based architecture.

Main Concepts: Components, Modules and Services

Angular is composed of three foundational blocks: Components, Modules and services.

Components

Components are like Lego blocks of the application. Each component holds a portion of the user interface and its behaviour. Components serve as the bridge between the application data and what the user experiences on the screen.[7]

Modules

In Angular modules serve as containers that group related components, directives, and services together that can be combined with other modules. It plays an im-

portant role in improving maintainability and re-usability, key concepts of Angular development.[8]

Services

Angular services use typescript classes with injectable decorators. The decorator informs Angular that the class function as a service and can be injected into other components that need that service.[9]

4.1.2 Angular architecture

Model-view-controller is a pattern that separates the application into three main components: the model, the view, and the controller. The controller is responsible for the interaction between the model and the view.

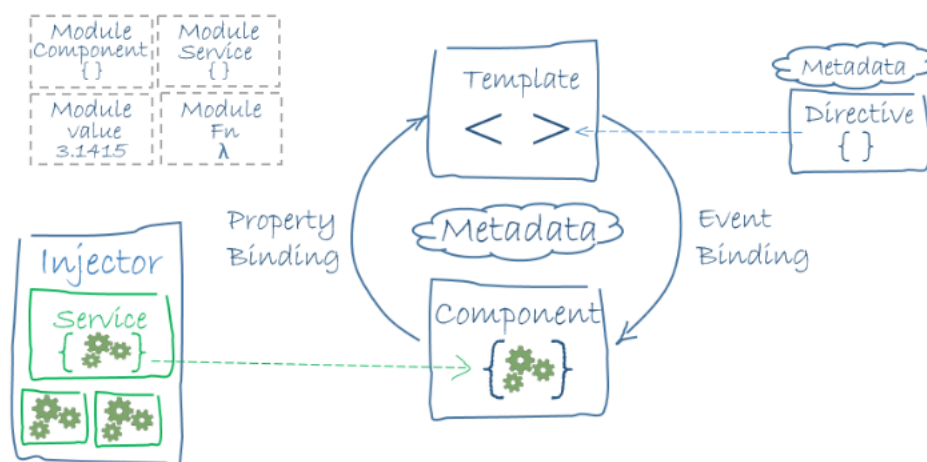


Figure 4.1: Angular application architecture

4.1.3 Advantages of using Angular

Component-Based Architecture

As mentioned earlier in discussions Angular organises its functionalities into components. These components have the ability to communicate with each other enabling updates to sections without affecting the rest of the application.

Mobile-Friendly Approach

Angular incorporates techniques such, as lazy-loading, which means loading parts of the application (like images) only when they are needed. This ensures that users do not experience long waiting times.

Two-Way Data Binding

With Angular two way data binding data can seamlessly flow between the component and the view allowing for synchronization.

Asynchronous Programming

By utilizing programming executes code in a non-sequential manner and employs multi-threading to enhance performance. This speeds up operations and prevents system freezes, providing users with a seamless experience.

Single-Page Applications

Angular creates a dynamic single-page application which can be navigated without page reloads, improving the user experience with better user interaction and engagement.

Code Re-usability

The component-based architecture of Angular promotes the re-usability of UI components saving development time.

Dependency Injection

With dependency injection in place, Angular allows for the creation of objects that rely on other objects. This improves modularity and efficiency, within the app.

Angular Material

Angular's documentation offers a range of built user interface components and modules that adhere to Google's Material Design principles. This greatly facilitates the developer's work, simplifying the design process and enabling application development.

Angular CLI

Angular command line interface gives the developer the ability to generate Angular projects, modules, services, and components with a single command,

which helps reducing configuration errors, and gives the developer the freedom to dive into creative aspects of the project, focusing on innovation and functionality rather than getting bogged down by initial setup complexities.

4.2 Data Visualization

Data visualization involves the creation of visual representations of data to communicate information clearly and efficiently. This technique is used to identify patterns, trends, and correlations that would be difficult to identify in raw data. It is an important step in the data analysis process and could be used to communicate information clearly and efficiently.

4.2.1 Chart.js

Chart.js is a powerful, efficient, and open-source JavaScript library that offers capabilities for generating attractive charts and graphs, which can be easily integrated into web pages. It stands out for its user-friendly nature, making it an excellent choice for data visualization tasks. Utilizing HTML5 canvas technology, it ensures that charts are responsive and adjust seamlessly to the container's dimensions. Additionally, it boasts compatibility across all contemporary web browsers and enjoys support from an extensive developer community. Chart.js acts as an artist's palette for developers, equipped with a versatile API that supports an array of chart types, including line, radar, and pie charts, among others. Moreover, it provides extensive customization possibilities, enabling the creation of distinct and eye-catching charts.[10]

Table 4.2.1 is showing some of the features of Chart.js[10]:

Feature	Description
Easy to use	Aesthetically charts can be created without the need of extensive configuration. The library follows a declarative approach allowing the developer to define the data and settings of the chart in a single object.
Responsive	Charts generated by Chart.js are responsive, adapting to different devices, and makes sure that the visualization is readable across various platforms.
Customization	Chart.js provides a high degree of customization. Colors, fonts, and other visual elements can be changed creating unique charts.
Interactivity	Chart.js has great support tooltips and animation, enabling the user to explore the data points, adding a layer of engagement to the project.
Cross-browser compatibility	Chart.js is compatible with all most of the browsers.

Table 4.1: Chart.js Features

Chart.js stands as a dependable and powerful solution for visualizing data, characterized by its simplicity and a wide array of customization features. Additionally, it benefits from the strong support of an extensive developer community, offering a significant advantage.

4.2.2 D3.js

D3.js or Data-Driven Documents has the ability to bind data to the DOM (Document Object Model) and creates interactive and dynamic visualization. It provides a lower level and more granular approach to data visualization, giving the developer more control over the visualization process.

Table 4.2.2 is showing some of the features of D3.js[11]:

Feature	Description
Data-Driven	D3.js is data-driven, meaning that it can bind data directly to HTML or SVG elements. This type of control is advantageous for creating highly customized visualizations.
Modular	D3.js is modular, composed of many small modules that can be used independently or combined together to create a custom visualization.
Extensible	D3.js is extensible, meaning that it can be extended to support any type of visualization.
Community	D3.js is supported by a large community of developers, which is a great advantage.

Table 4.2: D3.js Features

D3.js is a great tool for visualizing data. It provides a high degree of customization and control over the visualization process. However, it is more complex and requires more time to learn and implement. It is also not supported by all browsers, which is a disadvantage.

In conclusion, while D3.js is a great tool for visualization of data, it is more complex and requires more time to learn and implement. It is also not supported by all browsers, which is a disadvantage. Considering the scope of the project and the time constraints, Chart.js was chosen as the tool to visualize data, as it is easy to use and provides a high degree of customization.

4.3 AI module - Data analysis

An amazing fact is that ninety percent of the data generated in the world was generated in the last two years. In the early 2000s, the amount of data being generated exploded exponentially on the same rates as the internet and social media usage. Organizations found themselves facing a massive volume of data that was very hard to process. Then the concept of Big data was created to describe this large volume of data. It refers to data that is so large and complex that traditional methods of data processing are not sufficient to process it. [12]

4.3.1 Hadoop

As the volume of data grew rapidly across the globe, organizations needed a way to manage it to gain valuable insights. That's where Hadoop came in. In 2006, a team of engineers from Yahoo developed Hadoop inspired by Google's MapReduce. Hadoop has introduced a new way of handling data called distributed processing.

Now, to process all data it wouldn't use single machines anymore, instead, Hadoop uses multiple computers, allowing large amounts of data to be processed way faster than before.

Hadoop primarily consists of two core elements: HDFS (Hadoop Distributed File System) and MapReduce. HDFS distributes the data across various computers, while MapReduce handles parallel data processing, enabling organizations to store and analyze vast volumes of data efficiently.

Hadoop was a great advance in big data processing, however, it had a few limitations. One of the biggest problems was that it relied on storing data on disk. This slowed down data processing because every time a job ran it had to save its data to disk, read it back, process it, and save it back to disk. Another problem is that Hadoop processes data only in batches. This means a new job couldn't be submitted before the other job ended. With Hadoop, storing and processing large amounts of data became possible, even with some limitations it was an important step in big data processing.[13]

4.3.2 Apache Spark

As described Hadoop has a few limitations, there was a need to process all this data faster and in real-time. And here is where Apache Spark comes into action. In 2009, researchers at the University of California developed Apache Spark as a research project. At this point, RDD (Resilient Distributed Dataset) was introduced and deemed to be a very powerful concept.[14] RDD is the backbone of Apache Spark, storing the data in memory for faster access and processing. It will not read and write the data from the disk, instead, Spark processes the entire data just in memory. The meaning of memory here is the RAM (Random Access Memory) stored inside the computer. What makes Spark so powerful is that it can process data in real-time, unlike Hadoop, which processes data only in batches.

Spark also gives the ability to write code in various programming languages such as JAVA, Scala, Python and R.

