FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Automatic Website Generation for Low-ranking Football Clubs: An Empirical Study

Beatriz Costa Silva Mendes

Master in Informatics and Computing Engineering

Supervisor: André Restivo

July 26, 2023

Automatic Website Generation for Low-ranking Football Clubs: An Empirical Study

Beatriz Costa Silva Mendes

Master in Informatics and Computing Engineering

Approved in oral examination by the committee:

President: Filipe Correia

Referee: Luis Borges Gouveia

Supervisor: André Restivo

July 26, 2023

Resumo

Na era digital atual, os *websites* tornaram-se fontes de informação valiosas, frequentemente atualizadas com dados exaustivos de jornalistas, indivíduos certificados e contribuições do público. Os clubes de futebol compreendem a importância de manter *websites* atualizados para cativar adeptos e investidores, gerar receitas e aumentar a visibilidade. No entanto, existe uma discrepância digital. Por um lado, os clubes de escalões superiores a usufruem de *websites* mais completos. Em contrapartida, clubes mais pequenos enfrentam limitações devido a restrições de recursos e a plataformas desatualizadas, o que realça a necessidade de websites atualizados e baseados em dados, independentemente da dimensão ou da classificação do clube.

Esta dissertação propõe uma solução inovadora para colmatar esta lacuna através do desenvolvimento de uma plataforma automatizada que utiliza dados publicamente disponíveis para gerar *websites* atualizados, de fácil utilização e personalizáveis para equipas desportivas locais. O objetivo é fornecer um meio prático e económico para os clubes criarem e manterem *websites* que reflitam a identidade da sua marca, mantendo-se atualizados com a natureza evolutiva do desporto.

A abordagem de validação envolve a verificação manual, a comparação de registos de data e hora com os *websites* oficiais dos clubes e a recolha de comentários dos adeptos para avaliar a atualidade dos *websites* criados. As avaliações centradas no utilizador e a *System Usability Scale* [8] medem a usabilidade, enquanto as avaliações manuais examinam a facilidade de criação e a atração visual.

A avaliação empírica produziu resultados positivos relativamente à atualidade, à usabilidade e ao aspeto visual dos *websites* criados. Os adeptos dos clubes da primeira e da segunda divisão reconheceram a facilidade de utilização e a atualidade dos *websites* gerados, ultrapassando os *websites* oficiais. Além disso, manifestaram preferência pelos *websites* desenvolvidos devido à sua maior facilidade de utilização. A aplicação da *System Usability Scale* forneceu ainda provas quantitativas que sustentam a excelente usabilidade dos *websites* criados.

Os resultados confirmam a hipótese, destacando benefícios como a melhoria da comunicação, a redução dos esforços de manutenção, a melhoria da experiência do utilizador e o reforço da identidade da marca. A solução proposta reduz o discrepância digital, garantindo a igualdade de oportunidades para todos os clubes e promovendo a transparência e o crescimento do desporto.

Esta dissertação contribui para o desenvolvimento de *websites* desportivos, demonstrando a viabilidade e os benefícios de processos automatizados e de dados acessíveis ao público. Permite que os clubes reforcem as relações com os adeptos, envolvam as partes interessadas e promovam o sucesso financeiro no mundo digital atual.reconhecer que pertencem a instituições diferentes e se acham que foram ambos gerados usando a mesma plataforma.

Palavras-Chave: Dados, *Website*, Customizável, Geração Automática, *UX/UI*

Classificação ACM: *Human-centered computing → Human-computer interaction (HCI) → Interaction paradigms → Web-based interaction*

Abstract

In today's digital era, websites have become valuable sources of information, frequently updated with comprehensive data from journalists, certified individuals, and crowd contributions. Football clubs understand the importance of maintaining up-to-date websites to engage fans and investors, generate revenue, and enhance visibility. However, there is a digital divide. On the one hand, higher-ranking clubs benefit from more comprehensive websites. In contrast, smaller clubs face limitations due to resource constraints and outdated platforms, underscoring the need for updated and data-driven websites, irrespective of the club's size or ranking.

This Dissertation proposes an innovative solution to bridge the divide by developing an automated platform that uses publicly available data to generate up-to-date, user-friendly, and customisable websites for local sports teams. The objective is to provide a practical and cost-effective means for clubs to create and maintain websites that reflect their brand identities while staying current with the evolving nature of the sport.

The validation approach involves manual verification, comparing timestamps with official club websites, and gathering supporters' feedback to assess the generated websites' currency. User-centric evaluations and the System Usability Scale [8] gauge usability, while manual assessments examine the ease of generation and visual appeal.

The empirical evaluation yielded positive results regarding the generated websites' up-to-dateness, usability, and visual appeal. Supporters from lower-ranking and first-league clubs acknowledged the user-friendliness and currency of the generated websites, surpassing the official ones. Furthermore, they expressed a preference for the developed websites due to their enhanced usability. The application of the System Usability Scale further provided quantitative evidence supporting the excellent usability of the generated websites.

The findings affirm the hypothesis, highlighting benefits such as improved communication, reduced maintenance efforts, enhanced user experience, and reinforced brand identity. The proposed solution bridges the digital divide, ensuring equal opportunities for all clubs and fostering sports transparency and growth.

This Dissertation contributes to sports website development by showcasing the feasibility and benefits of automated processes and publicly accessible data. It empowers clubs to strengthen relationships with supporters, engage stakeholders, and drive financial success in today's digital world.

Keywords: Data, Website, Customisable, Automatic Generation, UX/UI

ACM Classification: Human-centered computing → Human-computer interaction (HCI) → Interaction paradigms → Web-based interaction

Acknowledgements

First and foremost, I want to thank my supervisor, André Restivo, for his guidance and support during the most challenging moments of this past year. Your mentorship was crucial to the successful completion of this Dissertation. Additionally, I am grateful to *ZOS* personnel, particularly Pedro Dias, Marco Sousa, and Fernando Pereira, for providing me with a suitable space to work on my project and for their invaluable assistance in implementing the proposed solution.

Next, I want to thank my "Turtles" colleagues who have accompanied and supported me since beginning my academic journey at FEUP. Your presence and assistance whenever I needed it made my college life genuinely fantastic. I will forever cherish the memories of Taco Tuesdays, quarantine gaming nights, and "Quintas Mágicas".

I want to extend my gratitude to my three best friends, Guilherme, Inês, and Luana, fondly known as "Carpetes." We have known each other for many years, and you have been there by my side since the beginning of my academic journey, celebrating all of my victories. Your advice during the most challenging moments has often been the guiding light in my darkest days, and I am eternally grateful to each of you. The memories of our trips, adventures, and "sueca" afternoons will always hold a special place in my heart. Thank you for being such amazing friends.

I want to express my heartfelt gratitude to my boyfriend, Tiago, for his unwavering support throughout this long and challenging journey of writing my dissertation. Without your constant support, advice, and lightheartedness, I would not have been able to reach the point where I am today. Thank you for being my rock and for providing me with love, inspiration, and motivation every step of the way.

Furthermore, I want to thank my family, especially my parents and brother, for providing me with the means and support necessary to thrive academically, especially my mother, who showed me the importance of perseverance, even in the most challenging times.

Lastly, I want to express my gratitude to the remaining family members, friends, and professors. This accomplishment would not have been possible without your support, encouragement, and guidance.

Beatriz Mendes

"What is this brief mortal life,
if not the pursuit of legacy?"

Lord Corlys Velaryon

Contents

List of Figures

List of Tables

Abbreviations and Symbols

API	Application Programming Interface
DOM	Document Object Model
HTTP	HyperText Transfer Protocol
RDBMS	Relational Database Management System
SUS	System Usability Score
SWAM	Sport Website Acceptance Model
SWQ	Sports Website Quality

Chapter 1

Introduction

In this chapter, we will delve into several key aspects that form the foundation of this Dissertation.

Firstly, starting in Section 1.1, we will explore the evolution of the Internet and its profound impact on businesses, examining how organisations have adapted to the ever-growing digital landscape. Moreover, we will specifically focus on the relevance of websites for sports clubs, particularly small clubs. The motivation driving this research stems from recognising websites as crucial assets for sports clubs and the unique challenges smaller clubs face in creating high-quality online platforms (Section 1.2 (p. 2)). We will address the problem at hand in Section 1.3 (p. 3), shedding light on the factors contributing to the difficulties encountered by small clubs in establishing websites of sufficient quality. In Section 1.4 (p. 3), we specify this Dissertation's goal, which consists of a comprehensive solution tailored to the specific needs of small clubs, bridging the gap and empowering them to develop and maintain websites of exceptional quality. Finally, we will provide a comprehensive overview of the Dissertation's structure in Section 1.5 (p. 4), offering readers a clear roadmap to navigate the subsequent chapters and cohesively explore the research.

1.1 Context

The Internet has revolutionised how we access and share information, impacting various aspects of our lives. With just a few clicks, we can explore websites that offer a wealth of data, making them indispensable for netizens seeking credible and comprehensive information. Platforms like

IMDb[1] and Wikipedia[2] exemplify this with their vast databases and frequent updates by journalists and certified individuals. Additionally, crowdsourcing allows users to contribute and verify data, tapping into the collective knowledge of online communities [21].

As the Internet continues to expand, its global user base has reached approximately 5.16 billion people, representing a significant portion of the world's population. Over the past year, internet users have grown by 1.9% [28]. Portugal, in particular, has witnessed pervasive internet adoption, with 8.73 million users, accounting for 85.1% of the country's population [29]. These statistics highlight the widespread influence of the Internet, shaping how we communicate, access information, and connect with others.

Beyond information access, the Internet has transformed the innovation and entrepreneurship landscape. The emergence of powerful digital technologies, platforms, and infrastructures has revolutionised value creation and capture across various sectors. Businesses have unprecedented opportunities to leverage digital advancements, explore new markets, and establish competitive advantages. Entrepreneurs can harness the digital ecosystem to develop disruptive business models, utilise data-driven insights, and engage with a global customer base. Consequently, the Internet has become a catalyst for innovation, redefining the dynamics of the modern business landscape [35].

In the context of small businesses, having an effective website is a crucial step towards embracing e-commerce. Research suggests that a clear online strategy, supported by a well-designed website, increases the likelihood of transitioning to e-commerce. Unfortunately, many small business owners overlook their web presence when formulating their business strategy. Recognising the transformative potential of the Internet and its digital advancements, small business owners need to prioritise building a robust online presence aligned with their strategic goals. A strategically designed website serves as a gateway to e-commerce, enabling businesses to expand their reach, attract customers, and fully tap into the vast opportunities offered by the digital marketplace [19].

Just as small businesses need a website, sports clubs are no exception, especially considering that they can be considered businesses themselves [16, 36]. Today, it has become standard practice for many football clubs to maintain a regularly updated website. Dedicated staff or team managers typically carry out this task, ensuring that fans and investors are informed about the team's latest news and events [22]. Furthermore, having a website provides an additional avenue for clubs to generate revenue through ticket sales and merchandise [36] while enhancing their visibility.

1.2 Motivation

In the modern era, a website has become an indispensable tool for the successful operation of football clubs, particularly for smaller clubs that rely on additional sources of revenue and visibility to thrive. More than ever, these clubs recognise the importance of a well-designed website. It serves

[1] https://www.imdb.com/
[2] https://www.wikipedia.org/

as their lifeline, offering a platform to engage with supporters and sponsors and provide timely updates and news about the team.

For smaller clubs, a website is a means of communication and a lifeline for their growth and development. It levels the playing field, enabling them to compete on an equal footing with larger clubs by reaching a wider audience and expanding their fanbase. With limited resources, smaller clubs often struggle to secure sponsors and generate revenue. However, a well-designed website opens doors to new opportunities. It allows them to showcase their talents, attract potential sponsors, and explore innovative income-generating ways.

Moreover, a sports website becomes a vital revenue-generating channel for smaller clubs. These clubs can tap into a global market by selling tickets and merchandise online, extending their reach beyond their local fanbase. The website provides fans a convenient and secure online shopping experience, ultimately driving higher sales and increasing revenue streams.

1.3 Problem

Well-established football teams have the resources and financial means to invest in custom website development, enabling them to establish a professional online presence. These teams can allocate a dedicated budget for website management and employ individuals specifically responsible for website development, design, content management, and regular updates. With a team of professionals, these clubs can ensure that their website remains updated with the latest news, match results, player profiles, and other relevant information. They can also leverage advanced technologies and design elements to enhance the user experience and engage fans on a deeper level.

In contrast, lower-ranking football teams often face resource limitations, including budget constraints. Consequently, they may need help to invest in custom website development or hire professionals for website management. Instead, these smaller teams may rely on volunteers with the technical skills and knowledge to design, develop, and maintain their websites. However, the website may not receive regular updates and maintenance due to limited resources and the need for volunteer assistance, leading to outdated content and a bad user experience.

Furthermore, smaller teams may need more technical expertise and experience to create visually appealing and user-friendly websites. As a result, their websites have poor aesthetics, navigation difficulties, and limited functionality.

1.4 General Goal

Our thesis aims to create an automated web generation tool for low-ranking football teams, utilising publicly available data to update their websites while automatically maintaining the team's identity. The primary goal is to develop a user-friendly, cost-effective, efficient, and scalable solution tailored to teams of any size.

By harnessing publicly available data, we believe it is possible to automate the generation of up-to-date and user-friendly websites for local sports teams. This approach offers an efficient and cost-effective solution, enabling teams to create and maintain websites that remain current with the latest news and games. Furthermore, the tool will prioritise excellent usability and customisation, ensuring each team's website reflects its unique brand identity.

Achieving our objective would lead to the realisation of an efficient and cost-effective solution, empowering local sports teams to establish and maintain websites that consistently showcase the latest news and games. By adopting an automated approach, the management of websites would be streamlined, eliminating the need for manual updates and freeing up teams to concentrate on other operational aspects.

Moreover, successful implementation would strengthen the online presence of low-ranking clubs, enhance audience engagement, and elevate the visibility and recognition of their club. This tool can level the playing field and provide equal opportunities for all teams to establish a solid online presence by bridging the digital divide between first-league and low-ranking clubs.

1.5 Document Structure

The subsequent chapters of this dissertation are organised as follows:

- **State of the Art** (Chapter 2 (p. 5)): Covers website quality, including sports website quality and aesthetic quality, as well as the use of external data in web development and organisations' online presence;
- **Problem Statement** (Chapter 3 (p. 15)): Addresses the problem under study, mainly focusing on Portugal Clubs' Website Analysis, and presents the hypothesis, research questions, solution proposal, and methodology;
- **Implementation** (Chapter 4 (p. 27)): Focuses on the implementation of the solution, covering various aspects such as the tools and technologies used, the implemented architecture, the utilisation of *zerozero* endpoints for real-time data updates on club websites, the website generator tool employed, and the structure of the websites including implemented pages and components;
- **Empirical Evaluation** (Chapter 5 (p. 49)): Outlines the objectives of the validation, questionnaires used for assessment, preliminary assessment, results for low-ranking clubs and first-league clubs, discussion, and validity threats;
- **Conclusions** (Chapter 6 (p. 77)): Summarises the conclusions, contributions and potential future improvements;

Chapter 2

State of the Art

This chapter provides an overview of the current state of the art in website creation. Since we have yet to find much information specifically regarding website creation, we analysed literature regarding related fields.

Section 2.1 examines how website quality impacts user engagement, brand perception, and overall success in the digital realm, emphasising the significance of visually appealing, user-friendly, and technologically advanced websites. Section 2.2 (p. 9) delves into incorporating external data sources in web development, exploring the benefits and challenges of leveraging publicly available data to enhance website functionality and efficiency. Additionally, in Section 2.3 (p. 9), we investigate the digital challenges faced by small businesses and sporting teams, considering existing case studies and research to gain insights into strategies for establishing a robust online presence.

2.1 Website Quality

There is no official definition of website quality, as it can differ depending on the purpose and target audience of the website. Nevertheless, a generally accepted definition of *website quality* is the degree to which a website meets its goals and provides a positive user experience for its audience [38].

For this research, we will study two fields regarding website quality: sports website quality and the aesthetic quality of a website.

2.1.1 Sports Website Quality (SWQ)

Sports websites provide fans essential team information, such as news, scores, statistics, and player profiles. They offer a crucial platform for strengthening ties between the team and its supporters and generating extra revenue from selling goods and tickets. Given its significance, the effectiveness of these websites becomes crucial since it can affect user intentions and behaviours, which in turn affects the team's success as a whole.

This subsection comprehensively reviews relevant literature to examine existing studies and identify significant findings thoroughly. The findings will be succinctly summarised and presented, emphasising key takeaways and their relevance to the sports website's quality topic.

Weisheng Chiu *et al.* [12] Building a better website is crucial in attracting consumer attention, especially in the sports industry, where large amounts of information are consumed. A comprehensive assessment of the quality of sports websites is necessary to fulfil user needs, as they are usually informative and multi-functional. The five quality dimensions defined are Information quality, Interaction quality, Design quality, System quality, and Fulfillment quality. They are crucial in determining a consumer's perception of the website. Fulfilment quality, which measures consumer enjoyment, fun, and pleasure, should be regarded as the top priority for a sports website as it directly affects consumer preference to revisit. Sports organisers need to create an enjoyable environment for fans, and a visually appealing website design is crucial in retaining and attracting more people.

Muge Elden *et al.* [17] Websites are essential in sports marketing as they are used to persuade consumers. Persuasive websites must have quality elements such as informativeness, usability, credibility, visual design, involvement, and reciprocity. Usability criteria like up-to-dateness, navigation, and understanding are essential. However, credibility criteria such as the club history and "About Us" sections may be lacking. Sports marketing websites, to increase engagement, should have interactive features and be improved in the "feedback", "newsletter", and "online brochure" subdimensions since it will ensure the continuity of the corporate brand in the user's mind.