Main components and Features

The Apache Spark ecosystem includes the below components[15]:

- Spark SQL: Previously referred to as Shark, Spark SQL operates as a distributed framework designed to handle both structured and semi-structured data. It supports analytical and interactive applications for streaming as well as historical data, offering compatibility with various data sources including JSON, Parquet, and Hive tables.cite8988541

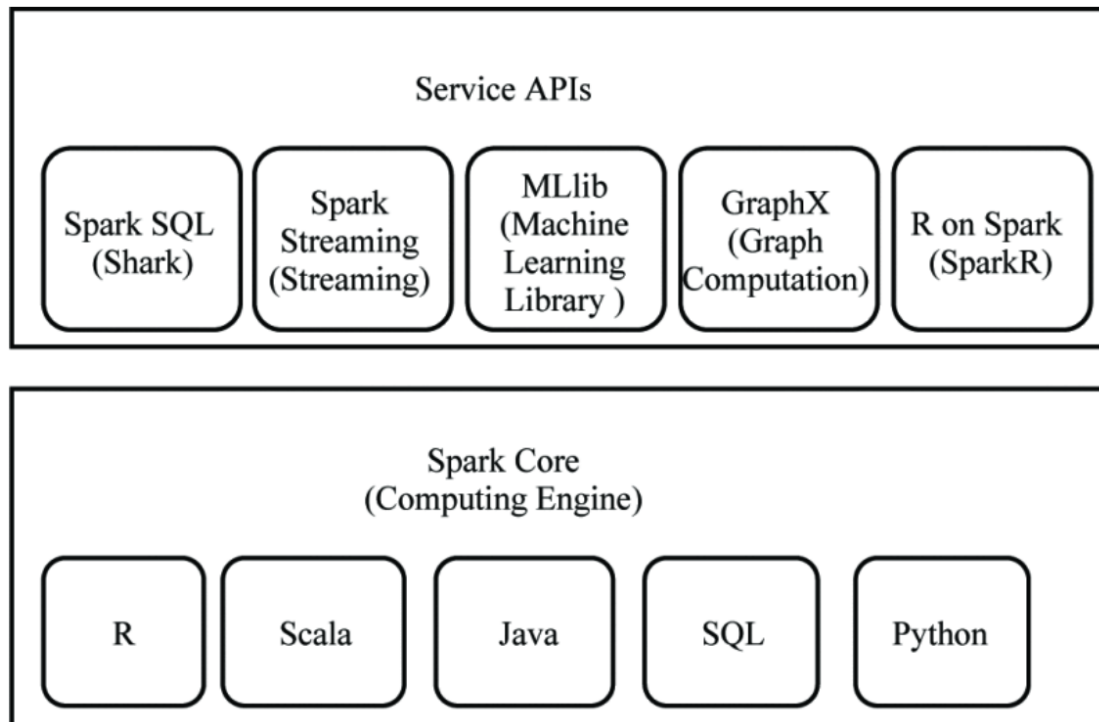


Figure 4.2: Apache Spark Ecosystem

- **Spark Streaming:** This component enables the processing of live data streams. As a scalable and fault-tolerant system, Spark Streaming accommodates both batch and streaming data workloads. It leverages Apache Spark's rapid scheduling capabilities by organizing incoming data into small batches for processing, which are then transformed and can be sourced from live feeds or data sources like Twitter, Apache Kafka, IoT sensors, and Amazon Kinesis.[15]
- **Spark Core:** Serving as the foundational layer of the Spark ecosystem, Spark Core is crucial for distributing data processing across multiple nodes, ensuring both efficiency and reliability. The core functionalities of Apache Spark, including a comprehensive API set and programming support for Scala, Java, and Python, are built atop this layer. Spark Core employs in-memory computing techniques to enhance performance and address the limitations associated with the MapReduce model.cite8988541
- **MLlib:** This library is focused on making machine learning scalable and accessible. It encompasses a wide range of machine learning algorithms, includ-

ing those for regression, classification, clustering, and linear algebra. MLlib not only offers advanced algorithms but also foundational machine learning primitives like the generic gradient descent optimization algorithm, alongside utilities for model evaluation and data handling. It supports development in Java, Scala, and Python. [15]

With all these components working together, Apache Spark became a powerful tool for processing and analysing Big Data. The coordination managed by Spark across the cluster of computers is what makes it so powerful.

Spark manages and coordinates the execution of tasks on data across a cluster of computers.

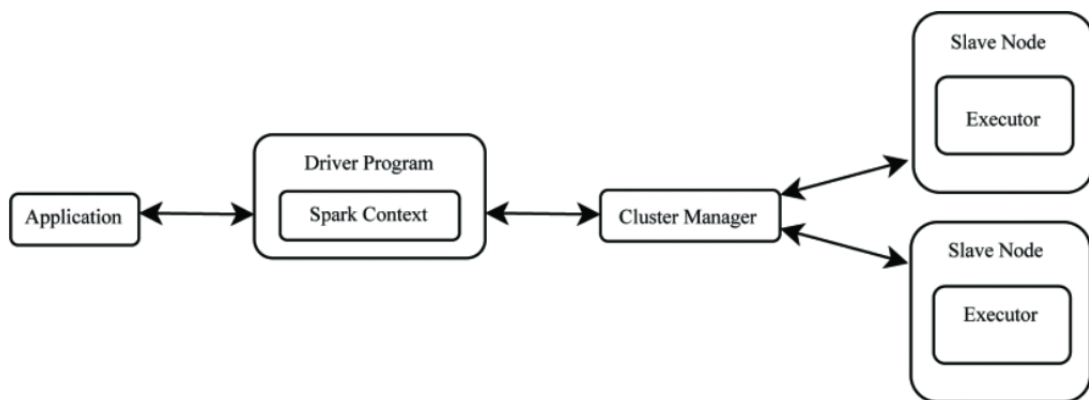


Figure 4.3: Apache Spark Architecture

Figure 4.3 shows the Spark Architecture. The master node is responsible for managing the cluster and distributing the work across the worker nodes. It is the entry point for the cluster and is responsible for allocating resources to the applications, allowing for communication and coordination with other nodes within the cluster, known as Slave nodes. Within the slave node, there are many numbers of executors which act as workers. Whenever a job is initiated, the driver process will make sure it goes through the Apache application properly. It analyzes the work that needs to be done, divides it into smaller tasks and assigns it to the executors. The driver process is the heart of the Apache Spark application and executors are the real workers, which execute the work assigned by the driver and report back the progress and results of the computation.[15]

4.3.3 TensorFlow

Google brain team developed TensorFlow, which is an open-source library for numerical computation and large-scale machine learning. It is not only powerful but also flexible, allowing for the creation of custom machine learning models. It is also highly scalable, allowing for the training of models on multiple CPUs, GPUs, and TPUs (Tensor Processing Units).[16]

- **Flexible Architecture:**
 - TensorFlow provides developers with the capability to execute computations across multiple CPUs or GPUs using a unified API, whether on desktops, servers, or mobile devices, thereby making it versatile for various applications.
- **Comprehensive Library:**
 - This platform offers an extensive array of tools, libraries, and community support, enabling researchers to push forward the cutting-edge in machine learning (ML), while also allowing developers to effortlessly create and implement applications powered by ML.
- **High-Level APIs:** With high-level APIs like Keras, TensorFlow is very accessible for beginners by simplifying how to build and train models. It also provides a low-level API for more advanced users, allowing for greater flexibility and customization.
- **Visualization with TensorBoard:** As part of TensorFlow, there is a tool called TensorBoard, which is a suite of visualization tools that makes it easier to understand, debug, and optimize complex neural networks. It simplifies the process of understanding, debugging, and optimizing complex neural networks.
- **Scalability:** TensorFlow's ability to perform computations on multiple CPUs or GPUs, making it fit for large-scale machine learning tasks.
- **Large Community and Support:** Having a vast community, TensorFlow benefits from a plethora of tutorials, documentation, and active community support which aids in solving problems and improving the framework.

Figure 4.4 [17] shows the TensorFlow Architecture. The architecture is designed to be flexible and scalable, allowing for computations on a variety of devices and the ability to handle large-scale machine learning projects.