Youngjin Hur *et al.* [23] Online sports consumption is growing, and fans visit websites for information, merchandise, and interaction with other fans and fantasy leagues. Key variables include SWQ, satisfaction, and loyalty. In SWQ, their dimensions, such as information, interaction, design, system reliability and fulfilment, shape fans' perceptions and satisfaction. High e-satisfaction leads to increased website visits and page views. E-loyalty and commitment to continue using a site are crucial for long-term profitability and word-of-mouth referrals. Research shows that e-satisfaction fully mediates the relationship between SWQ

and e-loyalty, with high quality leading to increased loyalty. However, even a visually appealing site may not be satisfying if it does not have information that is not interesting to the users.

With the analysis of the related literature mentioned above, we can affirm that the quality of sports websites is crucial for successful sports marketing and engaging with fans [12, 17]. Persuasive elements such as informativeness, usability, credibility, visual design, involvement, and reciprocity are essential for building a high-quality website. Including interactive features and improved feedback, newsletter, and online brochure subdimensions also play a role in increasing engagement [17]. The overall quality of a sports website can be determined by five key dimensions: Information quality, Interaction quality, Design quality, System quality, and Fulfilment quality. Fulfilment quality is considered the top priority, directly affecting consumer preference to revisit the website [12]. With the growth of online sports consumption, website quality, satisfaction, and loyalty are becoming increasingly vital variables in the sports industry [23]. Research has shown that e-satisfaction fully mediates the relationship between website quality and e-loyalty, emphasising the importance of maintaining high website quality for long-term success [23].

2.1.2 Aesthetic Quality

The Interaction Design Foundation has the following definition for the concept of aesthetics [1]:

> "Aesthetics is a core design principle that defines a design's pleasing qualities. In visual terms, aesthetics includes factors such as balance, colour, movement, pattern, scale, shape and visual weight. Designers use aesthetics to complement their designs' usability, and so enhance functionality with attractive layouts."

Aesthetic quality is a critical aspect of website design that significantly impacts user experience. A well-designed website needs an appealing and memorable visual experience for the user to provide valuable information and functionality. In this subsection, we will explore the impact of aesthetic quality on user behaviour and how they contribute to website quality.

This subsection will examine relevant literature and emphasise the crucial discoveries.

Farah Alsudani *et al.* [6] Promote web credibility by enhancing one factor that affects it: aesthetics. The interface of a web page, and ultimately a website, can be considered an artistic composition. To achieve a more appealing web interface and produce better immediate judgments on web credibility, web designers must apply the concept of "Unity". This concept can be achieved using balance, harmony, contrast and dominance design principles. Additionally, she concluded that "Unity" affects users' immediate judgments on web credibility and attracts their attention, encouraging them to stay on a website. In summary, improving the aesthetics of a website can lead to an increase in web credibility.

Kiemute Oyibo *et al.* [37] Designing a successful website in human-computer interaction requires balancing aesthetics, usability, and credibility. Aesthetic design impacts user attitudes more than usability, and website credibility is based on trustworthiness and design expertise. Gender differences affect the impact of aesthetics on usability, and females are more critical in evaluating websites. Designers must balance hedonic and utilitarian concerns and consider gender differences and the importance of visual aesthetics.

Lisbeth Thorlacius *et al.* [41] Aesthetics play a vital role in the successful design of websites, including supporting content and functionality, appealing to the target audience's taste, creating desired sender images, and addressing website genre requirements. User-friendly web design has five components: learnability, efficiency, memorability, error handling, satisfaction, and understandability. The quality of a website is evaluated based on these components and their functionality. Young people prefer visual communication, making it essential for web designers to create aesthetically pleasing designs that align with the sender's image, functionality, genre, and target audience.

Based on the aforementioned related literature analysis, we can confirm that it is essential to balance aesthetics, usability, and credibility in website design [17, 37]. Aesthetics significantly impact user attitudes, making it an essential aspect of website design [37]. Gender differences and hedonic and utilitarian concerns should also be considered when designing a website [37]. Enhancing aesthetics can improve web credibility through "Unity", achieved through design principles such as balance, harmony, contrast, and dominance [6]. The aesthetic design supports content and functionality, appeals to the target audience, creates desired sender images, and addresses website genre requirements [41]. Web design should be evaluated using five key components: learnability, efficiency, memorability, error handling, satisfaction, and understandability to ensure a website is user-friendly [41]. Considering these factors, a successful website design can be achieved that balances aesthetics, usability, and credibility to provide a positive user experience.

In conclusion, the quality of sports websites is vital for sports marketing success and fan engagement. A high-quality sports website should incorporate persuasive elements such as informativeness, usability, credibility, visual design, involvement, reciprocity, and interactive features. The overall quality of a sports website can be evaluated based on five key dimensions, with Fulfilment quality being considered the top priority. With the increasing growth of online sports consumption, website quality, satisfaction, and loyalty are becoming crucial variables in the sports industry. Maintaining high website quality is crucial for long-term success, as e-satisfaction fully mediates the relationship between website quality and e-loyalty. In website design, it is essential to balance aesthetics, usability, and credibility and consider gender differences and hedonic and utilitarian concerns. A successful website design should be user-friendly and evaluated based on five key components, with aesthetics significantly impacting user attitudes. Considering these factors, a successful website design that provides a positive user experience can be achieved.

2.2 Use of External Data in Web Development

Given the large amount of data accessible online, several websites are using external sources to enrich their databases. Web scraping enables organisations to harvest data from other websites without needing APIs, or code is one of the most well-known techniques for achieving this task. Nevertheless, there are moral and legal issues with this convenience that need to be taken into account.

We will examine numerous articles to comprehend the effects of web scraping on web development and review the existing technologies that allow this technique. We will highlight the most important discoveries and summarise the data.

Vlad Krotov *et al.* [33] Using web scraping to extract data from websites provides many opportunities for researchers and practitioners. However, it is vital to consider the legal and ethical implications of such actions. For this reason, reviewing the legal and ethics and privacy literature is recommended to identify areas of concern and questions that must be addressed. Asking questions about potential harm and how to prevent it helps ensure that web scraping is ethical and legal.

Rabiyatou Diouf *et al.* [15] Web scraping can be done using various methods, including earlier identified data location (mimicry), weight measurement of words in each branch of a web page's DOM tree (weight measurement), comparing differences in content between two pages of the same website (differential), and training algorithms with large samples of manually analysed web pages (machine learning). Regardless of the approach used, the ultimate goal of web scraping is to capture web data and present it in a more structured format. Some of the most common application areas for web scraping include job search engines, recommendation systems, advertising, health, and journalism. Despite its wide range of applications, web scraping remains most needed in journalism and is often performed using fewer specialised tools than other sectors. Regarding the tools, browser extensions and software, and platforms can be used to perform the web scraping tasks.

In summary, web scraping is a practical technique for harvesting and processing data from websites with various applications across multiple industries [33], especially journalism [15]. However, it is crucial to consider its legal and ethical implications [33]. Web scraping can be performed using various platforms, with specialised tools needed [15]. As web scraping continues to grow, it is crucial to balance the benefits of data collection with the need to maintain ethical and legal standards [33, 15].

2.3 Organisations' Online Presence

To succeed in the digital era, sporting organisations and small companies need a solid online presence. The Internet and social media platforms have created new avenues for organisations like

corporations and clubs to advertise. Sports leagues and small businesses need a strong web presence to succeed in the digital age. Thanks to the Internet and social media platforms, organisations like businesses and clubs have more ways to promote. A well-managed internet presence may also assist small businesses and sports teams in expanding their markets, gaining new clients and supporters, and keeping their current ones. As a result, small companies and athletic organisations need to comprehend the essential elements that affect the effectiveness of their online presence and create plans that effectively exploit the potential provided by digital media.

This section examines the various factors that lead to small enterprises' and athletic clubs' success, the difficulties they encounter and the best practices for in-depth analysis of the research that has already been done on the topic.

2.3.1 Small Businesses

A robust online presence in today's digital age is crucial for small businesses. Almost everyone in the world has internet access and is slowly turning to online platforms to research and purchase products and services, so it is essential for small businesses to keep up with this transition. In order to have the best online presence possible, it is crucial to research to decide the next steps for a small business to improve its channels.

In this subsection, we aim to analyse an article that explores a new model for small businesses to use to improve their online presence, followed by some highlights.

Stephen Burgess [9] The new model is proposed to represent small businesses' web presence better. The "web presence pyramid model" (Figure 2.1 (p. 11)) not only considers how a small business may have an online presence through its website, third-party web portals, business directories, and social media but also recognises the non-linear nature of these presences. The model classifies the web presence of small businesses into four categories: Basic Web Presence, Added Value Websites, Online Database Websites, and Payment Web Presence, with a core set of features such as contact details, business information, and ways to communicate with the business being present across most of these presences.

In summary, to better represent small businesses' web presence, a new web presence pyramid model has been proposed by Stephen Burgess. This model categorises small business web presence into four categories: Basic Web Presence, Added Value Websites, Online Database Websites, and Payment Web Presence. These categories share a core set of features, including contact details, business information, and ways to communicate with the business.

2.3.2 Sport Clubs

As we verified in the previous subsection, a solid online presence is crucial for small businesses and applies to small sports clubs [16]. A website allows the club to reach a bigger audience and provide comprehensive information about the club that cannot be found easily on social media. The online presence will increase the club's visibility, reputation, and credibility. Additionally,

Web Presence Pyramid Model

	E-services	Distribution	D	Distribution	E-services	
	Payment	Online database	T	Online database	Payment	
Social networking	Reviews	Contact the Business	C	Contact the Business	Reviews	Social networking
Added Value	Business information	Contact details	I	Contact details	Business information	Added Value
Website features			**Virtual Spaces**	**Third Party Website features**		

Figure 2.1: Web Presence Pyramid Model proposed by Stephen Burgess [9]

having an online presence allows fans and supporters to engage with the club, making it easier for the club to establish a connection with their target audience and gather valuable feedback that can be used to improve the club's operations.

This subsection will analyse relevant related work and expose the key findings.

Stephen Burgess *et al.* [10] The comparison between upper-level and lower-level sporting clubs have revealed the existence of a digital divide between them in terms of Internet technologies. While most local sporting clubs rely on volunteers to keep them running, upper-level clubs are likelier to adopt club-initiated websites and social media sites, using these platforms' information, transactions and communications space. On the other hand, lower-level clubs often need help to keep up with the demands of a digital presence, needing more time, finances, and expertise to exploit the potential of these technologies effectively. As a result, there is a need for different strategies to support the volunteers in lower-level clubs, enabling them to use these technologies with minimal time and effort, bridging the digital divide in the sports community.

Stephen Burgess *et al.* [11] There is a digital divide between upper- and lower-level clubs. Upper-level clubs are likelier to adopt club-initiated websites and social media, whereas lower-level clubs often need more resources. It is crucial to consider the unique needs of each club and avoid the assumption that all local sporting clubs have the same level of Internet adoption and usage. The results suggest that local clubs without a governing body website should consider having a club-initiated website to provide basic information to potential members and that more studies are needed to understand club volunteers' use of Internet-based applications fully.

Youngjin Hur *et al.* [22] A sports website in the study contains sport-related products or services where fans can access game results and player statistics, purchase team merchandise, and share their opinions. The author developed a SWAM (Figure 2.2 (p. 12)) to address the gap in understanding the factors influencing fans' adoption of sports websites. The model

considers the critical variables of perceived ease of use, usefulness, enjoyment, and trustworthiness as the basis for beliefs that drive the intention to use the websites. Additionally, it proposes that sports involvement and psychological commitment to a team can affect behavioural intention in sports website consumption. The model aims to comprehensively explain how fans perceive and accept sports websites and how these factors influence their use.

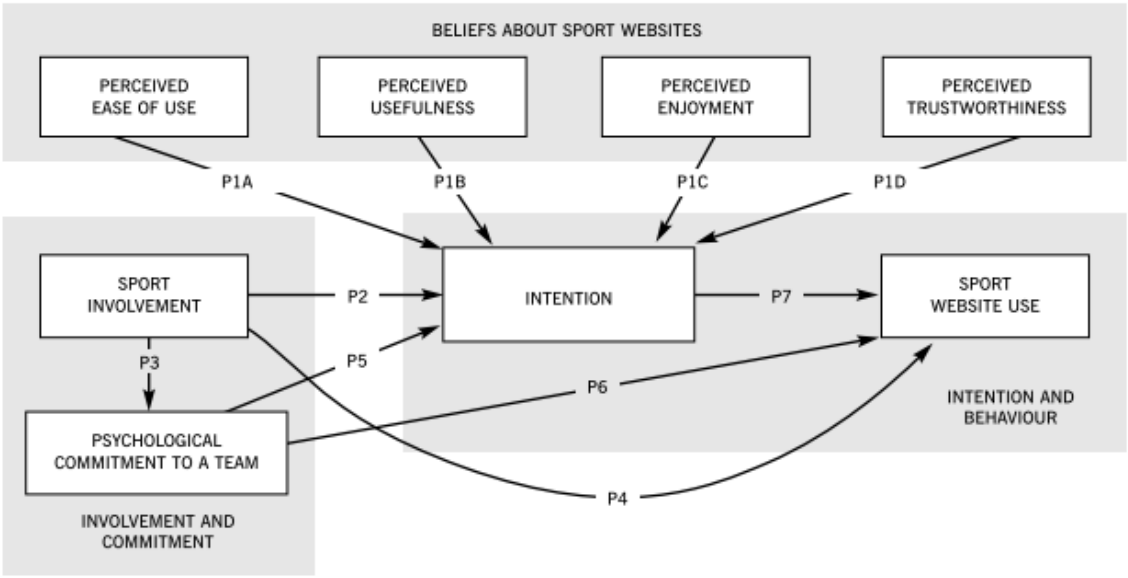


Figure 2.2: The SWAM proposed by Youngjin Hur *et al.* [22]

In conclusion, the digital divide between upper- and lower-level sporting clubs in adopting Internet technology is evident. Upper-level clubs have more resources to adopt club-initiated websites and social media. In contrast, lower-level clubs need different strategies to support their volunteer members in utilising these technologies [10, 11]. The SWAM has been developed to understand the factors influencing fans' adoption of sports websites. The model considers critical variables such as perceived ease of use, usefulness, enjoyment, and trustworthiness as the basis for beliefs that drive the intention to use sports websites. Additionally, sports involvement and psychological commitment to a team can shape behavioural intention in sports website consumption. The SWAM model aims to comprehensively explain how fans perceive and accept sports websites and how these factors influence their use [22].

This section focused on the importance of a web presence for small businesses and understanding the factors influencing fans' adoption of sports websites. The new web presence pyramid model proposed by Stephen Burgess helps to better represent small businesses' web presence by categorising it into four categories - Basic Web Presence, Added Value Websites, Online Database Websites, and Payment Web Presence. All categories share core features such as business information, contact details, and ways to communicate with the business. The digital divide regarding

sporting websites highlights the need for different strategies to support the latter in adopting Internet technology. The SWAM provides insights into fans' adoption of sports websites and the factors influencing their usage, such as perceived ease of use, usefulness, enjoyment, trustworthiness, sports involvement and psychological commitment to a team.

2.4 Summary

In conclusion, this state-of-the-art chapter explored the importance of high-quality sports websites (Section 2.1 (p. 5)), the use of external data in web development (Section 2.2 (p. 9)), and the web presence of small businesses and local sports teams (Section 2.3 (p. 9)).

High-quality sports websites should incorporate persuasive elements such as informativeness, usability, credibility, and interactive features. They can be evaluated based on five critical dimensions, with Fulfillment quality being the top priority.

External data is used in web development mainly through web scraping, a practical data collection technique, but its legal and ethical implications must be considered.

The web presence pyramid model proposed by Stephen Burgess categorises small businesses' web presence into four categories. The SWAM provides insights into fans' adoption of sports websites. The chapter highlights the need to balance aesthetics, usability, and credibility in website design and support lower-level sporting clubs adopting internet technology to bridge the digital divide.

Chapter 3

Problem Statement

This chapter introduces the problem statement, focusing on the digital divide in Portugal's sports sector. An extensive literature review shows a significant disparity between higher-ranking and lower-ranking teams regarding their digital capabilities [10, 11]. This divide stems from the contrasting levels of resources available to these two categories of teams, particularly concerning the creation and maintenance of modern digital platforms. Well-established teams, bolstered by substantial financial resources and organisational support, have the means to develop and sustain sophisticated digital platforms that cater to the needs of their fanbase and stakeholders. They can invest in professional website development, dedicated staff, and advanced technologies, ensuring their online presence remains cutting-edge and engaging.

In Section 3.1 (p. 16), we delve into a comprehensive analysis of the digital divide, explicitly examining its various manifestations within the sports industry in Portugal. By understanding the intricacies of this divide, we can gain insights into the challenges faced by lower-ranking teams and the opportunities for improvement. Building upon this understanding, Section 3.2 (p. 21) explores the impact of the digital divide on sports clubs, shedding light on its implications for their online presence and overall competitiveness. With a clear focus on addressing this divide, we present our research questions in Section 3.3 (p. 21), which aim to uncover potential solutions

and strategies for bridging the gap. Furthermore, we present our hypothesis, a guiding principle for our research endeavours. Section 3.4 (p. 23) then outlines the proposed solution, offering an overview of the approach we aim to implement to address the digital divide and empower lower-ranking teams. Finally, in Section 3.5 (p. 24), we outline the methodology employed to assess our proposed solution's effectiveness, ensuring that our research efforts yield meaningful insights and practical recommendations. By establishing this comprehensive framework, we pave the way for the subsequent chapters, which will delve into a detailed analysis of each section's subject matter, contributing to advancing knowledge in the field of sports club websites.