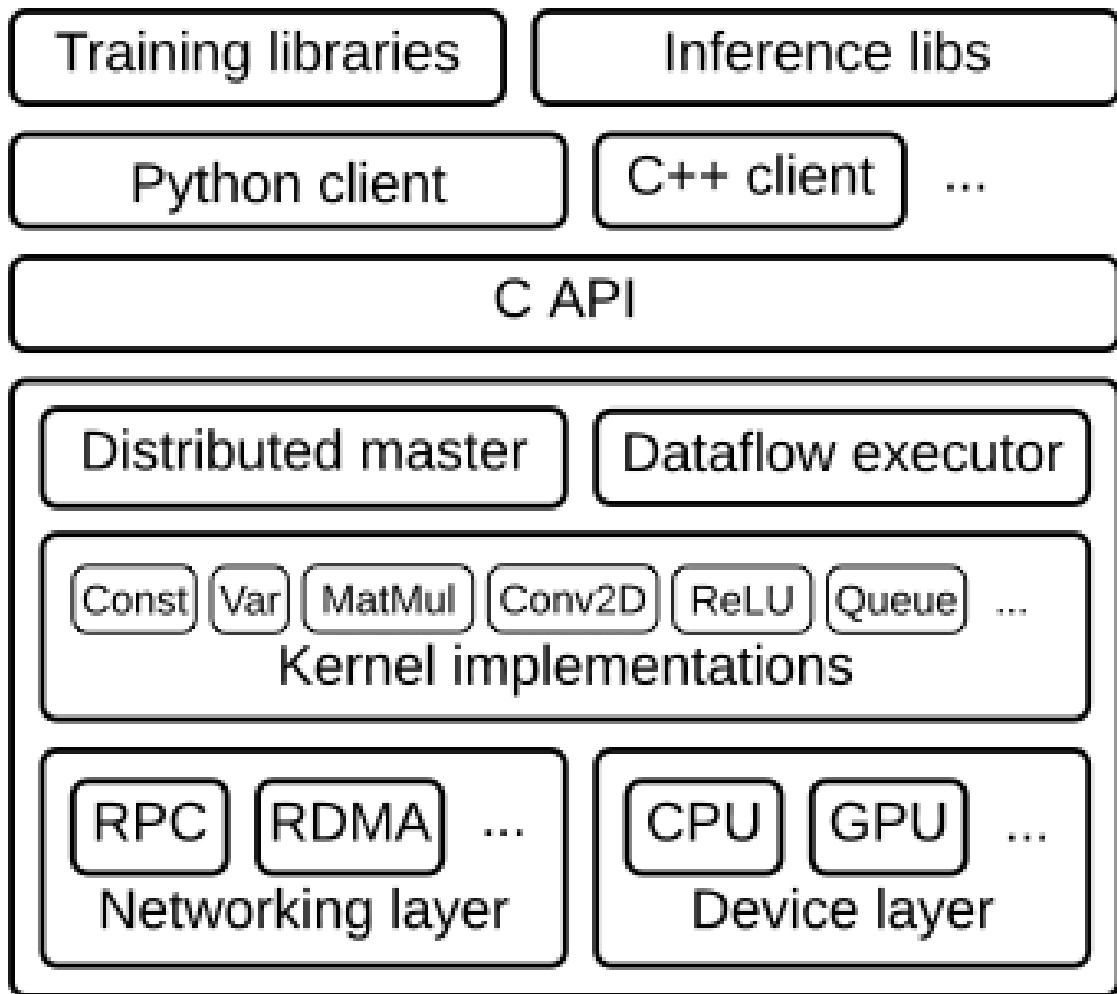


Figure 4.4: TensorFlow Architecture

Given these powerful features and advantages, TensorFlow was chosen to be used in conjunction with Apache Spark, for this project to handle complex machine learning tasks, including but not limited to, neural network design and training. As the dataset grows, TensorFlow's performance scales with the size of the data, processing with high speed.

4.3.4 Node JS

NodeJS, a cross-platform, open-source JavaScript runtime, enables the execution of JavaScript code beyond the confines of a web browser. Constructed atop Chrome's V8 JavaScript engine, it facilitates the development of swift and scalable network

applications. With its event-driven and non-blocking architecture, NodeJS efficiently manages multiple concurrent operations, preventing the halt of other processes. This characteristic renders it perfectly suited for real-time applications that are data-heavy and operate on a distributed network of devices.[18]

Advantages of Node.js

NodeJS is designed as an asynchronous, event-driven JavaScript runtime ideal for crafting scalable network applications. Its efficiency and lightweight nature make it an excellent option for data-heavy, real-time applications.[19] A significant benefit of NodeJS is its use of a single-threaded event loop, enabling it to manage multiple operations concurrently without hindering the execution of additional tasks.

A key feature of NodeJS is its extensive library ecosystem. With the Node Package Manager (NPM), it boasts the world's largest collection of open-source libraries, hosting over a million packages.[20] This abundance of resources significantly streamlines the development process, allowing developers to leverage existing modules and focus on innovation.

NodeJS also benefits from its single programming language across both the server and client sides. This allows for the sharing of code and data between the server and the client, which is a great advantage, reducing the learning curve associated with NodeJS development. [19] Furthermore, NodeJS's uniform programming language for both server and client sides simplifies code and data sharing across them, offering a streamlined learning process for development. In terms of performance, NodeJS excels due to its foundation on Chrome's V8 JavaScript engine, a high-performance engine that ensures speed and efficiency, especially in scenarios requiring real-time data processing. [19]

Disadvantages of Node.js

NodeJS is not without its criticisms. Critics often point out the callback hell, a situation where the code becomes unreadable due to the excessive use of callbacks. Although it has been largely mitigated by the introduction of Promises and `async/await` syntax, it remains a valid criticism. [21]

Table 4.3.4 is summarizing the advantages and disadvantages of Node.js[19]:

Advantages of Node.js	Disadvantages of Node.js
Non-blocking and event-driven: Allows for high concurrency and scalability, making it suitable for real-time applications.	Single-threaded: it is single-threaded, and can lead to blocking if not handled correctly, impacting CPU-bound tasks.
JavaScript as a single language: Developers can use JavaScript for both server-side and client-side development, reducing context switching.	Callback hell: the code could be hard to read and maintain as there are nested callbacks (Promises and async/await can minimize this problem).
Large and active community: The Node Package Manager makes thousands of open-source libraries and modules available.	Limited support for multi-core processors: Node.js does not fully utilize multi-core CPUs out of the box.
Speed: it is known by its high-performance Node.js is built on the V8 JavaScript engine.	Less suitable for CPU-intensive tasks: The event-driven and single-threaded characteristics, it could not be the best option for CPU-bound operations.
Lightweight and fast startup: Node.js applications typically have lower memory consumption and quicker startup times.	Maturity and stability: Some developers argue that Node.js, compared to more established platforms, may have less maturity and stability in certain use cases.

Table 4.3: Advantages and Disadvantages of Node.js

4.3.5 PM2

PM2, or Process Manager 2, is a production process manager for NodeJS applications. With a rich feature set, including monitoring, load balancing, and error handling. It is designed to simplify the deployment and management of NodeJS applications in production environments with the ability to have the application running forever, restarting it with no downtime. The standout feature of PM2 is its ability to keep processes running in the background indefinitely. This is particularly important for web applications that requires constant availability. PM2 automatically resurrects crashed applications, ensuring that system glitches do not result in prolonged downtime.[22] This automatic restart capability is a safeguard against potentially costly application crashes that could otherwise lead to a poor user experience or even lost revenue. [22] PM2 also provides a built-in load balancer that can distribute incoming requests across multiple instances of the

application. This improves performance and scalability, allowing the application to handle more requests, evenly distributing traffic across the instances, which can be particularly beneficial when running on multi-core systems. [23] However, PM2 is not one-size-fits-all solution. It is designed for NodeJS applications, and is not suitable for applications written in other languages. It is also not suitable for applications that require a high degree of customization. [22] Despite the considerations, PM2's benefits have outweighed its limitations, and it was chosen as the technology for the project. It provides many features, including monitoring, load balancing, and error handling. It is designed to simplify the deployment and management of NodeJS applications in production environments and to keep applications alive forever, restart them without downtime and simplify common system administration tasks. Its installation and setup process is straightforward, requiring minimal configuration. All these features make PM2 an ideal choice for the project, as it provides the necessary tools to manage and monitor the NodeJS application in a production environment.[19]

4.3.6 Docker

Docker has emerged as a revolutionary tool for software development. It provides a platform for developers to build, ship, and run applications in containers. Comparing with virtual machines (VMs), containers are lightweight, portable, and efficient, making them an ideal solution for deploying applications across different environments, allowing for the isolation of applications and their dependencies into containers. With the aim to encapsulate the application's code, libraries, and dependencies in a single object, Docker ensures that the application runs reliably and consistently across different environments. Due to its simplicity and portability, Docker has become the de facto standard for containerization. It is supported by all major cloud providers, including Amazon Web Services, Microsoft Azure, and Google Cloud Platform. Its lightweight nature compared to traditional virtual machines makes it ideal for cloud computing. Containers share the machine's OS kernel, and do not require an OS per application, which greatly reduces the memory footprint and improves performance. [24]

However, Docker's container model is not without challenges. In the context of this project, the average size of a single Docker image exceeding 400MB presents a concern regarding bandwidth utilization. Continuous build and deployment processes involving large container images can consume substantial network resources, leading to potential cost overruns. [24] Another challenge is the security of Docker containers. Docker containers share the same kernel, which means that a vulnerability in the kernel can affect all containers. [25] While Docker's layering and image caching mechanisms can mitigate some of the network overhead, it is still a concern. In conclusion, Docker is a powerful tool that offers numerous ben-

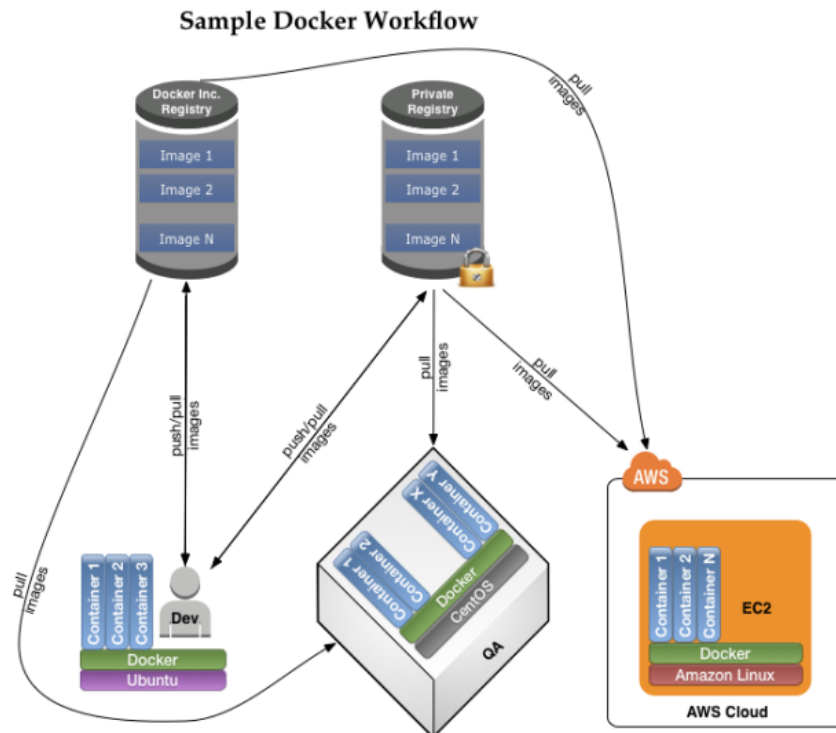


Figure 4.5: Development workflow with Docker

efits for application deployment, from development to production. However, it is not without its challenges. The large size of Docker images can lead to bandwidth utilization issues, and the shared kernel model can lead to security concerns. Considering the scope of the project, time constraints and budget, Docker wasn't chosen as the technology for the project as the cost would be too high.