3.1 Portugal Clubs' Website Analysis

In order to examine the digital divide prevalent among football clubs in Portugal more in-depth, a comprehensive analysis was conducted using the data collected by Mascarenhas [34]. Mascarenhas examined clubs' websites across different rankings in Portugal, including the *I Liga*, *II Liga*, *Liga 3*, *Campeonato de Portugal*, and *Distrital*.

The analysis focused on various aspects of these websites, encompassing their existence, secure connection, and functionality. Specifically, Mascarenhas assessed whether the websites were functioning or not, and on the functional websites, he further investigated the presence of four key sections, namely "News", "Next Games", "Latest Results", and "Players." Additionally, Mascarenhas delved into the recency of updates of each website. He categorised them into three different timeframes: updates within the previous seven days, updates between seven and thirty days ago, and updates exceeding thirty days.

Mascarenhas devised a rating scale ranging from 0 to 9 to facilitate the classification of the websites. Points were allocated based on specific criteria as follows:

- He awarded websites with a **secure connection** 1 point, while the absence of a secure connection yielded 0 points;
- The presence of the **"Next Games" section** garnered 1 point, while its absence resulted in 0 points;
- Likewise, the presence of the **"Latest Results" section** earned 1 point, whereas its absence led to 0 points;
- The existence of the **"News" section** accumulated 1 point, while its absence yielded 0 points;
- Similarly, the presence of the **"Players" section** was assigned 1 point, while its absence resulted in 0 points;
- The **level of up-to-dateness** was taken into consideration, with websites classified as "Outdated" receiving 0 points, those deemed "Not Recently Updated" (updated in the last thirty

days) obtaining 1 point, and those confirmed as "Updated" (updated in the last seven days) acquiring 2 points;

- The **aesthetics level** of the websites was also evaluated, with websites categorised as "Bad" earning 0 points, those considered "Could be better" receiving 1 point, and those perceived as "Good" accumulating 2 points.

The scores of the existing websites are presented in Table 3.1.

Table 3.1: Scores obtained by the existing websites after from *I Liga*, *II Liga*, *Liga 3*, *Campeonato de Portugal* and *Distrital* their analysis

Ranking		**Score**			
		0	**1-3**	**4-6**	**7-9**
I Liga	nº	1	0	0	17
	%	5.56%	0.00%	0.00%	94.44%
II Liga	nº	1	0	8	7
	%	6.25%	0.00%	50.00%	43.75%
Liga 3	nº	2	1	6	6
	%	13.33%	6.67%	40.00%	40.00%
Campeonato de Portugal	nº	18	15	5	7
	%	40.00%	33.33%	11.11%	15.56%
Distrital	nº	277	120	56	28
	%	57.59%	24.95%	11.64%	5.82%
Total	nº	299	136	75	65
	%	52.00%	23.65%	13.04%	11.30%

Observing the rankings, we can see that most websites from the *I Liga* (94.44%) and *II Liga* (43.75%) obtained scores in the range of 7-9, indicating a relatively higher level of quality and functionality. In contrast, websites from *Liga 3* and *Campeonato de Portugal* had a more varied distribution of scores. While a significant proportion of websites in *Liga 3* (40.00%) achieved scores in the range of 4-6, the *Campeonato de Portugal* websites showed a relatively balanced distribution across the different score ranges.

Notably, the websites from the *Distrital* division had the lowest scores overall. Over half of the websites in this division (57.59%) received scores of 0, indicating a lack of functionality or minimal web presence. Additionally, many websites in the *Distrital* division (24.95%) achieved scores in the range of 1-3, reflecting a relatively limited level of quality and functionality.

These results indicate an apparent discrepancy in website quality and functionality between higher-ranking divisions (*I Liga* and *II Liga*) and lower-ranking divisions (*Liga 3*, *Campeonato de Portugal*, and *Distrital*). While the websites of higher-ranking teams generally exhibited better scores and more comprehensive features, there is room for improvement across all divisions, particularly in the lower-ranking ones.

3.1.1 Websites' Functionality/Existence

To understand Mascarenhas's research findings comprehensively, we first analysed the presence and functionality of websites among football clubs. In The Table 3.2, we have a representation of the collected data.

Table 3.2: Status of the websites from the clubs from *I Liga*, *II Liga*, *Liga 3*, *Campeonato de Portugal* and *Distrital*

Ranking		**No Website**	**Website**		**Total**
			Not Functioning	**Functioning**	
I Liga	nº	0	1	17	18
	%	0.00%	5.56%	94.44%	
II Liga	nº	2	1	15	18
	%	11.11%	5.56%	83.33%	
Liga 3	nº	10	2	13	25
	%	40.00%	8.00%	52.00%	
Campeonato de Portugal	nº	11	17	28	56
	%	19.64%	30.36%	50.00%	
Distrital	nº	415	136	345	896
	%	46.32%	15.18%	38.50%	
Total	nº	438	157	418	1013
	%	43.24%	15.50%	41.26%	

The data shows again that the digital divide among Portuguese football clubs is quite significant. Most clubs in higher-ranking leagues, such as *I Liga* and *II Liga*, have functioning websites, with 94.44% and 83.33% functionality rates, respectively. However, it is worth noting that a small percentage of clubs in these rankings still need a website, indicating a potential gap in their digital presence.

Moving down the rankings, the proportion of clubs without websites increases significantly. In *Liga 3*, 40% of the clubs still need a website, and in *Distrital*, nearly half of the clubs (46.32%) do not have a website. This finding indicates a substantial disparity in online representation, as these clubs may need help reaching a wider audience and showcasing their achievements and events. Additionally, the relatively low percentage of functioning websites (38.50%) among *Distrital* clubs suggests that even those with an online presence have room for improvement in website functionality and user experience.

3.1.2 Websites' Features

Subsequently, we conducted a more detailed analysis of the data collected by Mascarenhas regarding the sections/features available on the websites of Portuguese clubs across different rankings. The only sections analysed were "Next Games", "Latest Results", "News", and "Players". Figure 3.1 (p. 19) provides an overview of the distribution of these sections' presence on club websites.

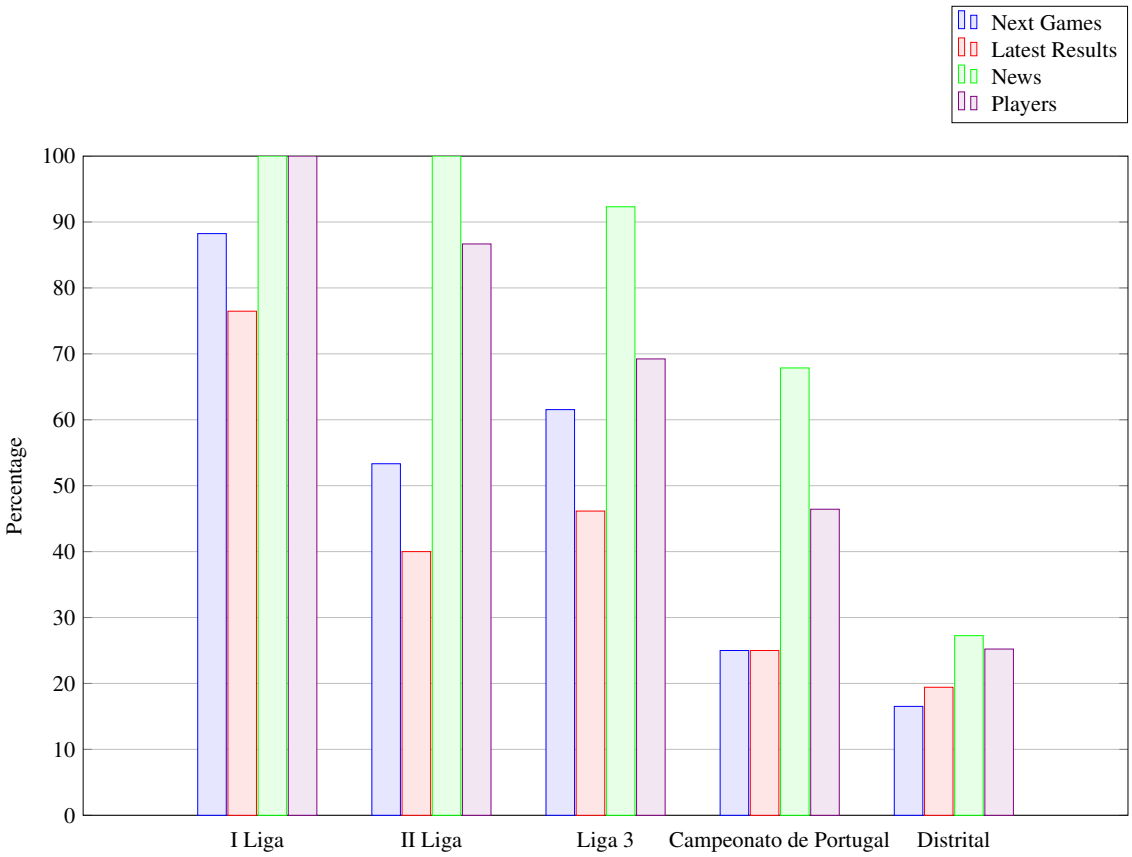


Figure 3.1: Percentage of functioning websites from the clubs from *I Liga*, *II Liga*, *Liga 3*, *Campeonato de Portugal* and *Distrital* that have sections for Next Games, Latest Results, News and Players

The analysis of the results proves the existence of variations in the presence of website sections among football clubs of different rankings. The *I Liga* demonstrates the highest percentage of functioning websites with sections for "Next Games" (88.24%) and "Latest Results" (76.47%). Additionally, 100% of functioning websites in the *I Liga* and *II Liga* have sections for "Players". On the other hand, the *Distrital* shows the lowest percentages for all sections, with 16.52% for "Next Games", 19.42% for "Latest Results", and 25.22% for "Players".

Additionally, it can be concluded that the most common section among club websites is "News". Conversely, the least recurring section overall is "Latest Results", often limited to displaying the outcome of the most recent match rather than encompassing the entire history.

Notably, the most significant disparity observed in these graphs is the evident difference in the number of sections between clubs of higher and lower rankings, reinforcing the presence of a digital divide. These sections play a crucial role in the existence of club websites, as they serve as platforms for sharing information and establishing closer connections with fans and potential investors.

3.1.3 Websites' Level of Up-to-dateness

Finally, Mascarenhas analysed each of the functional websites to determine their last update dates and categorised them into three groups: "Outdated" (not updated for more than thirty days), "Not Recently Updated" (updated more than seven days but less than thirty days ago), and "Updated" (updated within the last seven days). The data collected by Mascarenhas was organised in Table 3.3 for easier comparison among different rankings.

Table 3.3: Analysis of the last update on the websites from clubs from *I Liga*, *II Liga*, *Liga 3*, *Campeonato de Portugal* and *Distrital*

Ranking		**Days since last update**		
		>= 30	**7 - 30**	**<=7**
I Liga	nº	0	0	17
	%	0.00%	0.00%	100.00%
II Liga	nº	2	2	11
	%	13.33%	13.33%	73.33%
Liga 3	nº	4	1	8
	%	30.77%	7.69%	61.54%
Campeonato de Portugal	nº	19	1	8
	%	67.86%	3.57%	28.57%
Distrital	nº	315	8	22
	%	91.30%	2.32%	6.38%

According to the table, in the *I Liga*, all websites have been updated within the past seven days, demonstrating high timeliness in news updates. Similarly, in *II Liga*, most websites (73.33%) have been updated within the past seven days, indicating a relatively up-to-date website.

Moving to *Liga 3*, a lower percentage of websites (61.54%) have been updated within the past seven days, indicating a slight delay in updates compared to the higher-ranked leagues.

In the *Campeonato de Portugal*, a significant portion of websites (67.86%) has not been updated in the past thirty days, suggesting a potential lack of regular updates for these clubs.

The *Distrital* exhibits the lowest frequency of updates, with only a small percentage of websites (6.38%) being updated within the past seven days, indicating a significant delay in updates for clubs in this ranking.

3.1.4 Conclusions

In conclusion, the analysis of the collected data highlights a significant digital divide between larger and smaller clubs in Portugal, particularly in terms of website presence, functionality, and update frequency. As we descend the rankings, there is an apparent decrease in the percentage of clubs with websites, with nearly half of *Liga 3* clubs lacking a website. This discrepancy becomes even more apparent when examining the functionality of existing websites and the update frequency, with a substantial proportion of lower-ranking clubs experiencing non-functional platforms or websites that have not been updated in the last thirty days.

The digital divide extends beyond websites' mere presence, encompassing the content's quality and currency. The scarcity of sections such as "Next Games", "Latest Results", "News", and "Players" suggest that lower-ranking clubs may struggle to provide comprehensive and up-to-date information to their online visitors. On the other hand, higher-ranking clubs demonstrate a higher commitment to maintaining updated websites, as evidenced by the consistently recent updates within the *I Liga*.

A digital divide in Portuguese football clubs has broader communication, engagement, and transparency implications. Smaller clubs with limited digital resources must establish an effective online presence, hindering their ability to engage with fans and potential investors. This divide may perpetuate existing power imbalances, with larger clubs enjoying greater visibility, access to resources, and growth opportunities, making it crucial to bridge this digital divide.

3.2 Problem Under Study

In this section, we aim to delve deeper into the problem we seek to address in this Dissertation: the existing digital divide between local and high-ranking clubs, particularly concerning their websites.

The digital divide between clubs with more significant financial resources and those with fewer resources, as evidenced by the analysis of the previously presented literature review and the results obtained by Mascarenhas, is conspicuous. Addressing this division is crucial, as effective communication with supporters, potential investors, and the general public is essential for clubs. A website is vital for establishing this connection, enabling clubs to share all necessary information in a centralised space accessible to interested parties.

However, the primary challenge lies in the financial constraints faced by smaller clubs, preventing them from hiring professionals to develop and maintain their websites, resulting in instances where websites remain outdated for extended periods exceeding a year. Additionally, even in cases where volunteers are available, the need for a user-friendly website development and maintenance solution further exacerbates the issue. The absence of a website due to financial limitations or a lack of volunteers with the requisite skills also deprives clubs of potential revenue streams that could arise from an effectively managed website.

3.3 Hypothesis and Research Questions

We believe it is possible to use publicly available data to generate updated and user-friendly websites for local sports clubs with minimal effort. The objective is to provide local sports clubs with an easy and cost-effective means of creating and maintaining a website that remains up-to-date with the latest news and games while ensuring good usability to facilitate user access to the desired information and customisation to reflect the clubs' brand identities.

By harnessing publicly available data sources, such as news articles, game schedules, and player statistics, the generated websites can be continuously updated in real-time, ensuring the

information presented remains current and accurate. This approach eliminates the need for manual updates, saving time and resources for local sports clubs.

Moreover, our belief emphasises the significance of user-friendliness in website design, which encompasses intuitive interfaces, streamlined navigation structures, and efficient information retrieval mechanisms. By prioritising usability, the generated websites aim to enhance the overall browsing experience for users, enabling easy access to desired content and fostering engagement.

Additionally, customisation is essential since it allows each local sports team to incorporate their unique branding elements, including colours and logos. This customisation aspect not only enhances the visual appeal of the websites but also reinforces the individual identity of each team, contributing to increased recognition and credibility.

Overall, our belief in the feasibility of automatically generating up-to-date and user-friendly websites for local sports teams is driven by its potential benefits, including improved communication, reduced maintenance efforts, enhanced user experience, and strengthened brand identity.

This work will be based on the following hypothesis:

> *"It is feasible to develop a platform that utilises publicly accessible data to automatically generate up-to-date and user-friendly websites for local sports teams, providing them with an efficient and cost-effective solution for creating and maintaining websites that remain current with the latest news and games while ensuring good usability and customisation for brand identity."*

Based on this hypothesis, we can guide our research in order to answer the following research questions:

RQ1. *Is it possible to automatically generate websites using publicly available data that are more* ***up-to-date*** *than the club's previous ones?*

We consider a website up-to-date when it accurately reflects current information, events, and data, with indicators such as timely content updates, accurate schedules and current contact information.

RQ2. *Is it possible to automatically generate websites using publicly available data that are more* ***user-friendly*** *than the club's previous one?*

We characterise a website as being user-friendly by its intuitive navigation, responsive design, fast loading times, visually appealing layout, and accessible features, all aimed at enhancing the user's experience and enabling easy access to information and actions.

RQ3. *Is it possible to automatically generate websites using publicly available data that require* ***minimal effort*** *for their creation?*

A website requires minimal effort to be generated when it utilises automatic processes, such as incorporating pre-designed templates and skins and leveraging existing readily available content.

RQ4. *Is it possible to automatically generate websites using publicly available data that allow* ***customisation****?*

A website allows customisation when it allows users to choose, personalise and modify various aspects of the website by choosing between different layouts and skins, for example.

3.4 Solution Proposal

The proposed solution incorporates a comprehensive architecture (Figure 3.2) that leverages the content available in *zerozero* [4] to automate the generation and customisation of club websites. *zerozero* is an extensive online platform renowned for its reliable and diverse football-related data, encompassing match schedules, player statistics, team profiles, league standings, and news articles. By harnessing the wealth of content offered by *zerozero*, our solution aims to revolutionise the creation and management of club websites, providing dynamic and informative platforms that engage supporters, sponsors, and stakeholders with accurate and up-to-date information about their respective teams.