4.3.7 NGINX

NGINX works as a reverse proxy server, load balancer, and HTTP cache. It is a high-performance web server that can handle high volumes of concurrent connections.

It is a lightweight and efficient solution that can handle high volumes of concurrent connections. It is also highly configurable, allowing for the customization of its behaviour to suit the needs of the application. It was originally created by Igor Sysoev in 2004, with the goal of solving the C10K problem, which refers to the challenge of handling 10,000 concurrent connections. [26] Due to its vast capabilities, it has become one of the most widely used web servers in the world, often used as a load balancer and as a reverse proxy in addition to its role as a

web server.

Table 4.3.7 is showing some of the features of NGINX[26]:

Aspect	Description
Type	Web server software
Primary Use	Serving web content, load balancing, reverse proxy, and more
Performance	Highly efficient in handling high concurrency with low memory footprint
Architecture	Event-driven, asynchronous, and non-blocking which contributes to its ability to handle a large number of simultaneous connections easily
Scalability	Scalable to support growth in traffic and applications
Security Features	Offers robust security features including rate limiting, client request filtering, and SSL/TLS termination
Flexibility	Highly configurable for a wide range of web and mail server tasks
Open Source/Commercial	Available in both open-source and commercial versions (NGINX Plus)

Table 4.4: NGINX Features

Due to its vast capabilities NGNIX was chosen for our project as reverse proxy and load balancer. It is a lightweight and efficient solution that can handle high volumes of concurrent connections. It is also highly configurable, allowing for the customization of its behaviour to suit the needs of the application.

4.3.8 MySQL

MySQL uses structured query language (SQL) for database access, providing high performance and reliability. It uses structured query language for database interaction.

Table 4.3.8 is showing some of the features of MySQL[27]:

Aspect	Description
Data Handling	Efficiently manages large datasets, crucial for training machine learning models
Query Performance	Fast query execution, beneficial for data retrieval and preprocessing in machine learning workflows
Scalability	Easily scales with data volume and complexity, supporting the growing needs of machine learning applications
ACID Compliance	Ensures data integrity and consistency, vital for the accuracy of machine learning outputs
Advanced Analytics	Supports SQL extensions for advanced analytics, facilitating machine learning data processing tasks
Data Storage Options	Offers various storage engines, allowing optimization based on the specific needs of machine learning models
Community and Support	Strong community and extensive documentation, aiding in troubleshooting and optimization for machine learning projects
Integration Capabilities	Easily integrates with popular machine learning frameworks and languages, streamlining the development process

Table 4.5: MySQL

In the context of this project, the dataset was originally in a JSON format in Firestore which were found to be of a high complexity while performing compound querying. To facilitate querying and the difficulties of Firestore, MySQL was chosen to be the database for the project, where all the data for analysis is sourced directly from it, having a structured replica of the Firestore Database obtained from the Legacy System.

By regularizing the data in MySQL, a more standardized and organized dataset was achieved. This standardization significantly simplified the data preprocessing steps in the machine learning workflow.

4.3.9 Express JS

Express.js, often simply called Express, is a streamlined and adaptable framework for Node.js designed to enhance web and mobile application development. It comes equipped with a comprehensive suite of tools and features that cater to the needs of developers in creating sophisticated web and mobile applications. It is commonly used for building server-side applications and APIs due to its simplicity,

performance, and scalability. In this project, Express.js was chosen for several strategic reasons, which are outlined below.[28]

Feature	Advantage for Our Project
Minimalist Framework	Express.js provides essential web application features without dictating any specific architecture, allowing for flexibility and customization in our project.
Middleware Support	The use of middleware modules enables us to extend the functionality of our application easily and efficiently.
Routing System	Its powerful routing system helps manage requests and responses effectively, a crucial aspect for our project's RESTful API design.
High Performance	Known for its high performance, Express.js enhances the responsiveness and speed of our web application.
Community Support	Being one of the most popular Node.js frameworks, it has strong community support, ensuring access to a wide range of resources and troubleshooting assistance.
Easy Integration	Express.js seamlessly integrates with other technologies and databases, which is vital for the diverse tech stack of our project.
Simplicity	Its simplicity and ease of use accelerate development and reduce the learning curve for new team members.

Table 4.6: Express.js Features

It was chosen as a back end web application for this project due to its simplicity, performance, and scalability. Being widely used for building APIs and server-side applications, it offers a great advantage.

Chapter 5

System Design

Provide a detailed explanation of the overall system architecture [?], i.e. the HOW of the project. Use UML, system architecture diagrams, screenshots, code snippets and algorithms to illustrate your design.

Legacy Architecture

The entire system architecture, as designed in the previous projects, is visualized in Figure 5.1. This system consists of three modules that interact with both an external API and user devices. These modules are detailed as follows:

- **Test Application**
A Unity-designed desktop application on a Windows Platform where users can play tests for different game scenarios.
- **Node Js Web Application**
An Express Node JS Web application deployed on Amazon EC2 Virtual machine, serving as a callback endpoint for OAuth 2.0 Authentication for the Polar Flow API.
- **Firestore Data Storage**
Permanent storage medium for user's test and Biometric data.

User Devices:

- Smart Watch
- Smart Phone

The actions in the Figure Figure 5.1 can be summarized as follows:

- sync: passing of biometric data to a smartphone from a Polar watch.

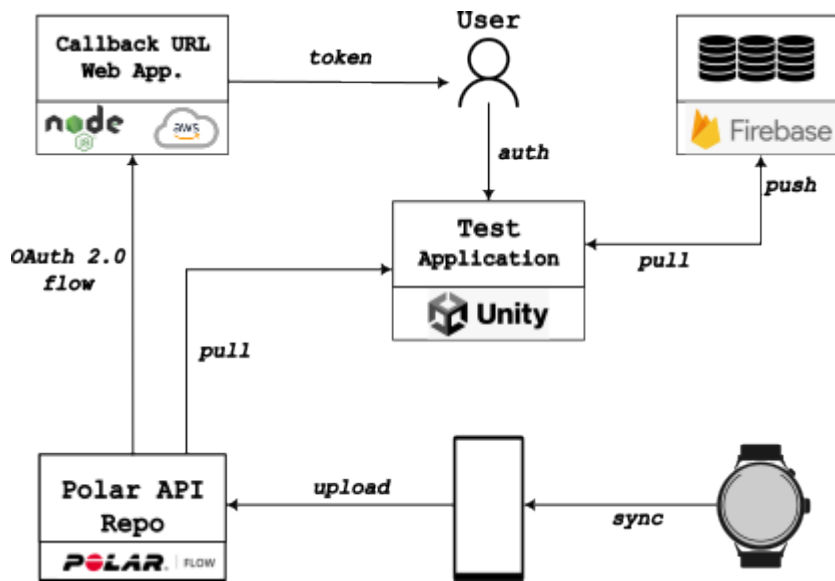


Figure 5.1: System Architecture

- **upload**: uploading of biometric data to Polar API repository.
- **OAuth 2.0 flow**: Multi step OAuth 2.0 authentication protocol.
- **token**: getting verification token from the Callback URL by the user.
- **pull (Polar API - Test Application)**: retrieving biometric data from Polar API by the Test Application.
- **auth**: The user supplies a token to the Test Application as a final step of authentication.
- **pull (Firestore - Test Application)**: retrieving user test/biometric data for rendering by Test Application.
- **push**: saving user's test/biometric data to a Datastore.

Proposed Architecture

The final system architecture will include new modules to facilitate data analytics using AI models. The new system calls for a relational database replica of the Firestore storage to ensure data consistency, predictability and structures suitable for relevant statistical analysis. A.I models as a mini-service that performs operations with supplied data and returns the result of the analysis. A Web Application that functions as a Controller that interfaces the data and A.I modules, Executive

Dashboard that renders user data and results from the A.I models and any other information about the system.

Figure 5.2 shows the proposed architectural layout for the System.