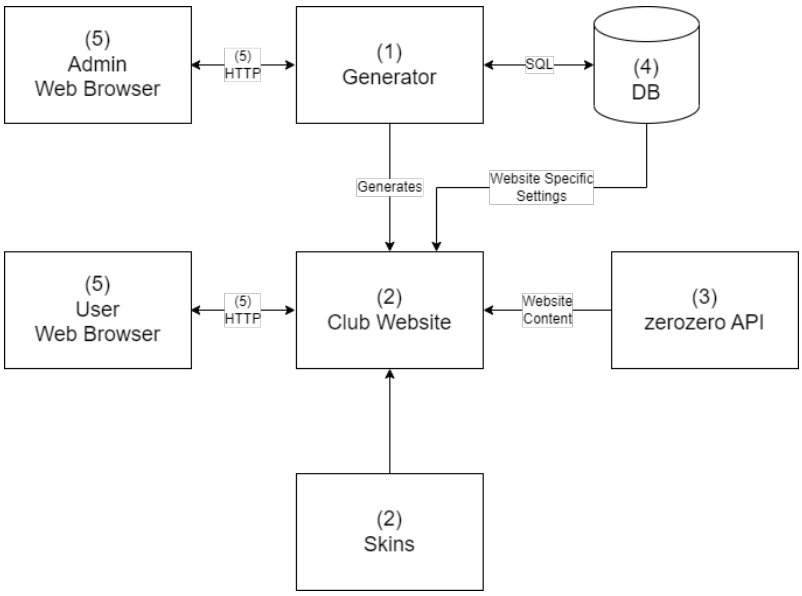


Figure 3.2: Diagram representing the proposed solution

The architectural blueprint of our solution encompasses the following interconnected components, each playing a crucial role in the system's operation:

1. **Generator**: This component takes centre stage in the architecture and generates the unique domain and associated PHP files for each Club Website. It interfaces with the Database to store pertinent website settings, such as the club ID and selected skin, ensuring seamless integration and customisation.

2. **Club Website**: Comprising a set of customisable PHP files, the Club Website component offers flexibility in design and layout by utilising templates and skins. It interacts with the

zerozero API, leveraging the club ID obtained from the Database, to retrieve comprehensive content, including match results, league classifications, news updates, and more, ensuring that the generated websites remain current, informative, and aligned with the latest developments in the football domain.

3. ***zerozero* API:** This integral component serves as the conduit for retrieving content from *zerozero*. Through the API, the Club Websites seamlessly interact with *zerozero*'s vast repository, enabling the retrieval of comprehensive and up-to-date information, which includes match schedules, player statistics, team profiles, league standings, and news articles, ensuring the generated websites are enriched with relevant and engaging content.

4. **Database (DB)**: The Database component stores specific website settings and domain information as a centralised repository, which includes the club ID, selected skin, and other relevant parameters. The Database facilitates efficient configuration and customisation of the Club Websites, providing administrators with a streamlined approach to tailor the websites to their respective club preferences.

5. **HTTP Connection**: The HTTP Connection component facilitates seamless communication between the Club Websites, the Generator, and the browsers. Leveraging the HyperText Transfer Protocol (HTTP), this component ensures smooth data transfer and an uninterrupted user experience.

Recognising the prevalent media coverage disparity experienced by smaller clubs with limited financial resources, our solution also integrates the Prosebot [2] tool offered by *zerozero*. Prosebot employs advanced natural language algorithms and extensive data searches to synthesise factual content by extracting information from game sheets, statistical data, historical records, milestones, and match-related news. By incorporating Prosebot into our solution, we empower the generated websites of smaller clubs to access up-to-date and relevant content, thereby enhancing their online presence and bridging the media representation gap across clubs with varying financial capabilities.

3.5 Methodology

To validate our hypothesis, we intend to conduct our validation process to answer the previously mentioned research questions affirmatively. This section will specify how we want to make this validation.

A comprehensive approach will be utilised to validate the first research question (**RQ1.**), involving manual verification of website timestamps and gathering feedback from club supporters. The manual verification process will compare the latest news on the generated websites with official club websites, while supporter feedback will assess the perceived currency of various website elements. The first research question is validated if both methods confirm that the generated websites are up-to-date and more current than the official ones.

To validate the second research question (**RQ2.**), a user-centric approach will be adopted, focusing on gathering feedback from supporters of low-ranking clubs regarding the usability of the developed website. Participants will assess the ease of performing tasks on the developed website compared to the official website, and the System Usability Scale (SUS) will be used to gather more specific insights. The second research question is confirmed if the feedback indicates that participants find the developed website user-friendly and receive positive ratings on the SUS.

To validate the third research question (**RQ3.**), a thorough manual assessment will examine the ease of generating club websites, irrespective of their ranking, league, or national affiliation. The generated websites should present comprehensive information, have automatic updates, and have no errors. These criteria indicate the successful generation of websites with minimal effort. By generating numerous websites effortlessly and meeting the criteria above, the research question can be confidently answered in the affirmative.

An investigation will validate the fourth research question (**RQ4.**) to determine the feasibility of creating multiple visually appealing skins for a single website, allowing for customisation. These skins must exhibit distinct visual characteristics while retaining essential elements that contribute to a club's identity, such as colours and logos. The successful creation of skins that maintain the club's identity while offering diverse visual variations for a single website provides affirmative evidence in addressing the research question.

If all of these research questions are effectively addressed with positive outcomes and necessary measures are taken to mitigate potential threats to validity, we can confidently assert the presented hypothesis's validity. Through meticulous examination and validation of each research question, we ensure that the hypothesis is substantiated by robust evidence and withstands scrutiny within the scientific framework.

3.6 Summary

This chapter serves as an introduction to the problem statement, examining the digital divide within the sports industry in Portugal (Section 3.1 (p. 16)), its impact (Section 3.2 (p. 21)), presenting research questions and a hypothesis (Section 3.3 (p. 21)), proposing a solution (Section 3.4 (p. 23)), and outlining the validation methodology to address the divide and empower lower-ranking teams (Section 3.5 (p. 24)).

The problem under study revolves around the significant digital divide observed in football clubs, particularly their online presence. We confirmed this divide by analysing the currency and functionality of the Portuguese football clubs' websites. This divide is characterised by higher-ranking clubs having more robust websites. In comparison, smaller and lower-ranking clubs need more resources and have outdated platforms that lack comprehensive and up-to-date information. Bridging this digital divide ensures equitable opportunities for all clubs, enhances stakeholder communication, and fosters transparency within the football community.

We believe it is possible to create and maintain updated and user-friendly websites for local sports clubs using publicly available data with minimal effort. The generated websites will

provide real-time updates and accurate information without requiring manual updates by utilising data sources such as news articles, game schedules, and player statistics. The emphasis on user-friendliness ensures intuitive interfaces, streamlined navigation, and efficient information retrieval for a seamless browsing experience. Additionally, customisation options allow each club to incorporate unique branding elements, reinforcing its identity.

The research questions formulated for this study address vital aspects such as the timeliness of information, usability, ease of generation, and customisation options. By investigating these research questions, we aim to validate our hypothesis.

The proposed solution integrates the extensive content available in *zerozero* to automate the creation and customisation of club websites. The solution's architecture comprises several interconnected components, including the Generator, Club Website, *zerozero* API, Database, and HTTP Connection, which collaborate to generate dynamic and informative websites.

A comprehensive validation approach will be implemented to support our hypothesis. It involves manual verification and gathering feedback from club supporters. The currency and superiority of the generated websites will be confirmed by comparing timestamps and obtaining supporter feedback. Usability will be assessed using the SUS, while a manual assessment will evaluate website generation ease and the feasibility of visually appealing skins. Successful validation of these factors will provide reliable evidence to support the hypothesis and establish its validity in the scientific context.

Chapter 4

Implementation

In response to the previously presented problem, we have developed a solution that will be explained in detail throughout this chapter. The proposed solution utilises the rich content available in the *zerozero* to automate the generation and customisation of club websites. By harnessing the wealth of content provided by *zerozero*, our solution aims to revolutionise creating and managing club websites, offering dynamic and informative platforms that effectively engage supporters, sponsors, and stakeholders with accurate and up-to-date information about their respective teams.

Firstly, we will address the implemented architecture in Section 4.1. Subsequently, we will explore the tools and technologies used to implement our solution (Section 4.2 (p. 29)). We will delve into the website generator tool used in Section 4.3 (p. 31). Finally, Section 4.4 (p. 33) will explain the gateway and front-end of the implemented solution for the club websites.

4.1 Architecture

To demonstrate the architecture of the implemented solution, we opted to use a deployment diagram, also known as an implementation diagram [5]. The deployment diagram is appropriate

because it shows the relationships between the system's software and hardware components and the processing's physical distribution [24]. Figure 4.1 illustrates the deployment diagram, providing a comprehensive overview of the system's architecture and configuration.

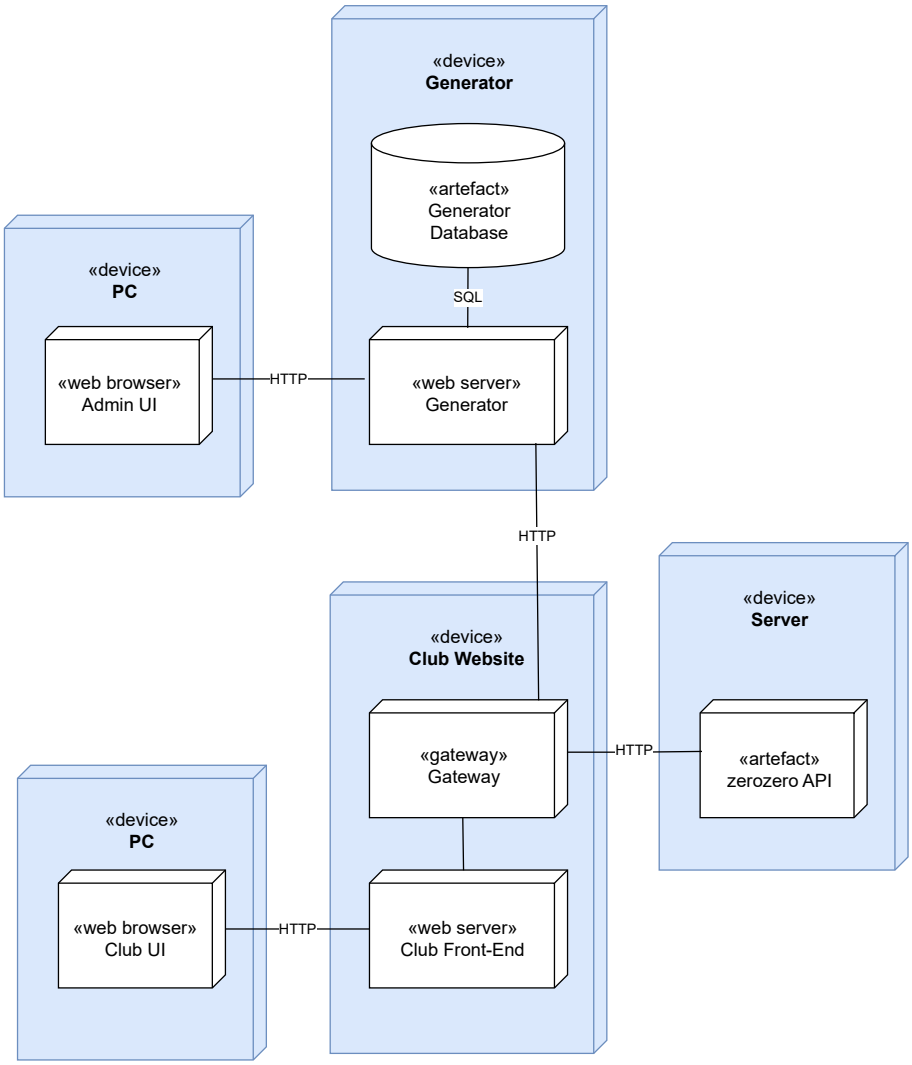


Figure 4.1: Deployment Diagram representing the architecture of the implemented solution

The deployed architecture of the implemented solution is illustrated in Figure 4.1, which consists of three devices:

- The device hosting the generator and its database;
- The device hosting the club website (including the club front-end and gateway components);
- The device hosting the *zerozero* API

The permissions and access levels vary for accessing the generator and the club website, leading to the representation of two different additional devices: one with the web browser of the administrator and another with the web browser of the regular user.

The club website device can be divided into two artefacts: the club front-end and the gateway. The club front-end web server primarily focuses on the client-side logic of the application, encompassing the design and development of the website's pages and components. It renders the user interface, handles user interactions, and presents the website's content.

Conversely, the gateway is crucial in providing data to the club front-end and facilitating communication with the generator's database and the *zerozero* API. It interacts with the *zerozero* API and the website generator through HTTP requests and responses, with responses consistently formatted as JSON. The gateway component retrieves relevant data from the *zerozero* API and facilitates the exchange of information between the generator's database and the club website.

The generator device consists of two artefacts: the generator itself and its database. The website generator web server is instrumental in the solution's architecture as it generates the club websites and performs cache clearance operations. When a website is generated, the generator's database is updated to store the domain of the newly created website and the selected settings. The generator web server communicates these settings to the club website's gateway through HTTP requests and responses, with responses transmitted in JSON format.

Furthermore, the generator's database is a persistent storage mechanism for the generator's data. It maintains records of the generated websites, including their respective domains and settings. The generator's database is pivotal in ensuring the proper functioning and configuration of the generated club websites.

Lastly, the device hosting the *zerozero* API represents a vital external component within the solution. It is utilised to acquire up-to-date information and data about the clubs. The club website device communicates with the *zerozero* API to retrieve relevant data, which is then processed and utilised in generating and presenting the club websites.

4.2 Tools and Technologies

This section will discuss the tools and technologies employed to implement the proposed solution and the reasons for their selection.

4.2.1 Gateway

In the gateway of the club websites, PHP was exclusively employed as the primary programming language for communication with both the *zerozero* API and the website generator API.

PHP was the intermediary for communication with the *zerozero* API, utilising HTTP requests to retrieve essential data for constructing the club websites. The JSON responses received from the API were processed and dynamically incorporated into the websites to ensure they were always up-to-date.

Additionally, PHP facilitated communication between the website generator and the club websites through the execution of HTTP requests. This interaction primarily occurred when cache clearance requests were made either by users through the generator interface or periodically to

maintain regular updates and synchronisation between the generator and club websites. Furthermore, PHP was also responsible for handling requests from the club websites to retrieve their specific settings, such as the chosen skin and domain, ensuring that each website could be customised according to the club's preferences.

The company chose Apache HTTP Server[1] for request routing, which was fundamental to the gateway infrastructure. It efficiently managed incoming HTTP requests and ensured web content delivery to clients. With its robust routing capabilities, Apache directed requests to the appropriate PHP scripts, enabling seamless execution of gateway processes and data retrieval.

Overall, PHP and Apache worked in tandem to facilitate seamless communication with external APIs, handle data retrieval and synchronisation, and efficiently manage request routing, contributing to the dynamic and user-friendly club websites.

4.2.2 Club Front-end

In the project's initial stages, the company preferred utilising a PHP template engine, specifically Latte[2], for the front-end development of the club's websites. The choice of Latte was by the company's familiarity with this technology. However, a comprehensive investigation into available template engines revealed that Latte might not be the optimal choice due to our lack of familiarity and the relatively small community support, potentially leading to implementation challenges that could not be easily overcome.

Furthermore, the requirement to use PHP for the club front-end narrowed down the available technology options, such as ReactJS[3], which were excluded. Two alternatives were presented in evaluating the best template engines: Twig and Blade, the templating engine used in Laravel. Blade emphasises simplicity, performance, and seamless PHP integration, making it suitable for PHP developers seeking a lightweight and familiar templating solution [40]. Another influential factor was our knowledge of Blade, which streamlined the implementation process. On the other hand, Twig offers enhanced security, separation of concerns, a strong template language, and a large community, appealing to developers aiming for a more comprehensive and secure templating engine [3].

Considering these factors, the company ultimately chose to implement the club front-end using Blade, specifically the `jenssegers/blade`[4] library, a standalone version of Laravel's Blade templating engine tailored for outside Laravel. This decision was made based on the company's requirement to exclude other functionalities provided by Laravel while maintaining the simplicity of Blade.

In addition to Blade as the chosen template engine, the club front-end development also incorporated Bootstrap 5.3.0 [5]. The utilisation of Bootstrap provided a robust and responsive framework for building consistent and visually appealing user interfaces across different devices and screen

[1] https://httpd.apache.org/
[2] https://latte.nette.org/
[3] https://react.dev/
[4] https://github.com/jenssegers/blade
[5] https://getbootstrap.com/

sizes. Additionally, we used CSS to complement these technologies to help design the web page and JavaScript for client-side scripting.

Overall, the careful evaluation of available options, considering factors such as familiarity, community support, performance, and security, led to the selection of Blade as the preferred template engine for club front-end development, supplemented by the utilisation of Bootstrap to enhance the visual and interactive aspects of the website, CSS for web page design and JavaScript for client-side scripting.

4.2.3 Database

The project's database, utilised exclusively by the website generation tool, is built using the MySQL[6] relational database management system (RDBMS). The decision to employ MySQL as the underlying technology for the database was a deliberate choice made by the company. The selection of MySQL was based on its reputation as a reliable and robust RDBMS known for its scalability, performance, and extensive community support [42]. The project team could harness MySQL's power effectively by leveraging phpMyAdmin[7], a widely recognised web-based database administration tool.

PhpMyAdmin was an effective and efficient solution for creating and managing the project's database. It is a user-friendly interface, and comprehensive features cover basic MySQL database and table operations, allowing for seamless database design, table creation, data insertion, and data manipulation tasks [14]. The tool's intuitive functionality enabled the project team to configure the necessary database schema, define relationships between entities, and ensure the integrity and consistency of the stored data.

4.3 Website Generator

To facilitate the implementation of our solution, the company has provided a tool developed in-house that functions as a website generator. The generator offers a simple and intuitive user interface, providing website viewing, creation, and cache clearance functionalities. Accessing the generator through the URL[8] gives users the access page, as depicted in Figure 4.2 (p. 32).

To access the homepage of the developed generator, users need to utilise a unique code assigned by the company. Upon entering the code, users are redirected to the initial page, which resembles Figure 4.3 (p. 33).

The homepage displays a comprehensive list of previously created club websites and their corresponding club ID in the *zerozero* database. Additionally, an option to clear the cache of these websites is available. To create a new website, users must click the "Create" button in the navigation bar, leading them to the website creation page depicted in Figure 4.4 (p. 34).

[6] https://www.mysql.com/
[7] https://www.phpmyadmin.net/
[8] https://clubes.pt

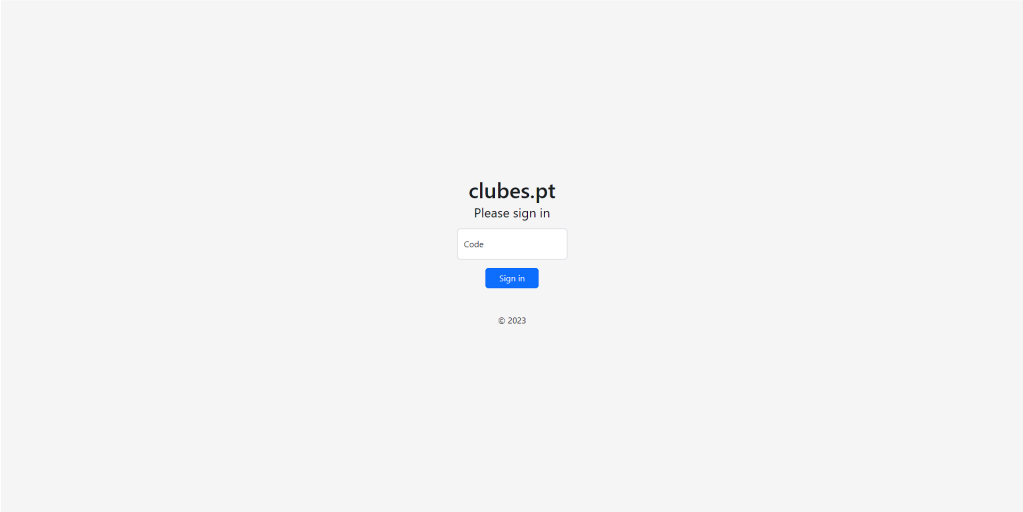


Figure 4.2: *clubes.pt*

Upon entering the website creation page for a specific sports team, users are initially prompted to select the club for which they intend to create a website. Currently, the generator allows any user to create a website for any club present in the *zerozero* database. Next, users are required to choose the desired domain for the website. It should be noted that the selected domain is always a subdomain `00.pt`. Once the domain is created, the website becomes immediately functional. Its functionality is achieved by directing and processing all requests to a subdomain `00.pt` in a unified location. Furthermore, users must select the preferred skin design for their website. At present, the generator offers a limited selection of two available skins.

Upon clicking the "Submit" button, a new entry is created in the generator's database, storing relevant information such as the club's ID, the chosen domain, and the selected skin. When accessing the generated website, `00.pt` uses API requests to retrieve club-specific information. The information on the website is based on the used domain, enabling the display of the respective club's details.

The website generator demonstrates simplicity and user-friendliness in its functionality, allowing users to view and create club websites efficiently. By leveraging subdomains and integrating with the *zerozero* database, the generator ensures seamless operation and personalised content delivery for each club's website.

It is worth noting that the generator provided by the company for this project functions solely as an initial prototype rather than a final product intended for distribution to clubs. This prototype serves as a proof-of-concept for demonstrating the feasibility and basic functionalities of the website generation process. As such, it should be regarded as an early-stage implementation that requires further refinement and development before being considered a market-ready solution for clubs.

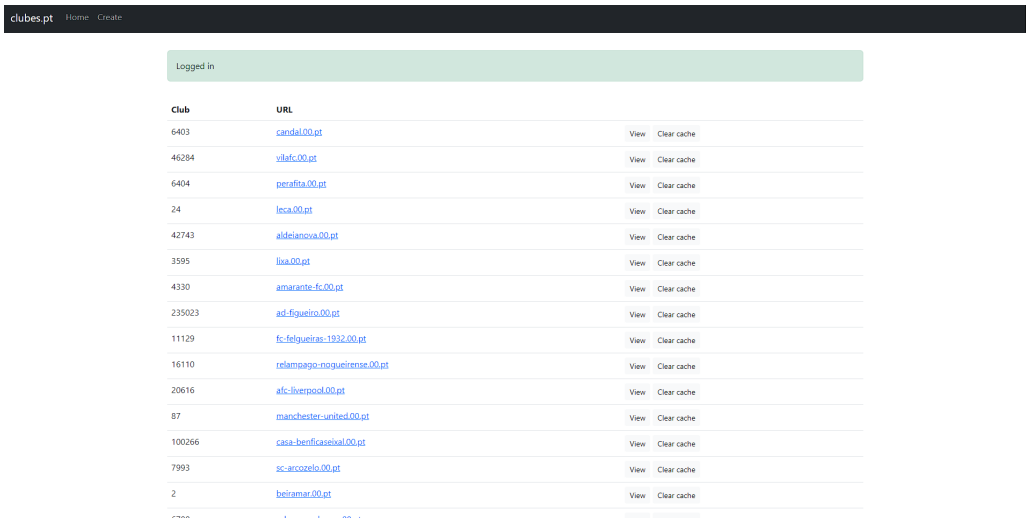


Figure 4.3: *clubes.pt* Homepage

4.4 Club Website

In this section, we aim to elucidate the functioning of both the gateway and club front-end of the solution we implemented. Concerning the gateway, we will discuss the endpoints of *zerozero* that were employed and the endpoints of the implemented websites. As for the club front-end, we will address the diverse pages implemented for each skin and their corresponding components.

4.4.1 Gateway

In this subsection, we will delve into the intricacies of the gateway functionality of the solution implemented by our team. We will explore the fundamental aspects, using endpoints provided by *zerozero*, a pivotal platform in our implementation. We will also investigate the implemented endpoints within the club websites, facilitating seamless user navigation and interaction. These endpoints enabled efficient data retrieval and website content display, enhancing the user experience.

For the gateway development, as mentioned in Section 4.2 (p. 29), we employed PHP as the primary programming language for facilitating communication between the *zerozero* endpoints and the generator. PHP served as a versatile tool in handling data exchange and processing tasks. Additionally, Apache HTTP Server was the routing server, directing incoming requests to the appropriate endpoints and ensuring smooth navigation within the system.

4.4.1.1 *zerozero* Endpoints

The endpoints provided by *zerozero* are essential for building a comprehensive and up-to-date website for a sports club. These endpoints cover various functionalities, allowing us to retrieve and display crucial information for various aspects of the club on a website automatically generated.

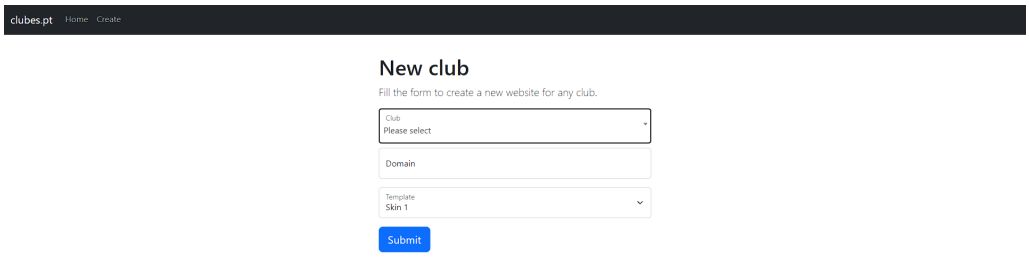


Figure 4.4: *clubes.pt* Website Creation Page

The endpoints include retrieving basic information about the club, such as its profile details. Additionally, team-specific information such as news articles, games, classification, and squad details can be accessed. Competition-related data can also be retrieved, including game details and classification standings. Sports information endpoints enable developers to access teams and competitions associated with a particular sport.

Furthermore, the game-specific endpoint allows developers to fetch detailed information about a specific match, including match reports and events during the game. The calendar endpoint provides upcoming game details, enabling the creation of a user-friendly calendar page to keep fans and users updated on the club's schedule.

Lastly, the news article endpoint allows developers to retrieve comprehensive information about specific news articles, including publication date, author, title, summary, full text, associated sports details, and image path.

These endpoints have been developed by the software developers' team at *zerozero*, providing convenient access to the necessary data for a sports club website implementation. By leveraging these endpoints, developers can create a dynamic and engaging online presence for the club, catering to the needs of fans and users.

Club Basic Information An API request can be made to retrieve essential basic information about a club. By giving the club ID as a parameter in the URL, the endpoint efficiently fetches the necessary data from the *zerozero* API. The retrieved information divides into two distinct sections:

Profile Provides crucial details for constructing the club's profile. It includes the club's unique identification number (ID), abbreviated short name, official full name, logo path or URL, physical address, and social media links such as Facebook, Twitter, YouTube, and Instagram.

Club Teams Encompasses information about all teams associated with the club. It includes details about the sport where each team participates, their team IDs, and their official names or designations. This comprehensive endpoint facilitates the seamless retrieval of necessary team-related details, enabling a comprehensive understanding of the club's teams.

Homepage The "Homepage" endpoint provided by *zerozero* is designed to retrieve the necessary information for building the homepage of a specific club based on its unique ID. The endpoint efficiently retrieves the essential data by replacing the ID in the request's URL with the club's unique ID. The retrieved information consists of news articles and game details, enabling the construction of a dynamic and engaging homepage for the club's website. The "Homepage" endpoint serves the following purposes:

News Retrieves the ten most recent news articles belonging to the specified club ID from *zerozero*. The information includes the news article's ID, publication date, title, author, news summary, full text, the path to the news image, the associated sport's ID, and the sport's name. This feature allows for the display of up-to-date news articles on the club's homepage.

Games Retrieves the ten most recent news articles from zerozero that belong to the specified club ID. The information includes the news article's ID, publication date, title, author, news summary, full text, the path to the news image, the associated sport's ID, and the sport's name. This feature allows for the display of up-to-date news articles on the club's homepage.

Team The *zerozero* API provides a dedicated endpoint for retrieving essential information to build a page for a specific team based on its unique ID. By requesting the designated URL, the endpoint efficiently fetches the required data. This endpoint serves multiple purposes, enabling the extraction of various details related to the team. The available functionalities include the following:

Name Retrieves the team's name, providing a straightforward way to identify and display the team's name.

News Retrieves the ten most recent news articles about the specified team ID from *zerozero*. The information includes the news article's ID, publication date, title, author, news summary, full text, the path to the news image, the associated sport's ID, and the sport's name. This feature facilitates the display of team-specific news articles, ensuring the latest updates are easily accessible.

Games Retrieves the last ten completed games and the next ten upcoming team games. The game information includes the game's ID, date, score, names and logos of the home and away teams, competition details, category of the game, and stadium information. This functionality allows for the display of recent game results and upcoming matches specific to the team.

Classification Provides the team's current standing and classification in all the competitions in which they are participating. The information includes the name of the competition and a grid that represents each line's position, team name, points, games played, victories, ties, losses, goals scored, goals conceded, and the colour of the row indicating the team's potential outcome at the end of the competition. This feature offers valuable insights into the team's performance in various competitions.

Squad Retrieves all the players from the team's squad. The information includes the player's name, age, number, position, photo path, and the path to the player's flag. This functionality allows for the comprehensive display of the team's squad, providing details about each player.

Competition To build the page dedicated to a specific competition, *zerozero* API offers a dedicated endpoint. This endpoint requires the edition ID (`idEdition`) and the participating team's ID (`idTeam`) to fetch the necessary data efficiently. The endpoint serves multiple purposes, allowing the extraction of various details related to the competition. The functionalities provided are as follows:

Name Retrieves the name of the competition, enabling the identification and display of the competition's name.

Games Retrieves the team's last played games and upcoming games within the competition. The game information includes the game's ID, date, score, names and logos of the home and away teams, competition details, category of the game, and stadium information. This functionality facilitates the display of recent and upcoming games specific to the team's participation in the competition.

Classification Provides the team's current standing and classification within the competition. The classification may be presented as a grid, showcasing the team's position, team name, points, games played, victories, ties, losses, goals scored, goals suffered, and the colour of the row indicating the team's potential outcome at the end of the competition if they maintain that position. This feature offers valuable insights into the team's performance and position within the competition.

Sport The *zerozero* API provides a dedicated endpoint to retrieve essential information for building a page dedicated to a specific sport. This endpoint requires the sport ID (`idSport`) and the participating team's ID (`idTeam`) to request the designated URL and efficiently fetch the necessary data. The endpoint serves two primary purposes, allowing the extraction of various details related to the sport. The functionalities provided are as follows:

Teams Retrieves all the teams of the club that practice the specific sport. The information includes the team ID and the name of each team. This functionality enables the display of all the teams associated with the club participating in the sport, providing a comprehensive overview of the club's involvement.

Competitions Retrieves the competitions in which at least one team from that sport is enrolled. The information includes the competition ID, edition ID, description of the competition, and the competition image URL. This feature allows for identifying and displaying the competitions associated with the sport, providing information about the competitions in which the club's teams participate.

Game The "Game" endpoint provided by the *zerozero* API retrieves essential information for building a page dedicated to a specific game. The necessary data can be efficiently fetched using the game ID in the designated URL for the API request. The endpoint serves three primary purposes, allowing the extraction of various details related to the game. The functionalities provided are as follows:

Info Retrieves information about the game, including the game ID, sport ID, match name, date, names and logos of the home and away teams, stadium information, competition details, and the home and away goals scored during the game. This functionality provides a comprehensive overview of the game, including the teams involved, the venue, and the goals scored, allowing for a detailed game page.

Report Retrieves the game report, including the players from the home and away teams and the coaches. Each player's information includes their number, name, country flag, and events with their respective minutes, such as cards, goals, assists, and more. This feature enables displaying detailed player and coach information, including their contributions during the game.

Events Retrieves the events needed to create the detailed match report. The information includes the minutes at which the events occurred, the corresponding text descriptions, and event image URLs. This functionality provides a comprehensive list of events during the game, such as goals, cards, and assists, allowing for creating a detailed and engaging match report.

Calendar The *zerozero* API provides a dedicated endpoint that retrieves the necessary information to build a news article page. We should use the designated URL with the ID of the news article we want to consult to retrieve the necessary data. This endpoint primarily serves the purpose of retrieving detailed information about a news article. The retrieved information includes the news ID, publication date, title, author, a small text summarising the post, the full text of the news, the path to the news image, the associated sport's ID, and the sport's name.

This endpoint allows developers to gather all the necessary data to construct a comprehensive news article page. The information retrieved from the endpoint provides details such as the publication date, author, title, and news article summary. The full text of the news, along with the associated sport's ID and name, is included. The path to the news image is also provided, allowing for the inclusion of relevant visual content.

News Article The *zerozero* API provides a dedicated endpoint that retrieves the necessary information to build a news article page. We should use the designated URL with the ID of the news article we want to consult to retrieve the necessary data. This endpoint primarily serves the purpose of retrieving detailed information about a news article. The retrieved information includes the news ID, publication date, title, author, a small text summarising the post, the full text of the news, the path to the news image, the associated sport's ID, and the sport's name.

This endpoint allows developers to gather all the necessary data to construct a comprehensive news article page. The information retrieved from the endpoint provides details such as the publication date, author, title, and news article summary. The full text of the news, along with the associated sport's ID and name, is included. The path to the news image is also provided, allowing for the inclusion of relevant visual content.

The endpoints are summarised in Table 4.1 (p. 39), providing a concise overview of the described endpoints.

4.4.1.2 Club Website Endpoints

The implemented endpoints, designed for the club website, were instrumental in enhancing its functionality and user experience. These endpoints, built using Apache as the web server, served as crucial access points to specific features and content within the website. They facilitated seamless interaction between the client-side and server-side components, allowing efficient data retrieval and manipulation.

The Apache-based web server utilised a meticulously crafted `.htaccess` file to enable the desired functionalities. This file, residing in the website's root directory, was pivotal in configuring the server's behaviour. It provided per-directory configuration settings and facilitated URL rewriting and routing [39], ensuring smooth navigation and content delivery within the club website.

The `.htaccess` file contained directives that defined the behaviour of the Apache server. It included rules for URL rewriting, error handling, and redirection. These rules, defined using the

Table 4.1: Endpoints used from the *zerozero* API

Endpoint	Method	Parameters	Description
/getClubBasicInfo	GET	ClubID	Gets the basic information for a specific club ID (profile and teams)
/getClubHomepage	GET	ClubID	Gets the information needed to build the homepage of the website for a specific club ID (news and games)
/getClubTeamInfo	GET	TeamID	Gets the information needed to build the page for a specific team ID (name, news, games, classification and squad)
/getClubCompetitionInfo	GET	EditionID, ClubID	Gets the information needed to build the page for a specific competition ID based on the participating team ID (name, games and classification)
/getClubTeamsbyMod	GET	ClubID, ModID	Gets the information needed to build the page for a specific sport ID based on the participating team ID (teams and competitions)
/getClubMatchInfo	GET	MatchID	Gets the information needed to build the page for a specific game ID (information, match report and events)
/getClubSchedule	GET	ClubID	Gets the information needed to build the calendar page for a specific game ID (games)
/getClubNews	GET	NewsID	Gets the information needed to build the news article page for a specific news ID (news article)

`RewriteRule` syntax, allow for mapping incoming URLs to specific files or scripts on the server. By intelligently rewriting the URLs, the endpoints within the club website could be accessed seamlessly, delivering the desired content to the users.