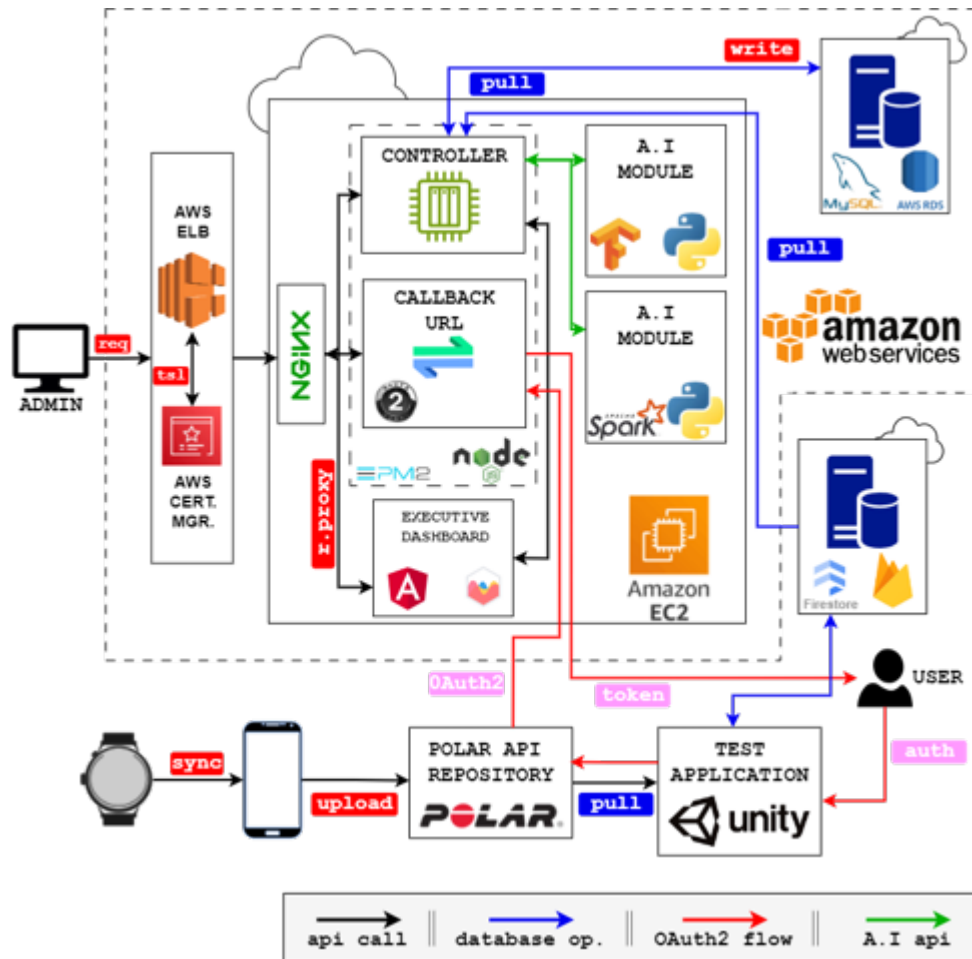


Figure 5.2: Web Application Architecture

Additional modules have been added to the web application namely back end ‘Controller’ & front end ‘Executive Dashboard’, and a ‘Relational Database’. The functions of these modules are explained below:

- **Controller:** This module is responsible for coordinating data fetch from the Relational Database and making sure that the Relational Database is always in sync with the Firebase Datastore. It is also responsible for filtering and structuring data being fed into the A.I. Modules and when necessary using the output for an A.I module as input for another A.I module.

- **Executive Dashboard:** This module provides a Graphical User Interface that displays the working dataset, and offers interfaces to filter, tune and if needed bias the data being sent to the A.I Modules for analysis. This module will be developed using the Angular Framework with integration of D3.js or Chart.js for displaying relevant charts. The choice for Angular Framework over other solutions namely ReactJS and Vue is because of the more structured Angular Framework which is more suited for collaborative work.
- **A.I Modules:** These modules are responsible for running analysis on data. The modules accept input from the 'Controller' and output a comprehensive summary of the results of the analysis done to the 'Executive Dashboard' module. For this purpose, TensorFlow and Apache Spark were chosen for the following reasons:
 - **Capabilities:** Apache Spark is capable of performing linear and logical regression[29] which are common tools used for this class of problem. TensorFlow also has the same capabilities and can perform regression neural network models[30].
 - **Accessibility:** Both tools are open-source and available to the general public[31].
 - **Python API interface:** Both tools provide a Python API. This eases deployment as Python is platform-independent and relatively easy setup in a Ubuntu Virtual machine.
 - **Large Community:** Both have a large community of contributors. Documentation and tutorials are vastly available on the internet.
 - **Relational Database:** Data for the analysis will be sourced directly from this Database which is a structured replica of the Firestore Database as obtained from the Legacy System. The rationale for this Module is to provide a flexible means of retrieving data in various forms and shapes as may be needed for analysis and also overcome the high complexities involved in Firestore for compound queries.
 - **Chart API:** Is responsible for displaying test results and biometric data to the users. This already developed module will be integrated into the 'Test Application' and undergo some modification to be able to display custom data as requirements change.

Chapter 6

System Evaluation

Evaluate your project against the objectives set out in the introduction. This chapter should present results if applicable and discuss the strengths and weaknesses of your system. This is a clear opportunity for you to demonstrate your critical thinking in relation to the project.

6.1 Working with Tables

Table 6.1 can be referenced with the label given to the table, i.e. `\ref{table:HexToBin}`. Note that \LaTeX will place the table wherever it deems fit. Don't bother trying to change where a table or figure is placed until your document is ready for final layout.

Hexadecimal to Binary					
Hex	Binary 2	Hex	Binary	Hex	Binary
1	00000001	B	00001011	15	00010101
2	00000010	C	00001100	16	00010110
3	00000011	D	00001101	17	00010111
4	00000100	E	00001110	18	00011000
5	00000101	F	00001111	19	00011001
6	00000110	10	00010000	1A	00011010
7	00000111	11	00010001	1B	00011011
8	00001000	12	00010010	1C	00011100
9	00001001	13	00010011	1D	00011101
A	00001010	14	00010100	1E	00011110

Table 6.1: Conversion from Hexadecimal to Binary

Chapter 7

Conclusion

Briefly summarise your context and objectives. Remind the reader about the overall rationale and goals of the project. Highlight your findings from the System Evaluation chapter.

Appendix

Ethics Application



Taught Programme Research Ethics Approval Application Form

Research undertaken by taught students must receive ethical approval unless deemed exempt. This application form may be completed by an individual student or by a Programme Board/Lecturer for a group of similar research projects.

This application is completed by:

Student: ✓

OR

Lecturer on behalf of Programme Board: ☐

PART A

Applicant Details	
Name:	Otito Mbelu, Rodrigo Almeida
Student ID:	G00397738, G00377123
Programme Title:	Computing in Software Development
Programme Stage:	Year 4
Research Supervisor's Name:	Dr. Damien Costello

Project Details	
Research Study Title:	Biometric Data Analysis in Digital Game Scenario
Research Study Summary (max 100 words):	
<p>The main hypothesis of this research is to find if there is any correlation between a user's biometric data and their performance in a first-person shooter game scenario.</p> <p>This research study will try to quantify how biometric data such as Heart Rate Variation, Heart Rate, active steps taken, quality and quantity of sleep, etc. affect select gaming skills like eye-to-hand coordination, fine-motor skills and reaction time.</p>	

Risk Checklist			
Please answer ALL the questions in each of the sections below – Tick YES or NO			
	Will the research study....?	YES	NO
1	Involve direct and/or indirect contact with human participants?	✓	
2	Involve analysis of pre-existing data which contains personal or sensitive information not in the public domain?		✓
3	Require permission or consent to conduct?	✓	
4	Require permission or consent to publish?		✓
5	Have a risk of compromising confidentiality?		✓
6	Have a risk of compromising anonymity?		✓
7	Collect/contain personal data i.e. any information that relates to an identified or identifiable individual?		✓
8	Collect/contain sensitive personal data e.g. health data, sexual orientation, race religion?	✓	
9	Contain elements which you OR your supervisor are NOT trained to conduct?		✓
10	Use any information OTHER than that which is freely available in the public domain?		✓
11	Involve respondents to the internet or other visual/vocal methods where participants may be identified?		✓
12	Include a financial incentive to participate in the research?		✓
13	Involve our own students or staff?	✓	
14	Take place outside Ireland?		✓
15	Involve participants who are vulnerable or at risk?		✓
16	Involve any participants who are unable to give informed consent?		✓
17	Involve data collection taking place BEFORE informed consent is given?		✓
18	Involve any deliberate deception or covert data collection?		✓
19	Involve a risk to the researcher or participants beyond that experienced in everyday life?		✓
20	Cause (or could cause) physical or psychological harm or negative consequences?		✓
21	Use intrusive or invasive procedures?		✓
22	Involve a clinical trial?		✓

Risk Checklist			
Please answer ALL the questions in each of the sections below – Tick YES or NO			
	Will the research study....?	YES	NO
23	Involve the possibility of incidental findings related to participant health status?		✓
24	Involve the remuneration of research participants?		✓

If, as a student, you answered **NO** to all the above questions your research supervisor will review, and if in agreement sign below to indicate that this form does not have to be submitted to the Taught Programme Research Ethics Committee.

<i>Name</i>		<i>Signed</i>		<i>Date</i>	
-------------	--	---------------	--	-------------	--

If you answered **YES** to any of the above questions, you need to complete part B below.

PART B

1	Project Overview
<p>Please give a brief overview of the study, including a summary of the aims and objectives.</p> <p><u>Help:</u> Describe the purpose of the research and what question(s) the project should answer.</p>	
<p>Introduction to PUBG: PUBG, short for Player Unknown's Battlegrounds, is a popular online multiplayer battle royale game developed and published by PUBG Corporation, a subsidiary of Bluehole Studio. In PUBG, up to 100 players parachute onto an island and scavenge for weapons and equipment to kill others while avoiding getting killed themselves. The game features a shrinking safe area to force players into close encounters, promoting tactical gameplay and intense firefights. The last player or team standing wins the game. PUBG became widely popular upon its release in 2017 and is available on various platforms, including PCs, consoles, and mobile devices¹. Since its launch in 2017, PUBG has become incredibly popular, with more than 280M active players. People of all ages and from all over the world enjoy playing this game. In PUBG, players need good hand-eye coordination and quick reflexes to compete well against others. The game offers a variety of weapons, and players can customize their controls to fit their preferences. This project builds upon the research done by university students, focusing on collecting biometric data to improve player performance in digital games.</p> <p>The original project, titled "Biometric Data Collection for Performance Optimization in a Digital Game Scenario", aimed to find out if a player's body data could help them play a video game better. The researchers set up a practice game similar to PUBG: Battlegrounds, with the same weapons and controls in a Unity Desktop Application. Biometric data was supplied by a Smart Watch. In this instance, the only data used to test the application was produced by the developers themselves and from the stakeholders, for this reason, it wasn't necessary an Ethics application as there wasn't a recruitment of volunteers at the time. The goal of the original project was to see if there was a connection between how players performed in the game and their body data. In our third-year project, our team was given the challenge of enhancing the visualization of data from both the test games and the Polar API. To achieve this, we developed a chart API, enabling users to conveniently view all the results on a single page. The chart API was specifically created to present the data collected during the initial research in a coherent and user-friendly chart format, integrating various user data into one unified visual representation.</p> <p>Project Objective The aim of this research project is to overcome the limitations mentioned in previous projects and complete the future development goals of both projects. These include:</p> <ul style="list-style-type: none"> • Chart API Integration: Integration of developed Chart API into the test application. • Offline Data Storage: Provision will be made for offline temporary file storage to improve the overall reliability of the whole system • PUBG API: Further research on new developments in the PUBG API for better user experience. <p>And eventually, seek to answer the following research questions:</p> <ul style="list-style-type: none"> - Can the user's current physical condition as indicated by their Biometric data, have any direct relationship with their performance in such a gaming scenario? - Can Biometric and test data help suggest the most suitable settings for different game scenarios? 	

2	Methodology
<p>Please give a description of the methodology, including any data collection and analysis methods.</p> <p><u>Help:</u> Give an outline of the study here. If the project is complex, you can also submit the research proposal/protocol (no more than 2-3 A4 sides) if this would help the reviewer's understanding of the project. Include details of your (or the Research Supervisor's) appropriate skills and qualifications to carry out this research. Consideration of how, and for what duration are stored should be provided under Section 7 below.</p>	
<p>This research is geared towards finding a correlation between users' biometric data to their performance in a first-person shooter gaming scenario. The quality of the user data collected will be critical in ensuring that the research aims are achieved. For Quality Assurance purposes, a systematic step-by-step approach has been designed for the data capturing process.</p> <p>An activity monitoring device will be distributed to volunteers in the form of a Polar Watch. Individuals are expected to use these devices at a pre-designated times prior to undertaking a test in the form of a First-Person Shooter Game. Results from the test will be paired with their biometric data for further analysis.</p> <p>Data of Interest</p> <p>For the purpose of the research, different categories of data will be collected from the volunteers through the monitoring device to help achieve the stated goals. These data relate to the volunteer's physical information, sleep data, daily activities, and nightly recharge.</p> <p>For the Test Game, volunteers are expected to play a total of 3 different categories of test to gauge their performance in these tests. The tests are Audio Test, Visual Test, and Fine Motor Test. The tests are designed to be played sequentially and will typically take a total of 15 minutes to conclude.</p> <p>Procedure</p> <p>The data collection process is expected to follow the below listed steps to maximize throughput and minimize the average time spent processing each participant.</p> <p>Location</p> <p>The data collection effort will be carried out at the Gym on the Atlantic Technological University (ATU) Galway campus.</p> <p>Time</p> <p>The following weekly schedule will be available for volunteers at various times that best suit their personal schedule.</p> <ul style="list-style-type: none"> · Tuesdays: 12:00 - 13:00, 15:00 - 16:00 · Wednesdays: 12:00 - 13:00, 15:00 - 16:00 · Thursdays: 12:00 - 13:00, 15:00 - 16:00 · Fridays: 13:00 - 15:00 	

Note: The timing is open for adjustment to suite volunteers

Biometric Data Procedure:

For the very first time processing a volunteer, some user information will be needed to have them registered with the activity monitoring device and their data is saved to the manufacturer's repository.

The Activity Monitoring Device to be used in this research is the Polar Vantage Smart Watch. The watch is capable of monitoring and capturing user's biometric data and saving the data to a repository where it can be accessed online via an API. The device is commercially available and accessible to the public. The rationale for using the Polar Vantage Watch is because it has been widely used in both academic and industrial research projects and validation. Most importantly, these devices are available to us in such quantity that can satisfy our research needs.

Physical Information

These are the first group of data to be collected about the individual to register the Monitoring device manufacturer. The data is stored with the manufacturer (Polar) which is GDPR compliant.

Sleep Data

According to the manufacturer's manual, the watch is capable of recording the quality and quantity of sleep and also showing how long spent on each stage of sleep. Light Sleep, Deep Sleep, and REM Sleep with the duration of each type of sleep. For the purpose of the research, an aggregated total of the various categories of sleep will be recorded and used.

Volunteers are expected to wear the watch to sleep the previous night before their scheduled test

- Sleep (minutes)

Daily Activities

According to the Activity monitor manufacturers' manual, "Polar device uses an internal 3D accelerometer to record your wrist movements. It analyses the frequency, intensity and regularity of your movements together with your physical information." calories, active calories, active steps and their respective durations are collected from through the device. For the purpose of the research, the active steps will be used.

Volunteers are expected to wear the previous day for this data to be available.

- Active Steps (count)

Nightly Recharge

From the Activity monitor manufacturers' manual, the Nightly Recharge is recorded as follows: "is an overnight recovery measurement that shows how well your body has coped with overall stress you have experienced lately." The parameters, measured during roughly the first four hours of your sleep are heart rate, heart rate variability and breathing rate. For the purpose of the research, the following parameters and derived parameters will be used:

- Heart Rate Average (bpm)
- Heart Rate Maximum (bpm)

· Heart Rate Variability (HRV)

Once again, volunteers are expected to wear the watch to sleep the previous night before their scheduled test.

Test Data Collection

The test data is generated at the completion of the test session in the Test Application. A test session consists of three categories of tests, which will be undertaken sequentially in different stages. For quality assurance, the test will be undertaken in a controlled environment using the same hardware in similar conditions over the course of the trials.

For the purpose of the trial, a special designated room has been reserved for the data collection effort. The following metrics for the different tests will be captured and will be subsequently used for the research effort.

Audio Test: designed to test users' auditory precision. This test is expected to generate.

Average Response Time: this is a measure of how quickly a user is able to react and identify the source of a sound with varying audibility (in decibels) in a 3-Dimensional space.

Visual Test: designed to test users' visual perception. At the end of this category of test, the following metrics were designed to measure users' performance.

Average Response Time: a measure of how quickly a user can identify and engage targets.

Shot Accuracy: a measure of the percentage of successful shots to the number of targets spawned.

Target Accuracy: a measure of the percentage of successful shots taken to a number of targets hit.

Fine Motor Test: designed to test users' perception of depth and eye to hand coordination.

Average Tracking Time: measure of average time it took for a player to successfully track, engage, and eliminate a target.

Accuracy: measure of the percentage of shots fired to the number of targets hit.

3	Main Ethical Considerations
---	------------------------------------

Please give a brief description of the main ethical considerations involved in the study.

Help: Highlight here the main ethical considerations for the study (which may concern, e.g., the type of participants, the sensitive nature of the study, the data collection process, and security-sensitive research) and advise how the main issues will be addressed. If the project is funded, give details here, and whether there are any potential conflicts of interest involved in the study.

NB: Section 5 below addresses: recruitment; voluntary participation; consent; and, the right to withdraw. Those details need not also be entered here.

There are no conflicts of interest, and the project is not funded.

4	Human Participants
<p>If the study includes Human Participants (or their data), please give a description of who will be included.</p> <p><u>Help:</u></p> <ul style="list-style-type: none"> • Please note this should include sample size/number of participants, whether the project will focus on any particular groups/individuals, if it will include any at risk or vulnerable participants, participants aged 16 years or under, etc. Please also specify the rationale for including/excluding groups of participants. • If the research involves secondary data not in the public domain, give details in this section. <p>For the research, physically active individuals that participate in work out activities at a minimum of 3 days a week are desired. Furthermore, the participants are expected to be regular console game players.</p> <p>We expect a minimum of 4 participants to make the research viable and a maximum of 20 volunteers which will be separated into a control group and intervention groups.</p>	

5	Recruitment, Voluntary Participation, Consent and Right to Withdraw
<p>If the study includes Human participants, please give a brief description of the recruitment process, how voluntary participation will be ensured, if (and how) informed consent will be obtained prior to participants taking part in the study, and the right of withdrawal from the research process.</p> <p><u>Help:</u></p> <ul style="list-style-type: none"> • This should include clear information on how participants will be identified, approached and recruited; whether the study will include any covert research or deliberate deception; whether help is required from a third party/ gatekeeper to access participants; what information will be given to participants, etc. • If expenses or any incentives are to be offered to participants, give full details. • If research involves students, colleagues and/or other employees then specify the rationale for this and how issues of coercion or feelings of obligation will be addressed. • <u>If data is held on participants, research using that data may require permission from the participant.</u> • Regarding withdrawal from the study, discuss the different stages/dates a participant could withdraw or withdraw their data, and how they could do this. <p>The poster below will be printed and placed around the ATU Campus. It will also be sent through email to all students registered on Galway Campus. The poster contains an electronic form that volunteers will be prompted to fill out when scanning the QR code from the poster or accessing the link https://forms.office.com/e/EmhfampDUT.</p> <p>Once volunteers register interest the email address will be kept as stated in the form and then we will contact them through their registered email.</p>	