Furthermore, the `.htaccess` file incorporated the ErrorDocument directive. This directive was crucial in handling the 404 (Not Found) errors. By specifying a custom error page through the ErrorDocument directive, the club website could provide users with a more user-friendly and informative experience when encountering missing or inaccessible pages.

The implemented .htaccess file can be observed in Appendix A (p. 87), showcasing the configuration directives and rules employed for the routing and manipulating URLs within the web server.

A summary of the implemented club website endpoints, configured through the `.htaccess` routing file is presented in Table 4.2 (p. 40). This overview offers a comprehensive view of the routing configuration and the associated functionalities provided by the club website endpoints.

Table 4.2: Club Website Endpoints

Endpoint	Method	Parameters	Description	Response
/	GET	-	Retrieves Homepage	200 OK
/sport/:id	GET	id	Retrieves the Sport page based on its ID	301 Moved Permanently
/competition/:sportID/:competitionID	GET	sportID, competitionID	Retrieves the Competition page based the given sport ID and competition ID	301 Moved Permanently
/team/:sportID/:teamID	GET	sportID, teamID	Retrieves the Team page given their sport ID and team ID	301 Moved Permanently
/game/:gameID	GET	gameID	Retrieves the Game page given its ID	301 Moved Permanently
/news/:newsID	GET	newsID	Retrieves the News Article page given its ID	301 Moved Permanently
/calendar	GET	-	Retrieves the Calendar page	301 Moved Permanently
/contacts	GET	-	Retrieves the Contacts page	301 Moved Permanently

4.4.2 Club Front-end

This subsection will explore the various pages and components implemented for the sports club website. We will provide an overview of the functionalities and features of each page, along with a description of the user interface components used. By understanding the purpose and composition of these pages and components, we can gain insights into the structure and user experience of the website. This section offers a comprehensive understanding of the implemented website pages and their respective components, showcasing how they contribute to creating an engaging and informative online presence for the sports club.

The design of the website was given to us by the company's design team, and our role was to implement precisely what was provided to us. We followed their design specifications and incorporated all the visual elements, layout, and styling per their instructions. Our focus was on translating their design into a functional website while ensuring the accuracy and fidelity of the implemented solution.

For this analysis, we will be consulting the website created for the club *Vila Futebol Clube*[9], a Portuguese low-ranking football club.

[9] https://vilafc.00.pt/

4.4.2.1 Pages

Each page is designed to serve a specific purpose and provide valuable information to website visitors. We will discuss the following pages: homepage, sports, team, news, competition, game, calendar, and contacts. By examining the components and features of each page, we aim to provide a comprehensive understanding of the content and functionality available to users on the sports club website.

Both skins' implemented pages are available for reference in Appendix B (p. 89).

Homepage — Figure B.1 (p. 90) and Figure B.2 (p. 91)

The homepage serves as the central hub of the sports club's website. It features a dynamic news section showcasing the most recent news articles about the club. The homepage also displays information about upcoming games and the results of previous games. Its content results from the request to the "Homepage" endpoint; see Subsection 4.4.1.1 (p. 35).

Sport Page — Figure B.3 (p. 92) and Figure B.4 (p. 93)

The sport page provides comprehensive information about a sport in which the club participates. It showcases the teams associated with the club that practice the sport. Additionally, the sports page provides insights into the competitions in which the club's teams are enrolled, including competition image, if available. This page is valuable for fans and enthusiasts seeking information about the club's involvement in a particular sport. The page's content is generated from the data retrieved through the request made to the "Sport" endpoint, see Subsection 4.4.1.1 (p. 37).

Team Page — Figure B.5 (p. 94) and Figure B.6 (p. 95)

The team page provides in-depth details about a specific team affiliated with the club. It offers a range of information to engage fans and followers, including news articles related to the team. Visitors can also find information about the team's next upcoming games, enabling them to plan their support accordingly. The team page further provides a glimpse into their performance by displaying previous game results and current classification within their respective competition. Additionally, visitors can explore the team's squad, discovering details about individual players, such as their positions and jersey numbers. The page's content is generated based on the data from the request to the "Team" endpoint, see Subsection 4.4.1.1 (p. 35).

News Article Page — Figure B.7 (p. 96) and Figure B.8 (p. 97)

The news page presents the details of a specific news article, comprising the title, text, image, and social media sharing options. Additionally, the page includes information about the publication date of the news article. Users can interact with the news content and conveniently share it across social media platforms. The page's content is dynamically generated using the data retrieved from the "News Article" endpoint, see Subsection 4.4.1.1 (p. 38).

Competition Page — Figure B.9 (p. 98) and Figure B.10 (p. 99)

The competition page provides valuable insights into the club's participation in specific competitions. It offers a glimpse into the club's performance by displaying the results of previous games in which the club has participated. Additionally, visitors can access information about upcoming games, allowing them to anticipate and follow the team's progress in the competition. The competition page also includes a classification section, providing a snapshot of the club's current standing within the competition. The content on this page is dynamically generated using the data obtained from the "Competition" endpoint, see Subsection 4.4.1.1 (p. 36), request, ensuring that it remains up-to-date and accurate.

Game Page — Figure B.11 (p. 100) and Figure B.12 (p. 101)

The game page is dedicated to providing detailed information about a specific game. It offers a comprehensive game overview, including the participating teams, date, and stadium. Visitors can explore the players involved in the game, discovering information about their performances and contributions. The game page may also include a report section, if available, offering a more in-depth analysis of the game going through the important game events. The page's content is dynamically generated using the data obtained from the "Game" endpoint, see Subsection 4.4.1.1 (p. 37), request, ensuring that it remains up-to-date and accurate.

Calendar Page — Figure B.13 (p. 102) and Figure B.14 (p. 103)

The calendar page is a convenient reference for the club's upcoming games. It displays the team's schedule, showcasing the dates, opponents, and competition details. The calendar lets fans and followers plan attendance and stay updated with the team's matches. The content displayed on this page is generated in real-time using the data obtained from the "Calendar" endpoint, see Subsection 4.4.1.1 (p. 38), request, ensuring that the information presented is always up-to-date and reliable.

Contacts Page — Figure B.15 (p. 104) and Figure B.16 (p. 105)

The contacts page is a hub for important contact information related to the sports club. It offers visitors access to essential details such as the club's address, phone number, email, and social media profiles. The contacts page may also feature a user-friendly contact form, facilitating direct communication between visitors and the club's administration. This page fosters engagement and interaction between the club and its supporters. The content presented on this page is sourced from the "Club Basic Information" endpoint, see Subsection 4.4.1.1 (p. 34), request, ensuring accuracy and reliability.

4.4.2.2 Components

The components implemented play a vital role in enhancing the functionality and user experience of the website. We will discuss the components used across different pages, including the homepage, sports page, team page, news article page, competition page, game page, calendar page, and contacts page. By examining these components and their integration within each page, we aim to provide a comprehensive understanding of the interactive and visual elements that contribute to the overall design and usability of the website.

Navigation Bar The implemented component consists of crucial elements that provide easy navigation and access to essential information on the sports club website. It includes logos that serve as hyperlinks to the club's social media platforms, such as Facebook, Instagram, and Twitter, facilitating seamless interaction with fans. The component also features the club's phone number for direct contact and a hyperlink to Google Maps, allowing visitors to locate the team's stadium quickly.

In addition, this component features the club's logo, providing a convenient hyperlink to the homepage and ensuring quick access to the main website. The prominent display of the club's name further strengthens its identity. The component's hyperlinks are organised into distinct sections, beginning with the homepage and followed by dedicated sections for each sport in which the team is involved. It also includes hyperlinks for the calendar and club contacts, facilitating easy navigation to these crucial sections of the website.

To enhance user experience, an expanded navbar is implemented, accessible through a dedicated button. The expanded navbar appears across all implemented pages and allows users to navigate seamlessly between different teams within the club, conveniently categorised. This dynamic navbar is constructed based on the data retrieved from the "Club Basic Information" endpoint request, ensuring its accuracy and relevance throughout the website.

Regarding the displayed information, there are no differences between the navigation bar in both skins; the variations lie solely in the aesthetic design.

Homepage Hero The homepage hero component is designed to captivate visitors and provide key highlights of the sports club. Skin 1 features a news carousel showcasing news articles with accompanying photos, the respective sport, the title, and a link to the news page for more details. The homepage hero for Skin 1 also includes a games carousel that displays information regarding previous and future games.

In Skin 2, the hero component presents a news carousel with news articles displayed alongside their corresponding photos. It provides information about the sport, title, date, and a brief description of each news article. This design variation offers a sleek and concise presentation of the latest news content on the homepage.

Both skins of the hero component serve as prominent sections exclusively on the homepage, effectively engaging visitors and keeping them updated on the club's news and game highlights.

Games Carousel The games carousel component provides an interactive display of the club's past and upcoming games, offering essential details for each match. In Skin 1, the games carousel is incorporated within the homepage hero, while in Skin 2, it functions independently.

Each box within the games carousel showcases information about a specific game, including the home and away team names and logos. The game score is also prominently displayed, allowing users to grasp the outcome quickly. Additionally, the sport and competition associated with the game are provided, giving context to the match. The game works as a hyperlink to its detailed page.

Further details, such as the stadium where the game took place and the scheduled time of the match, are included, enhancing the comprehensive information provided to the users. The games carousel focuses on the most recent past game, allowing users to review the club's performance and preview the next ten upcoming games.

This component is only available on the homepage for both skins.

News Section The news section component displays the sports club's latest news articles, delivering timely updates and valuable insights to visitors. Each news block features essential elements such as the news title, associated sport, and a summary. Users can click the provided link to access the detailed news page for comprehensive information. The publication time is also indicated, helping users determine the news article's recency.

Skin 1 may present several news blocks, while Skin 2 may have a different display. The aesthetics, including the visual layout of each news block, may also vary between the two skins. Nonetheless, both component variations serve the same purpose of showcasing the latest news and offering easy access to further details.

The news section component can be found on both the homepage and the team page, ensuring visitors stay informed about the sports club's latest news and developments.

Next Games Table The next games section component provides information about upcoming games for the sports club. The homepage includes a sport filter allowing users to narrow down the displayed games based on their preferred sport. The component is presented in the form of a table, with each row containing details such as the time of the game, the names and logos of the home and away teams, the team level, the competition, and the stadium.

In Skin 1 and Skin 2, the next games section component serves the same purpose and displays the same information. The only difference lies in the aesthetics, where the visual presentation may vary between the two skins.

The next games section component on the team, calendar and competition pages follows a similar structure to the homepage but without the sport filter functionality. It still presents the upcoming games in a table format, providing the relevant details of each game.

Footer The footer component appears on all website pages, which looks the same on both skins. It includes the club name and links to the terms and conditions and privacy policy, ensuring that visitors can access important information regardless of the page they are on.

Teams and Competitions Cards The teams and competitions cards component showcases a collection of cards, each representing a team or competition associated with the sports club. Each card displays the logo of the respective team or competition, if available. If a logo is unavailable, the card will show the club's logo instead. Additionally, the name of the team or competition is prominently featured on each card.

While there may be slight aesthetic differences between Skin 1 and Skin 2, the overall design and layout of the cards remain consistent. Both component variations serve the same purpose of presenting teams and competitions in a visually appealing manner, allowing visitors to quickly identify and navigate through the different entities associated with the sports club.

Results Table The results component is prominently featured on both team and competition pages, providing visitors with comprehensive information about game outcomes. Presented in a table format, each row of the table represents a specific game and includes vital details such as the date of the game, the names and logos of the home and away teams, the score (displayed in green for a win, red for a loss, and grey for a tie), the team level, the competition, and the stadium where the game took place.

The results component is a valuable resource for fans and followers, allowing them to track the team's performance or the competition's progress. While the aesthetics may differ slightly between skin 1 and 2, both component variations share the same functionality and purpose of presenting game results in an organised and visually appealing manner.

Classification The classification component provides an overview of the team or competition standings, offering valuable insights into their performance. Displayed in a table format, each row represents a team or participant, including their position, games played, victories, ties, defeats, goals scored, goals awarded, and total points. While there may be aesthetic differences between skin 1 and 2, both component variations fulfil the exact purpose of presenting the classification information clearly and effectively.

The classification component appears on both team and competition pages, ensuring that visitors can access the latest standings and stay updated on the progress of their favourite team or competition. Although currently limited to a table format, future work may expand the functionality of this component.

Squad The squad component provides an organised display of the team's roster, categorised by position. Each position within the component has its dedicated section showcasing individual player cards. These player cards offer essential information such as a player's photo, name, jersey number, age, and nationality. The squad component is a comprehensive

overview of the team's lineup, allowing fans and followers to familiarise themselves with the players and their respective positions. While the aesthetic presentation may vary between skin 1 and 2, both component variations maintain the same functionality and purpose of displaying the team's roster.

The squad component appears exclusively on the team page, offering visitors a dedicated space to explore the players and gain insights into the team's composition.

Game Header The game header component provides essential information about the ongoing or upcoming game being watched. It serves as a visual representation of the teams involved in the match, offering details that help viewers stay informed. The component prominently displays the logos and names of the home and away teams, instantly recognising the competing sides. In addition, it showcases the current score or "VS" if the score is unavailable.

The game header component also includes essential details such as the scheduled time of the game, the competition in which it takes place, and the stadium where the match is being held. These details offer viewers a comprehensive overview of the game and help them stay updated with the critical information surrounding the match.

While the aesthetic presentation may differ between skin 1 and skin 2, the core functionality and purpose of the game header component remain consistent. It appears exclusively on the game page, ensuring that viewers have access to the vital game-related information they need when following a specific match.

Team Information in Game The Team Information in Game component offers key details about each team participating in the game. It includes sections for the starting players, substitute players, and the coach. Each team member's jersey number, nationality flag, name, and game events, such as goals, cards, entries, and more, are displayed, along with the corresponding minutes. This component provides viewers with a comprehensive overview of the team's composition and player performance during the game.

Displayed exclusively on the game page, the Team Information in Game component enhances the viewing experience by providing crucial insights into the teams and their players. It lets viewers follow the lineup decisions, track player contributions, and better understand the game's dynamics. Although the component's appearance may differ slightly between skin 1 and 2, the underlying purpose and functionality remain consistent, enriching the game page with valuable team-specific information.

Game Report The Game Report component offers a comprehensive overview of the game's events, presenting viewers with a concise and structured summary. Each event is meticulously recorded with its corresponding time, represented by an icon and accompanied by a brief textual description. From goals and penalties to substitutions and fouls, the Game Report captures the significant moments that unfolded during the game, providing a valuable resource for fans and spectators.

The Game Report component enhances the viewing experience by offering real-time updates and insights exclusively on the game page. Its visually appealing layout and organised presentation allow viewers to easily navigate and comprehend the sequence of events, ensuring they stay informed and engaged throughout the game. Whether following the action live or catching up afterwards, the Game Report provides a reliable and convenient reference for reliving the game's key moments.

Contact Information The Contact Information component on the Contacts page provides the necessary details for connecting with the sports club. It includes the club's phone number, schedule, and stadium address, ensuring easy access to essential information for fans and visitors. With consistent information across both skins, the component offers a reliable resource for engaging with the club, reaffirming its commitment to open communication.

Appearing exclusively on the Contacts page, the Contact Information component is a dedicated point of reference for individuals seeking to establish contact or gather relevant details.

Email Form The Email Form component on the Contacts page allows visitors to send emails to the club's staff. Despite its non-functionality, the form demonstrates the club's commitment to direct communication and engagement. With only aesthetic differences between the two skins, the component maintains a consistent appearance and emphasises the club's willingness to receive messages. It serves as a reminder of the club's dedication to fostering open lines of communication and providing a streamlined means for individuals to reach out.

4.5 Summary

This chapter delves into the implementation of our solution, covering the architectural design (Section 4.1 (p. 27)), the tools and technologies utilised (Section 4.2 (p. 29)), the website generator (Section 4.3 (p. 31)), and the gateway and club front-end components of the club websites (Section 4.4 (p. 33)).

The deployed architecture of the solution involved three devices: the generator and its associated database, the club website device (consisting of club front-end and gateway components), and the device hosting the *zerozero* API. Access levels and permissions differed for administrators and regular users, leading to separate devices for their web browsers. The club front-end component focused on client-side logic, encompassing user interface rendering, user interactions, and content presentation. The gateway component communicated with the generator's database and the *zerozero* API, retrieving data and facilitating information exchange. The generator device generated club websites, updated its database with domain and settings information and served as persistent storage for generated website records. The *zerozero* API provided up-to-date club information. This architectural design facilitated effective communication, data processing, and website generation within the solution.

The gateway relied on PHP as the primary programming language, facilitating communication with the *zerozero* API and the website generator API through HTTP requests and JSON responses.

Apache HTTP Server was selected for request routing, ensuring efficient request management and content delivery. The club front-end development of the club websites was implemented using the Blade PHP template engine, chosen for its familiarity, performance, and seamless integration with PHP. Bootstrap 5.3.0 was utilised to enhance the visual and interactive aspects of the websites, complemented by CSS for web page design and JavaScript for client-side scripting. The MySQL database was chosen for its scalability, performance, and community support, managed via the user-friendly phpMyAdmin interface for streamlined database design and manipulation. These tools and technologies were carefully evaluated and selected to develop dynamic and user-friendly club websites.

The company provided an in-house website generator tool to facilitate the implementation of the solution. This generator offered a user-friendly interface for website creation, viewing, and cache clearance. Accessing the generator required a unique code assigned by the company, redirecting users to the homepage where previously created club websites and their respective club IDs were displayed. Users could clear the cache of existing websites or create new ones by selecting a club, subdomain, and skin design. The generator's database stored relevant information such as club ID, domain, and skin, while the generated websites retrieved club-specific information through API requests. The generator demonstrated simplicity and efficiency, providing personalised content delivery for each club's website through subdomains and integration with the *zerozero* database. However, it should be noted that the generator was in its prototype stage and required further refinement and development before being considered a market-ready solution for clubs.

Our team developed a solution that leverages the *zerozero* platform's endpoints to retrieve and present essential information about different aspects of a sports club on a dynamically generated website. The gateway implementation, employing PHP as the primary programming language and Apache HTTP Server as the routing server, facilitated seamless communication with the *zerozero* endpoints. This integration enabled efficient data retrieval encompassing club details, team-specific information, competition-related data, game specifics, calendar events, and news articles. We harnessed these endpoints and achieved a comprehensive and up-to-date online presence, enhancing user engagement and experience.

On the club front-end, our club websites featured a user-friendly and dynamically interactive interface, augmenting visitor engagement. Various intelligently designed pages and components fulfilled specific purposes in delivering valuable information. The homepage was a central hub, showcasing timely news updates and game-related content. Moreover, distinct pages such as sport, team, competition, and game pages provided insights into associated teams, competitions, and specific game details. Complemented by well-crafted components like the navigation bar, hero section, games carousel, news section, and footer, our front-end implementation significantly elevated the overall user experience. Visitors could conveniently access relevant information, stay abreast of news and game events, and actively interact with the platform.

Chapter 5

Empirical Evaluation

This empirical evaluation chapter assesses two critical dimensions related to the developed websites: the perception of up-to-dateness surpassing official websites and the evaluation of appearance and usability, which was achieved by designing two distinct questionnaires targeting different groups of football club supporters. The first questionnaire aims to gather feedback from supporters of low-ranking football clubs, including *GD Aldeia Nova*, *Leça FC*, *Perafita*, and *Vila FC*, to determine whether the developed websites were perceived as more up-to-date than their official counterparts and adequately represented the respective clubs. The second questionnaire targeted supporters of first-league clubs, namely *SL Benfica*, *FC Porto*, *Sporting CP*, and *SC Braga*, to compare the usability differences between the official and developed websites. This questionnaire also sought to evaluate the performance of these websites in accomplishing specific tasks and obtain feedback on navigation ease, feature intuitiveness, and overall user experience. The empirical evaluation chapter provides a comprehensive analysis of the gathered data, shedding light on the effectiveness and user perception of the developed websites compared to their official counterparts.

This chapter also considers the ease of creating automatic websites and the level of customisation available for each website. We plan to conduct a manual analysis to evaluate these aspects to determine if the implemented solution effectively meets these requirements. By examining the process of website creation and the range of customisation options provided, we aim to assess the user-friendliness and flexibility of the solution, ultimately gauging its ability to meet the diverse needs of different clubs.

We begin this chapter with 5.1 by providing a detailed account of the implemented questionnaires distributed to the participants. The preliminary assessment conducted to gauge initial findings is discussed in Section 5.2. Moving forward, the results obtained from the questionnaires, along with the outcomes of data tests, are presented in Section 5.3. Section 5.4 offers a comprehensive discussion addressing the research questions posed in this study. Finally, the validity threats inherent in our evaluation process, including mitigated and remaining threats, are elucidated in Section 5.5.

5.1 Questionnaires

The study employed two questionnaires to gather data and evaluate different aspects of the developed websites.

The first questionnaire targeted supporters of low-ranking football clubs (*GD Aldeia Nova*[1], *Leça FC*[2], *Perafita*[3], and *Vila FC*[4]). It aimed to assess whether the developed websites were perceived as more up-to-date than the official websites. Additionally, it sought to evaluate the visual appeal of the designed websites and determine if they adequately represented the respective clubs. We developed one questionnaire for these participants, who had to choose the club they supported.

The second questionnaire was distributed among supporters of first-league clubs (*SL Benfica*[5], *FC Porto*[6], *Sporting CP*[7], and *SC Braga*[8]). It aimed to assess differences in usability between the official and developed websites. Additionally, it aimed to compare the performance of the two websites in completing specific tasks. We developed four questionnaires for these participants, varying the club to separate the fans and mitigate possible negative results due to rivalry.

Only Skin 1 was tested for both questionnaires since it is the most stable Skin for participants to perform the tests.

Both questionnaires utilised scientific research principles and incorporated a range of question formats, including multiple-choice and Likert scale [27] responses, to gather data. The participants were selected based on their affiliation with the football clubs and their role as supporters. The confidentiality of the participants was prioritised by the company conducting the research.

[1] https://aldeianova.00.pt/
[2] https://leca.00.pt/
[3] https://perafita.00.pt/
[4] https://vilafc.00.pt/
[5] https://benfica.00.pt/
[6] https://fc-porto.00.pt/
[7] https://sporting.00.pt/
[8] https://sc-braga.00.pt/

The participants were requested to respond honestly to the questionnaires, as they were part of a validation process for an ongoing investigation. Anonymity was ensured to encourage sincere and unbiased responses further.

Additionally, both questionnaires were developed and administered in Portuguese. According to the company conducting the research, it was determined that most participants would feel more comfortable responding in their native language rather than in English. This decision was made to ensure a higher level of participant engagement and minimise language-related barriers that could hinder accurate and meaningful responses. The final questionnaires are in Appendix C (p. 107).

5.1.1 Common Questions Between Questionnaires

Specific questions were included in the low-ranking and first-league clubs' questionnaires to assess common aspects related to participants' demographics, football knowledge, and website usability. These questions aimed to gather essential information and evaluate users' perceptions across both questionnaire sets. The following questions were included in both questionnaires:

Demographic Information

Age: This question aimed to gather demographic information about the participants, allowing for further analysis considering different age groups.

Gender: This question aimed to collect demographic information regarding the participants' gender distribution.

Device: This question aimed to gather information about the devices participants use to access websites, providing insights into the influence of device type on their responses.

Football Knowledge

P1.1: This question focused on assessing participants' football knowledge, aiming to examine its potential influence on their subsequent responses.

P1.2: This question aimed to evaluate participants' knowledge regarding using sports team websites, exploring its potential impact on their subsequent responses.

Usability Assessment

The usability assessment in both questionnaires included the following questions, derived from SUS template, to evaluate the usability of the developed website:

SUS.1: *I would like to use this website frequently.*

SUS.2: *This website is unnecessarily complex.*

SUS.3: *This website is easy to use.*

SUS.4: *I would need assistance to be able to use this website.*

SUS.5: *The various functions of this website are well integrated.*

SUS.6: *This website is too inconsistent.*

SUS.7: *Most people would quickly learn to use this website.*

SUS.8: *This website is complicated/uncomfortable to use.*

SUS.9: *I felt confident using this website.*

SUS.10: *It is necessary to acquire new knowledge before being able to use this website.*

These questions were specifically chosen as they are part of the established SUS template for evaluating website usability. Each question addresses a different aspect of usability, such as ease of use, complexity, integration of functions, and user confidence. Collectively, these questions provide a comprehensive assessment of the usability of the developed website.

5.1.2 Low-ranking Clubs Questions

The low-ranking clubs' questionnaire included questions to assess various aspects related to participants' demographics, football knowledge, website usability, and perceptions of the developed website. The questions used in the questionnaire, besides the ones presented previously, along with their objectives, are described below:

Football Knowledge

P1.3: This question focused on assessing participants' knowledge of Portuguese football clubs outside the top division, aiming to examine its potential influence on their subsequent responses.

Club Affiliation

Club: This question aimed to verify whether participants' affiliation with a specific football club would affect their responses to the subsequent questions.

Feedback Statements regarding Website Usability

P2.1: This question aimed to determine whether participants could recognise that the developed website was updated with relatively recent news (within the last seven days).

P2.2: This question aimed to determine whether participants could recognise that the developed website was updated with relatively recent matches (within the last seven days).

P2.3: This question aimed to assess participants' perception of whether the developed website included all the necessary sections.

P2.4: This question aimed to gauge participants' perception of the developed website's attractiveness in terms of appearance.

P2.5: This question aimed to assess participants' ability to quickly find senior and youth team information on the developed website, examining the usability aspects.

Comparison with Official Website

P3.1: This question aimed to determine whether participants considered the developed website to have more up-to-date news than the official website.

P3.2: This question assessed whether participants believed the developed website had more up-to-date match information (upcoming matches and recent results) than the official website.

P3.3: This question aimed to evaluate participants' perception of the developed website having more relevant sections than the official website.

5.1.3 First-league Clubs Questions

The first-league clubs' questionnaire included a comprehensive set of questions to evaluate different aspects related to participants' demographics, football knowledge, website usability, and perceptions of the developed and official website. The following questions were used in the questionnaire, along with their corresponding objectives:

Football-related Knowledge

P1.3: In the domain of sports information websites (e.g., *zerozero*, online newspapers): Assessing participants' knowledge of sports information websites to determine its influence on subsequent responses.

Importance of Website Features

YouthTeams: Assessing if participants need to be able to follow youth teams (view results, upcoming games, news, squad) on a club's website.

Sports: Assessing if participants need to be able to follow all sports modalities (view results, upcoming games, news, squad) on a club's website.

Verification of Official Website Use

P2: *How many football teams can be viewed on the official website?* Task to verify if participants consulted the official website for the questionnaire.

Assessment of Official Website

P3.0: Regular usage of the official website: Determining if the participants' usage affects the results obtained.

P3.1: Currency of information on the official website: Determining if participants notice that the official website's information is up to date.

P3.2: Ease of obtaining information (news, games, results, squads) about other modalities besides football: Evaluating the usability of the official website by asking participants to perform a task and assess the ease with which they completed it.

P3.3: Ease of obtaining information (news, games, results, squads) about non-senior teams (in football): Evaluating the usability of the official website by asking participants to perform a task and assess the ease with which they completed it.

P3.4: Ease of obtaining information (news, games, results, squads) about non-senior teams in modalities other than football: Evaluating the usability of the official website by asking participants to perform a task and assessing the ease with which they completed it.

Verification of Developed Website Use

P4: *How many 11-a-side football teams can be viewed on the developed website?* Task to verify if participants consulted the developed website for the questionnaire.

Assessment of Developed Website

P5.1: Currency of information on the developed website: Determining if participants notice the developed website's information is up to date.

P5.2: Ease of obtaining information (news, games, results, squads) about other modalities besides football: Evaluating the usability of the developed website by asking participants to perform a task and assess the ease with which they completed it.

P5.3: Ease of obtaining information (news, games, results, squads) about non-senior teams (in football): Evaluating the usability of the developed website by asking participants to perform a task and assess the ease with which they completed it.

P5.4: Ease of obtaining information (news, games, results, squads) about non-senior teams in modalities other than football: Evaluating the usability of the developed website by asking participants to perform a task and assess the ease with which they completed it

5.2 Preliminary Assessment

Three pilot tests were conducted for each questionnaire to ensure clarity and understanding among participants. These pilot tests aimed to assess whether different participants comprehended the purpose of the questionnaires.

In the first questionnaire, participants were asked about the presence of recent news and games. However, during the pilot tests, it became evident that the "recent" concept was ambiguous. To address this, a clarification was added that news and games should be within the last seven days to be considered recent. Furthermore, participants needed clarification regarding distinguishing between the official and developed websites. To alleviate this confusion, hyperlinks to the websites were included in the questions, clearly identifying the official and designed websites.

In the second questionnaire, in addition to the concerns from the first questionnaire, participants had further queries regarding the teams to consider in the task-related questions. To address this, it was specified that the groups mentioned should be football teams. Additionally, for the developed website, it was determined that only 11-a-side football teams should be considered.

During the pilot tests for both questionnaires, participants expressed discontent with the usability questions about the developed website. This was primarily due to potential confusion arising from the rating scale used in the SUS Template [31]. The scale assigned a maximum score of 5 to odd-numbered questions and 1 to even-numbered questions, making some participants uncertain how to interpret and accurately respond to the questions.

However, despite the participants' dissatisfaction, it was decided to maintain the template for two primary reasons. Firstly, the template assists in identifying participants who may hastily complete the questionnaire without thoroughly reading the questions [20]. Secondly, maintaining consistency with the template allows for a standardised approach to measuring participants' usability.

The questionnaires were refined by conducting these preliminary assessments and making necessary adjustments based on participant feedback to enhance clarity, minimise ambiguity, and ensure a better understanding of the survey items.

It should be noted that the results from the three pilot tests conducted were not included in the analysis of the questionnaire responses. By excluding the pilot test responses from the dataset, the analysis focused solely on the responses collected from the main administration of the questionnaires, ensuring data validity and minimising potential biases. This approach allows for a more accurate evaluation of the targeted research objectives and enhances the reliability of the findings.

5.3 Results

In this section, we present the obtained results from the questionnaires distributed to participants and relevant statistics that aided their evaluation.

5.3.1 Low-ranking Clubs

A total of 35 individuals initially participated in the study. However, after reviewing the responses from the SUS questionnaire, it was determined that two participants provided consistently inconsistent answers. These participants responded similarly to all questions but contradicted themselves within their responses. As a result, these two individuals were deemed invalid and were

excluded from the analysis. Consequently, the final sample consisted of 33 participants whose responses were considered for further analysis. In the following sections, we will present the results and findings based on the responses of these 33 participants, providing insights into their perceived usability and user experience of the websites under evaluation.

5.3.1.1 Demographic Information

The initial segment of the questionnaire focused on examining the demographic characteristics of the participants. In terms of gender distribution, 26 participants, or 78.8%, identified as male, 6 participants, or 18.2%, identified as female, while 1 participant, or 3%, preferred not to disclose their gender. In terms of age distribution, the participants represented various age groups, with 15.2% (5 participants) falling within the 18-24 age range, 24.2% (8 participants) within the 25-34 age range, 48.5% (16 participants) within the 35-44 age range, and 12.1% (4 participants) within the 45-54 age range. Moreover, when considering the device used, most participants (72.7% or 24 participants) reported using a computer, while 27.3% (9 participants) used a mobile phone. Furthermore, the participants' club preferences varied, with 39.4% (13 participants) supporting *Leça FC*, 24.2% (8 participants) supporting *Vila FC*, 27.3% (9 participants) supporting *GD Aldeia Nova*, and 9.1% (3 participants) supporting *Perafita*.

5.3.1.2 Feedback Statements Regarding Website Usability

The present questionnaire included feedback questions to assess participants' opinions regarding the website. We analysed the feedback statements concerning the websites developed for their respective clubs. Participants were asked to rate their perception using a Likert scale of 1 ("Strongly Disagree") to 5 ("Strongly Agree"). The purpose was to assess whether the generated website was regarded as current, included all required elements, and was user-friendly. Descriptive statistics were computed to summarise the data, yielding the minimum, maximum, mean, median, mode, first quartile, third quartile and standard deviation for each feedback statement and can be visualised in Figure 5.1 (p. 57).

After analysing the presented results, we can assert that P2.2 received the highest mean score of 4.70, indicating that participants generally responded positively. Moreover, it had a low standard deviation of 0.85, indicating substantial agreement among the responders. Similar mean scores for statements P2.1, P2.4, and P2.5—ranging from 4.42 to 4.45—indicate favourable ratings overall. However, P2.1 had a higher standard deviation of 1.09, indicating more significant response variability. P2.3 received a slightly lower mean score of 4.15, suggesting a relatively favourable but slightly less favourable evaluation. Outliers were identified in several statements, P2.1 having one outlier (1), P2.2 having three outliers (1, 3 and 4), P2.4 having one outlier (3), and P2.5 having two outliers (2 and 1). These outliers indicate deviations from the overall pattern of higher scores. Overall, the data suggests generally positive evaluations for all statements, with P2.2 receiving the highest mean score and exhibiting the least variability among responses.

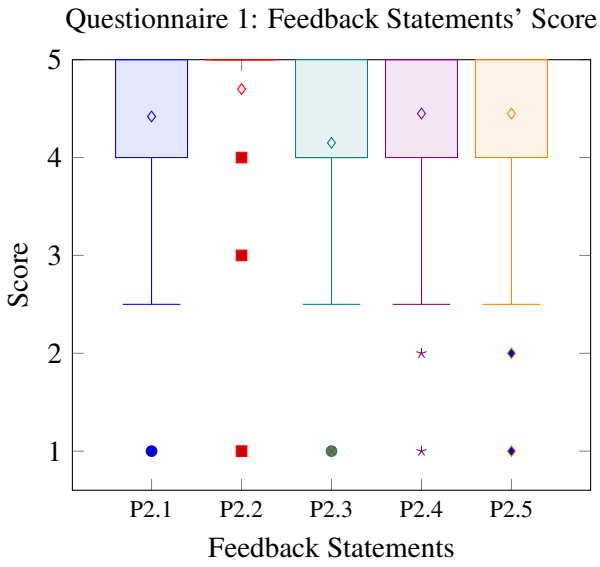


Figure 5.1: Boxplot of the feedback statements' scores for the low-ranking clubs' questionnaire.