GAMERS AND FITNESS ENTHUSIASTS
WE NEED YOU
RESEARCH PARTICIPANTS NEEDED

Atlantic Technological University
 Ollscoil Teicneolaíochta an Atlántaigh

SCAN THE QR CODE AND CHECK IF YOU QUALIFY

QUESTIONS OR TROUBLE ACCESSING IT?
 PLEASE EMAIL G00377123@ATU.IE

6	Risks and Benefits
	<p>Please give a brief description of how, when and where the research will take place and whether there are any risks and/or benefits involved.</p> <p><u>Help:</u></p> <ul style="list-style-type: none"> • This should include information on what participants will be required to do, the rationale for this and the level of risk involved. • When considering risks, please refer to risks to the participants (e.g., for research in sensitive areas, where there is a balance of power), the researcher, any other parties to the research; and also any health and safety issues for anyone involved (e.g., for lone researchers carrying out fieldwork).
	<p>The research is a software project that analyses data and does not involve the volunteers performing any physical activity other than does in their normal routine.</p> <p>Volunteers are expected to wear a biometric data monitoring device that records physical activities.</p> <p>The volunteers are expected to play a game that will be installed on their computer. On average the game is expected to last for 15 minutes.</p> <p>Data from the game and device monitor are collected during gameplay and stored in a secured database.</p> <p>Regression and Correlation analysis are done using these data.</p>

7	Personal Data, Anonymity and Confidentiality
	<p>Please specify what type of information/data will be collected/analysed and the source(s). In addition, specify if and how the anonymity of participants will be ensured, and information be kept confidential.</p> <p><u>Help:</u> This should include information on whether new information/data are being collected or uses data that are already in the public domain; whether the data includes personal data; whether the data includes sensitive personal data e.g. health data, sexual orientation, race, religion; how the data will be processed and stored; who will have access to it; who it will be shared with; how long data will be retained; how it will be destroyed; the Data Protection requirements for any sensitive personal data, etc. In addition, include whether there may be any requirements for disclosure of information to other parties due to professional practice or legal reasons. If there are limits to confidentiality, explain clearly how the participants would be advised about these limits and possible outcomes.</p>
	<p>There are basically two categories of data to be collected which are:</p> <ul style="list-style-type: none"> • Game test data that are generated during gameplay. • Biometric data is obtained from a biometric data monitor (watch). <p>Specifically, the data to be collected from the Biometric data monitor are exhaustively listed below:</p> <ul style="list-style-type: none"> • Maximum Heart Rate • Average Heart Rate • Active Steps • HRV • Sleep (hours) <p>To get a volunteer setup for data collection, the volunteer will have to register with the activity monitor manufacturer (i.e. Polar Flow). Such information like height, weight, etc will be taken to ensure the accuracy of data being collected by the activity monitor as advised in their documentation. These data are not of interest to the research and will be stored with the Activity monitor manufacturers.</p> <p>Volunteer's data collected will be stored in a Firestore database through Firebase which is a secured Backend-as-a-Service Cloud service that provides encryption, security, and availability. Individual user-specific data are stored with the alias/username they choose to register with.</p> <p>** It is important to note that no data will not be traceable to any individual as their data will be annotated with their chosen alias/username.</p>

8	Reporting and Dissemination
<p>Please give details of the planned dissemination and specify if the findings from the research will be published and whether any permission is required for this.</p> <p><u>Help:</u> This should include information on the methods of dissemination (e.g., dissertation/thesis) and/or what will be published and where (research papers, conference presentations). Specify if any permission is needed (e.g., from participants, clients, gatekeepers, etc.) prior to publication, and whether there are any potential issues relating to Intellectual Property Rights when creating or using materials.</p>	
<p>As indicated in section A. There are no plans at the moment to publish the research.</p>	

9	Location of research
<p>Will the research take place outside of Ireland?</p>	
<p>YES <input type="checkbox"/> NO <input type="checkbox"/> If yes, give details below.</p> <p><u>Help:</u> If yes, please specify where the research will take place. Research must comply with the laws of the country where it is taking place and also comply with local Data Protection and Intellectual Property legislation: you must confirm that your research is compliant with local requirements and how you have ascertained this. Advise if the project requires ethical approval in-country and how this has been ascertained. If approval is required, a copy of this should be included in the application or details of the process of how it will be obtained. Please make reference to insurance and indemnity cover for the project where relevant.</p> <p><u>Note:</u> If data is to be processed or stored outside the EEA contact dpo@atu.ie</p>	
<p>NO</p>	

10	Collaborative Projects
Is the research a collaborative project (i.e., it involves more than one institution)?	
YES <input type="checkbox"/> NO <input type="checkbox"/> If yes, give details below.	
<p><u>Help:</u> If yes, please specify the other institutions involved and if ethical approval needs to be / has been given by them. Please also specify what procedures have been put in place to ensure ethical compliance from all partners.</p> <p><u>Note:</u> If personal data is being shared between institutions then a data sharing agreement must be in place. Contact dpo@atu.ie</p>	
NO	

11	Any other permission or external ethical approval required to undertake the project
Please specify if the project requires any other ethical approval or permissions not mentioned previously in this application and how and when these will be obtained.	
<p><u>Help:</u></p> <ul style="list-style-type: none"> • Other permissions: ethical approval does not give the right of access to the Institute's students, staff or the use of Institute premises to carry out research, and you may need to contact an appropriate Institute gatekeeper for agreement to approach potential participants or for the use of premises, so please give details. • Gatekeepers: permission of a gatekeeper for initial access to participants may be required or to carry out data collection on their premises. • If the project requires approval from an external ethics committee, this should normally be obtained prior to submitting this application. • If a Disclosure and Barring Service check is required due to the specific participant group, give details. • Regarding insurance and indemnity cover, some projects will require individual confirmation of cover. See the Research Ethics Procedures document for more details. 	
NO	

SUPPORTING DOCUMENTATION: what to submit with the application

For projects involving human participants, you must submit, where appropriate, the Participant Information Sheet/s and consent form/s. You must also submit every communication a participant will see or receive. Failure to do so will cause delays to the application.

DECLARATIONS AND SIGNATURES**STUDENT**

I confirm that I will undertake this project as detailed in Part A and Part B of the application. I understand that I must abide by the terms of this approval and that I may not make any substantial amendments to the project without further approval. I understand that research with human participants or their data must not commence without ethical approval.

<i>Signed</i>	Rodrigo Almeida Otito Mbelu	<i>Date</i> 31/10/2023	
---------------	------------------------------------	-------------------------------	--

RESEARCH SUPERVISOR RECOMMENDATION FOR STUDENT PROJECT

I confirm that the committee has considered part A and part B of the application. The project is viable and the student has the appropriate skills to undertake the project. Where applicable, the Participant Information Sheet and recruitment procedures for obtaining informed consent are appropriate and the ethical issues arising from the project have been addressed in the application. I understand that research with human participants must not commence without ethical approval. I recommend this project for approval.

<i>Name</i>		<i>Signed</i>		<i>Date</i>	
-------------	--	---------------	--	-------------	--

Comment(s):

E.g. if similar research projects have been previously approved.

LECTURER ON BEHALF OF PROGRAMME BOARD

I confirm that the project will be undertaken as detailed in stage one and stage two of the application. I understand that I must abide by the terms of this approval and that I may not make any substantial amendments to the project without further approval. I understand that research with human participants or their data must not commence without ethical approval.

<i>Signed</i>		<i>Date</i>	
---------------	--	-------------	--

PROJECTS APPROVED BY THE RESEARCH ETHICS SUB-COMMITTEE

I confirm that this project was considered by the Taught Programme Research Ethics Committee and has received ethical approval.

<i>Chair</i>		<i>Signed</i>		<i>Date</i>	
--------------	--	---------------	--	-------------	--

This form will be retained for the purposes of quality assurance of compliance and audit for THREE years



Ollscoil
Teicneolaíochta
an Atlantaigh

Atlantic
Technological
University

PARTICIPANT INFORMATION SHEET

Biometric Data Analysis in Digital Game Scenario

We are Rodrigo Almeida and Otito Mbelu and we are students in Bachelor of Science (Honours) in Computing in Software Development at ATU Galway City. We are recruiting volunteers to take part in a research study. The aim of this study is to answer the following research questions:

- Can the user's current physical condition as indicated by their Biometric data, have any direct relationship with their performance in such a gaming scenario?
- Can Biometric and test data help suggest the most suitable settings for different game scenarios?

This participant information sheet outlines what the study involves and what will be required of you if you choose to volunteer to participate.

What is the purpose of this study?

The main hypothesis of this research is to find if there is any correlation between a user's biometric data and their performance in a first-person shooter game scenario.

This research study will try to quantify how biometric data such as Heart Rate Variation, Heart Rate, active steps taken, quality and quantity of sleep, etc. affect select gaming skills like eye-to-hand coordination, fine motor skills and reaction time.

What will be required of you?

For this study, you will be required to wear the Polar watch which is referred sometimes to "Activity Monitoring Device" in this text, a day prior to undertaking a test.

Location

The data collection effort will be carried out at the Gym on the Atlantic Technological University (ATU) Galway campus.

Time:

The following weekly schedule will be available for volunteers at various times that best suit your personal schedule.

- Tuesdays: 12:00 - 13:00, 15:00 - 16:00
- Wednesdays: 12:00 - 13:00, 15:00 - 16:00
- Thursdays: 12:00 - 13:00, 15:00 - 16:00
- Fridays: 13:00 - 15:00

Note: The timing is open for adjustment to suit you

Results from the test will be paired with their biometric data for further analysis.

What will happen to the information that is collected about me?

The information gathered from the study will be handled in complete confidence and cannot be traceable back to you. These data are stored in an encrypted database system with everything relating to a user linked with the alias (username) they choose to register with. When the study is finished, information will be kept on the researcher Luke Smyth.

What are the benefits?

The benefit of taking part in this research is that if such a correlation is found between your biometric data and your performance in the game you can use the results to improve your own performance when setting up in any First-shooter game. You may benefit by taking part in seeing how a final year project is run you may take some valuable knowledge and apply it to your own studies if applicable.

What are the risks?

There will be no risks as this study is based on the volunteer's normal activities and playing a game in a controlled environment.

What if you change your mind during the study?

Please be aware that individuals who volunteer to take part in the study have the right to withdraw from the project at any time without the need to provide a reason or notice period. Should you feel at any stage that you want to stop taking part in the study, then this is dealt with in a sensitive and confidential manner.

What happens at the end of the study?

At the end of the study, the data gathered will be disseminated for the purposes of a written academic report and submitted to ATU for assessment purposes. A summary of the results can be provided to you, upon request. All data will be held securely, in line with GDPR regulations, for a period of ten years, as per the ATU records retention policy.

What if you have more questions or do not understand something?

It is important that you feel completely at ease during the research. If you do not understand any aspect of the research, please contact us or my research supervisor to discuss any questions that you might have. Alternatively, you may contact the Head of Department, if you wish to speak to someone independent from the study.

Research Supervisor: Dr. Damien Costello

Head of Department: Dr. Gareth Roe

Thank you for taking the time to read this. I would be grateful if you would consider participating in this study.

Yours sincerely,

Student Name: Rodrigo Almeida, Otito Mbelu

Student Signature:

Date: *[Insert date of circulation of this document here]*

Contact Details: G00377123@atu.ie or G00397738@atu.ie



Ollscoil
Teicneolaíochta
an Atlantaigh

Atlantic
Technological
University

PARTICIPANT CONSENT FORM

Effects of Post-Action potentiation on sprint performance

Please tick the appropriate boxes

	YES	NO
1. I have read the Participant Information Sheet for this study and understand what's involved.	<input type="checkbox"/>	<input type="checkbox"/>
2. I have been given the opportunity to ask questions about the study.	<input type="checkbox"/>	<input type="checkbox"/>
3. I agree to take part in the study under the conditions set out in the Participant Information Sheet.	<input type="checkbox"/>	<input type="checkbox"/>
4. I understand that my taking part is voluntary and that I can withdraw my consent at any time before my data is de-identified or amalgamated with other data.	<input type="checkbox"/>	<input type="checkbox"/>
5. I understand that, in any report on the results of this study, my identity shall remain anonymous.	<input type="checkbox"/>	<input type="checkbox"/>

Signature of Participant:

Date:

Signature of Student:

Date:

Appendix D: Participant Withdrawal Form

PARTICIPANT WITHDRAWAL FORM

Reference Number:

Participant name or Study ID Number:

Title of Project: Biometric Data Analysis in Digital Game Scenario

Name of Principal Investigator: Rodrigo Almeida and Otito Mbelu

Name of the person to whom this form should be submitted: Gary Flynn

Participant to complete this section. Please initial one of the following boxes:

1. I confirm that I wish to withdraw from the study before data collection has been completed and that none of my data will be included in the study.	
2. I confirm that I wish to withdraw all of my data from the study before data analysis has been completed and that none of my data will be included in the study.	
3. I confirm that although the results of the study have already been produced and cannot change, I wish to be forgotten and that all of my personal data is deleted from verification records maintained by the university about the study. I understand that this means that only those data identifying me will be deleted.	

Your name is required to verify that you have withdrawn your data from the study as specified above. In the case of (3), above, we will need to retain this form until.....

It may be necessary to share this information with internal examiners, external examiners, and / or journal editors for the purposes of verification of findings and tracing results of studies to the raw data used.

This form will be stored securely until, when it will be destroyed, and will not be shared with anyone else.

Signature of participant:	Date:
Signature of person who will ensure that the stated data have been deleted:	Date:

Microsoft Form for Volunteer Recruitment

Research Participant Needed

Join our study exploring the link between biometric data and gaming performance. Help us understand how your body's responses impact your gaming abilities. Your input is crucial for our research. Thank you for participating!

* Required

* This form will record your name, please fill your name.

1. How many times a week do you exercise? (Select one option) *

- ☐ 1-3 times/week
- ☐ 3-5 times/week
- ☐ 5+ times/week
- ☐ I don't exercise

2. What is the average duration of your exercise sessions? *

- ☐ 30 minutes/session
- ☐ 45 minutes/session
- ☐ 60 minutes/session
- ☐ 75 minutes/session
- ☐ 90 minutes/session
- ☐ 90+ minutes/session

3. Do you own an activity tracker (E.g. fitness watch)? (Select one option) *

- ☐ Yes
- ☐ No

4. Do you play video games? (Select one option) *

☐ Yes

☐ No

5. What gaming platform do you use most frequently? (Select one or more options) *

☐ PC

☐ Xbox

☐ PlayStation

☐ Other

6. What kind of games do you enjoy the most? (Select one or more options)

☐ First Person Shooter

☐ Third Person Shooter

☐ Soccer

☐ Car Racing

☐ Other

7. What brand is your activity tracker? *

☐ Polar

☐ Garmin

☐ Apple

☐ Fitbit

☐ Other

8. I consent to be contacted by the researchers about my participation in this project and my data might be kept for the purpose of this research and will not be shared with any third parties.

(You have the right of withdrawal at anytime, by emailing G00377123@atu.ie) *

☐ Yes

☐ No

This content is neither created nor endorsed by Microsoft. The data you submit will be sent to the form owner.

 Microsoft Forms

Bibliography

- [1] Mihai Liviu Despa. Comparative study on software development methodologies. *Database Systems Journal*, 5(3), 2014.
- [2] Emily Ortega and CJK Wang. Pre-performance physiological state: Heart rate variability as a predictor of shooting performance. *Applied psychophysiology and biofeedback*, 43:75–85, 2018.
- [3] Marek Malik, J Thomas Bigger, A John Camm, Robert E Kleiger, Alberto Malliani, Arthur J Moss, and Peter J Schwartz. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European heart journal*, 17(3):354–381, 1996.
- [4] James McNames and Mateo Aboy. Reliability and accuracy of heart rate variability metrics versus ecg segment duration. *Medical and Biological Engineering and Computing*, 44:747–756, 2006.
- [5] Daniel Bonnar, Benjamin Castine, Naomi Kakoschke, and Gemma Sharp. Sleep and performance in eathletes: for the win! *Sleep Health*, 5(6):647–650, 2019.
- [6] Polar. Polar vantage v user manual, 2324.
- [7] Angular Components.
- [8] Angular Modules.
- [9] Angular Services.
- [10] Helder Da Rocha. *Learn Chart.js: Create interactive visualizations for the web with chart.js 2*. Packt Publishing Ltd, 2019.
- [11] Nick Qi Zhu. *Data visualization with D3.js cookbook*. Packt Publishing Ltd, 2013.
- [12] News. *Significance*, 9(4):2–3, 08 2012.

- [13] Ivanilton Polato, Reginaldo Ré, Alfredo Goldman, and Fabio Kon. A comprehensive view of hadoop research—a systematic literature review. *Journal of Network and Computer Applications*, 46:1–25, 2014.
- [14] Salman Salloum, Ruslan Dautov, Xiaojun Chen, Patrick Xiaogang Peng, and Joshua Zhexue Huang. Big data analytics on apache spark. *International Journal of Data Science and Analytics*, 1:145–164, 2016.
- [15] Eman Shaikh, Iman Mohiuddin, Yasmeen Alufaisan, and Irum Nahvi. Apache spark: A big data processing engine. In *2019 2nd IEEE Middle East and North Africa COMMunications Conference (MENACOMM)*, pages 1–6, 2019.
- [16] Bo Pang, Erik Nijkamp, and Ying Nian Wu. Deep learning with tensorflow: A review. *Journal of Educational and Behavioral Statistics*, 45(2):227–248, 2020.
- [17] Ki-Hyun Jung. A study on machine learning for steganalysis. 05 2019.
- [18] <https://nodejs.org/en/>. Node.js, Accessed: 23 January 2024.
- [19] Stefan Tilkov and Steve Vinoski. Node.js: Using javascript to build high-performance network programs. *IEEE Internet Computing*, 14(6):80–83, 2010.
- [20] Inc. (2021) npm. npm public registry, 2021.
- [21] Mike Cantelon, Marc Harter, TJ Holowaychuk, and Nathan Rajlich. *Node. js in Action*. Manning Greenwich, 2014.
- [22] David Gonzalez. *Developing Microservices with node. js*. Packt Publishing Birmingham, UK, 2016.
- [23] Olivier Aumage, Gabriel Antoniu, Luc Bougé, Vincent Danjean, and Raymond Namyst. Getting started with pm2. *LIP, ENS-Lyon*, 2001.
- [24] Dirk Merkel et al. Docker: lightweight linux containers for consistent development and deployment. *Linux j*, 239(2):2, 2014.
- [25] Inc. Docker. Docker hub, 2021.
- [26] Will Reese. Nginx: the high-performance web server and reverse proxy. *Linux Journal*, 2008(173):2, 2008.
- [27] Paul DuBois. *MySQL*. Addison-Wesley, 2013.
- [28] Evan Hahn. *Express in Action: Writing, building, and testing Node. js applications*. Simon and Schuster, 2016.

- [29] <https://spark.apache.org/docs/latest/ml-classification-regression.html>. Apache Spark, Accessed: 14 October 2023.
- [30] Jeff Heaton, Steven McElwee, James Fraley, and James Cannady. Early stabilizing feature importance for tensorflow deep neural networks. In *2017 International Joint Conference on Neural Networks (IJCNN)*, pages 4618–4624, 2017.
- [31] <https://www.databricks.com/blog/2016/01/25/deep-learning-with-apache-spark-and-tensorflow.html>. Apache Spark, Accessed: 11 October 2023.