These findings suggest that participants had a favourable perception of the developed website's recent match updates (P2.2) and generally positive evaluations of other aspects such as recognising recent news (P2.1), necessary sections (P2.3), website attractiveness (P2.4), and ease of finding team information (P2.5). However, there is variability in participants' responses to specific statements, indicating areas where improvements or further investigation may be beneficial.

5.3.1.3 Comparison with Official Website

Following the analysis, we investigated the significance of user preferences between the official and developed websites for questions P3.1, P3.2, and P3.3. This inquiry employed the Monte Carlo Simulation technique [32]. Monte Carlo methods, or Monte Carlo experiments, "encompass a wide range of computational algorithms that rely on repeated random sampling to obtain numerical results"[43]. These methods utilise randomness to solve problems that may, in principle, be deterministic.

In applied statistics, Monte Carlo methods are utilised to implement hypothesis tests efficiently. These tests offer advantages over exact tests, such as permutation tests, which are often computationally infeasible, while also providing greater accuracy than critical values derived from asymptotic distributions. This study's null hypothesis ($\boldsymbol{H_0}$) was defined as "The official website, and the developed website are equivalent." In contrast, the alternative hypothesis ($\boldsymbol{H_1}$) stated, "The developed website exhibits superior usability and more up-to-dateness than the official website." Consequently, a Monte Carlo simulation was performed.

The code implementation for this simulation is provided in Appendix D (p. 117). Through the execution of 1,000,000 iterations, the calculated p-value consistently yielded values less than 0.0001 for P3.1, P3.2, and P3.3. These results suggest that the observed data is highly improbable

under the null hypothesis at a significance level of 0.05. Thus, statistical evidence exists to reject the null hypothesis.

In light of these findings, participants perceived the developed website as more user-friendly and up-to-date than the official website.

5.3.1.4 Usability Assessment

In the final section of the questionnaire, we used the SUS Template to assess the website's usability more effectively. The participants responded on a scale of 1 ("Strongly Disagree") to 5 ("Strongly Agree") for each question.

In the first approach, we have constructed a boxplot for each question to analyse the data collected through the SUS Template questions to facilitate interpretation. The boxplot can be seen in Figure 5.1 (p. 57).

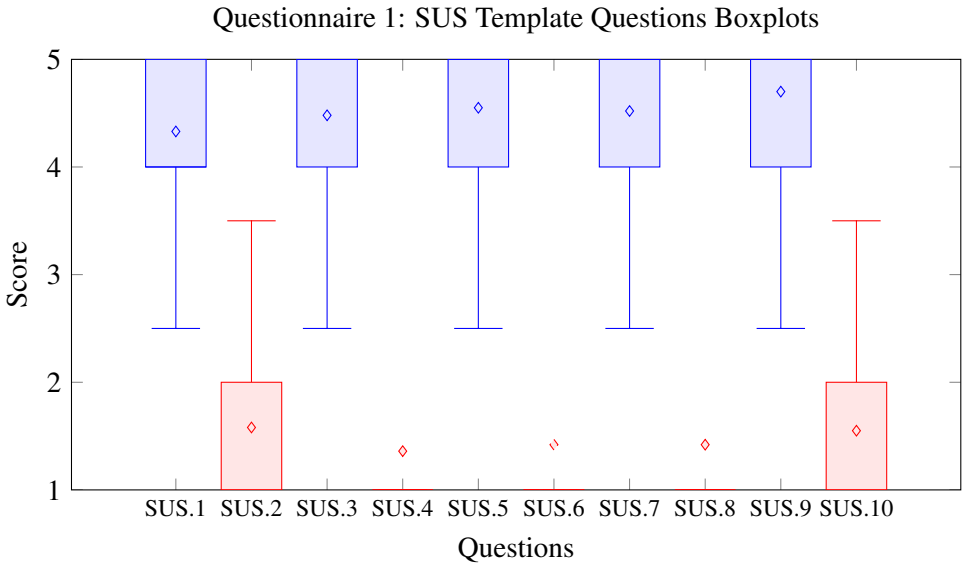


Figure 5.2: Boxplot of the results of each question of the SUS Template for the low-ranking clubs' questionnaire.

According to the data provided in the boxplots, the responses appear to be distinct patterns, dividing them into two groups: odds and evens.

- **Odd-numbered questions (SUS.1, SUS.3, SUS.5, SUS.7, and SUS.9)**: These questions have a consistent distribution, with scores predominantly clustering around the median of 5.00. The scores range from 2.50 to 5.
- **Even-numbered questions (SUS.2, SUS.4, SUS.6, SUS.8, and SUS.10)**: In contrast, these questions show a narrower range of scores, ranging from 1 to 3.50. The median score for these questions is 1.00.

In order to assess more accurately the usability of the low-ranking clubs, We had to calculate the SUS score by calculating the score of odd and even questions, as Brooke [8] outlined. We

summed the odd-number responses score for the odd score, which was obtained by subtracting 1 from the position on the response scale. In contrast, the even score was obtained by summing the even-numbered responses score, which we calculated by subtracting the position from 5. Finally, we summed the odd score to the event score and multiplied it by 2.5, resulting in the SUS Score. The SUS Score ranges from 0 to 100. Following Bangor's categorisation [7], the SUS Score can be associated with one of seven categories: "Worst Imaginable," "Awful," "Poor," "OK," "Good," "Excellent," and "Best Imaginable." The SUS Scores are displayed in a histogram in Figure 5.3, allowing for a comprehensive examination of the usability evaluation outcomes.

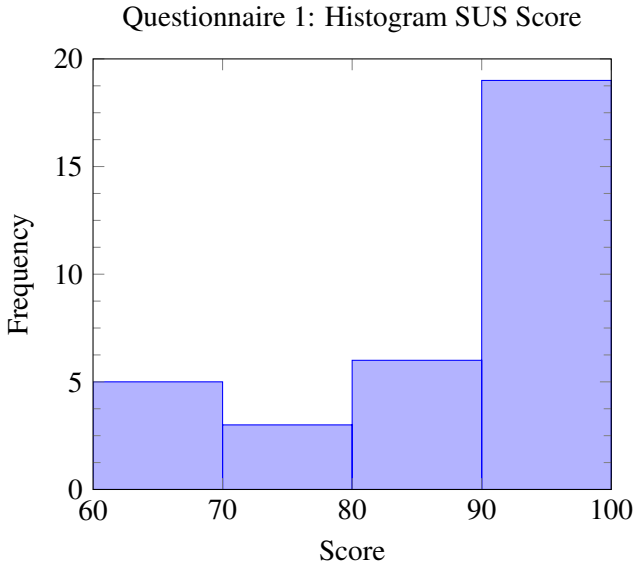


Figure 5.3: Histogram of the SUS Scores for the low-ranking clubs' questionnaire.

The analysis of the SUS Scores, obtained from the SUS Score calculation method described earlier, revealed valuable insights into the website's usability. The SUS Scores exhibited a range from a minimum of 60.00 to a maximum of 100.00, indicating a considerable variability in participants' evaluations. The mean SUS Score was 88.11, suggesting a positive perception of the website's usability among the participants. According to Bangor, 88.11 is considered "Excellent" usability.

According to the histogram, a significant proportion of responses indicated SUS Scores within the range of 90 to 100, providing further support for the previously discussed findings. The histogram demonstrates a prominent peak in frequency within this range, suggesting a prevailing trend towards high levels of perceived usability. This concentration of scores in the upper range reinforces the notion that a substantial portion of participants evaluated the website favourably regarding usability. The distribution exhibits a gradual decline in frequency as the scores move away from the peak, indicating a diminishing number of responses with lower ratings. The histogram's visual depiction aligns with the previous analysis, affirming the predominant perception of superior website usability among the study participants.

Based on the results from the SUS questions, the data reveals encouraging findings regarding the usability of the developed website. Participants' responses exhibited distinct patterns, with odd-numbered questions receiving predominantly positive evaluations. In contrast, even-numbered questions garnered lower scores, which is expected for optimal results in the SUS Template questionnaire. The overall SUS Scores indicated excellent usability, with a mean score of 88.11. The histogram analysis reinforced this positive perception, showing a significant concentration of responses in the high range of 90 to 100. These findings suggest that most participants viewed the website favourably in terms of usability, emphasising its effectiveness and user-friendly design.

5.3.1.5 Correlation Analysis

In order to gain additional insights for data interpretation, we aimed to investigate potential correlations among the responses provided for each questionnaire item. To accomplish this, we transformed all qualitative data into a quantitative form, allowing us to utilise the correlation matrix calculation method provided by the Python library pandas[9]. The following conversions were applied to the qualitative response questions:

- **Age conversion**: Age groups were numerically represented as follows: 18-24 (1), 25-34 (2), 35-44 (3), and 45-54 (4).
- **Gender assignment**: Gender categories were assigned numeric values: Male (1), Female (2), and Unspecified (3).
- **Device type representation**: Device types were encoded as Mobile (1) and Computer (2).
- **Club affiliation encoding**: Club affiliations were represented by the following numeric values: GD Aldeia Nova (1), Leça FC (2), Perafita (3), and Vila FC (4).
- **Preference between Official and Developed websites:**: The preference between the Official and Developed websites is indicated by the values: Official (1) and Developed (2).

Given that the data exhibited non-continuous variables and deviations from a normal distribution, we opted for the Spearman correlation method. The Spearman correlation coefficient measures the strength and direction of the monotonic relationship between variables that are usually censored and not normally distributed and ranges from -1 to +1 [25]. This approach allowed us to assess the association between variables despite their non-linear or non-normally distributed nature.

We employed the Python library seaborn[10] to construct a heatmap based on the correlation matrix values to enhance data visualisation. The resulting matrix is depicted in Figure 5.4 (p. 61),

[9] https://pandas.pydata.org/
[10] https://seaborn.pydata.org/

illustrating the strength and direction of the correlations among the variables. The heatmap provides a comprehensive overview of the interrelationships between different questionnaire items, enabling a deeper understanding of the potential associations within the dataset.

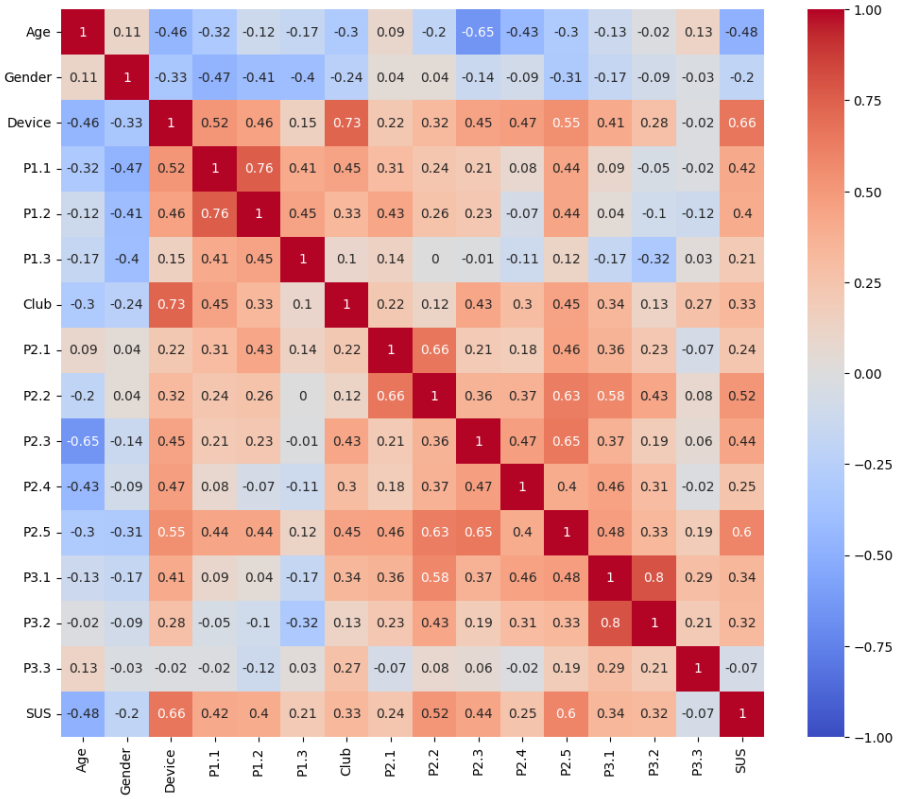


Figure 5.4: Correlation Matrix of the results obtained for the low-ranking clubs' questionnaire.

The identifiers for each question are determined based on the content of the respective question. In the case of questions with multiple sub-items, the identifiers are constructed as "P[*question number*].[*sub-item number*]". These identifiers are used for subsequent analysis of variable correlation.

The results highlight several significant positive and negative correlations among the variables examined. The following positive, relevant correlations were identified, shedding light on the interconnectedness and potentially shared constructs within the study:

- **Device and SUS (Spearman's r = 0.66)**: A positive association exists between the device used and perceived usability, indicating that when using a computer to answer the questionnaire, participants think the website has better usability.

- **Device and P2.5 (Spearman's r = 0.55)**: Higher device usage is positively correlated with the responses to P2.5, indicating that participants can more easily find information regarding senior and youth teams on the website when they use a computer to answer to the questionnaire.

- **P2.2 and P3.1 (Spearman's r = 0.58)**: The positive association between P2.2 and P3.1 suggests a potential alignment or similarity between the constructs represented by these survey items, suggesting that participants that consider that the developed website has all the needed sections also consider that the developed website is more up-to-date than the official one.

Although there are other correlations with bigger Spearman's r values, we concluded that the correlations were irrelevant for interpreting results after analysing the questions.

Contrary to the positive correlations, several negative associations emerged, underscoring the potential disparities and contrasting relationships within the investigated variables. The following negative correlations were identified, shedding light on the divergent nature and potential conflicts within the study:

- **Age and P2.3 (Spearman's r = -0.65)**: Negative correlation suggests that responses to P2.3 decrease as age increase, showcasing that as participants grow older, they increasingly perceive that the website lacks the necessary relevant sections.
- **SUS and Age (Spearman's r = -0.48)**: Negative association suggests that usability decreases as age increases, indicating that as participants grow older, they are less likely to perceive that the website has good usability.

Similarly to the positive correlations, correlations with smaller Spearman's r values exist, but after analysing the questions involved, we concluded that the correlations were irrelevant to the interpretation of results.

5.3.2 First-league Clubs

For the first-league questionnaire targeting supporters of *SL Benfica*, *FC Porto*, *Sporting CP*, and *SC Braga*, we collected 39 responses. However, nine individuals were excluded from the study due to inconsistencies in their replies or errors in task completion. As a result, we obtained 30 valid responses for analysis. This subsection will present the results and findings derived from these 30 legitimate responses. We will provide insights into the usability, navigation ease, feature intuitiveness, and overall user experience of the first-league clubs' official and developed websites.

5.3.2.1 Demographic Information

Men comprised 76.7% of the sample participants (23 participants), while women comprised 23.3% of the sample, according to gender distribution (7 participants). Regarding age, most respondents (53.3%) fell between 18 and 24. (16 participants). Moreover, just 3.3% (1 participant) of the population was between the ages of 45 and 54, as opposed to 16.7% (5 participants) and 26.7% (8 participants) in the 25–34 age range. Only a small percentage of participants (10.0%) used mobile phones, whereas the majority (90.0%) used computers to complete the questionnaire. Second place went to *Sporting CP* with 20.0% (6 participants), third place went to *SL Benfica* with 13.3%

(4 participants), and fourth place went to *SC Braga* with 6.7% of the evaluated clubs. *FC Porto* came in first with 60.0% (18 participants) (2 participants).

Notably, the choice of the device appeared as a critical factor potentially affecting the outcomes, but gender and age did not significantly influence the results.

5.3.2.2 Importance of Website Features

In this study, firstly, we investigated the perceived importance of information regarding youth teams and non-football sports on official club websites among participants. To gather participants' opinions, we analysed the percentage of responses indicating "Yes" and "No" to the YouthTeams and Sports questions. The results are presented in Table 5.1.

Table 5.1: Percentage of "Yes" and "No" answers to the questions YouthTeams and Sports

	Yes	**No**
YouthTeams	86.67%	13.33%
Sports	96.67%	3.33%

According to the findings in the table, a significant proportion of participants believed that the presence of youth team information and information on non-football sports is crucial on club websites. These participants assigned greater importance to including information about sports other than football.

5.3.2.3 Feedback statements regarding the official and the developed websites

Subsequently, participants engaged in a questionnaire comprising inquiries and feedback statements, necessitating their assessment on a Lickert scale for official and developed websites. This methodology facilitated a comparative examination of the outcomes between the two websites.

Respondents appraised the abovementioned inquiries using a Lickert scale from 1 ("Strongly Disagree") to 5 ("Strongly Agree"), whereby higher scores denoted more positive evaluations. We formulated the feedback statements to measure participants' perceptions regarding the timeliness of website updates (Q1), ease of accessing information about non-football sports (Q2), convenience in obtaining information about youth football teams (Q3), and ease of accessing information regarding youth teams in sports other than football (Q4).

Boxplots were generated to portray the acquired findings, wherein the outcomes associated with the official website (P3.1, P3.2, P3.3, and P3.4) were represented in a hue of purple. In contrast, the results concerning the developed website (P5.1, P5.2, P5.3, and P5.4) were depicted in a teal hue. These boxplots are visually represented in Figure 5.5 (p. 64).

Several observations can be made comparing the results between the official and the developed websites. Overall, the statistics indicate that the ratings for the designed website are higher and exhibit less variability compared to the official website. Specifically, focusing on question Q4, which pertains to obtaining information about youth teams in sports other than football, a significant difference is apparent. The scores for this question on the developed website range from

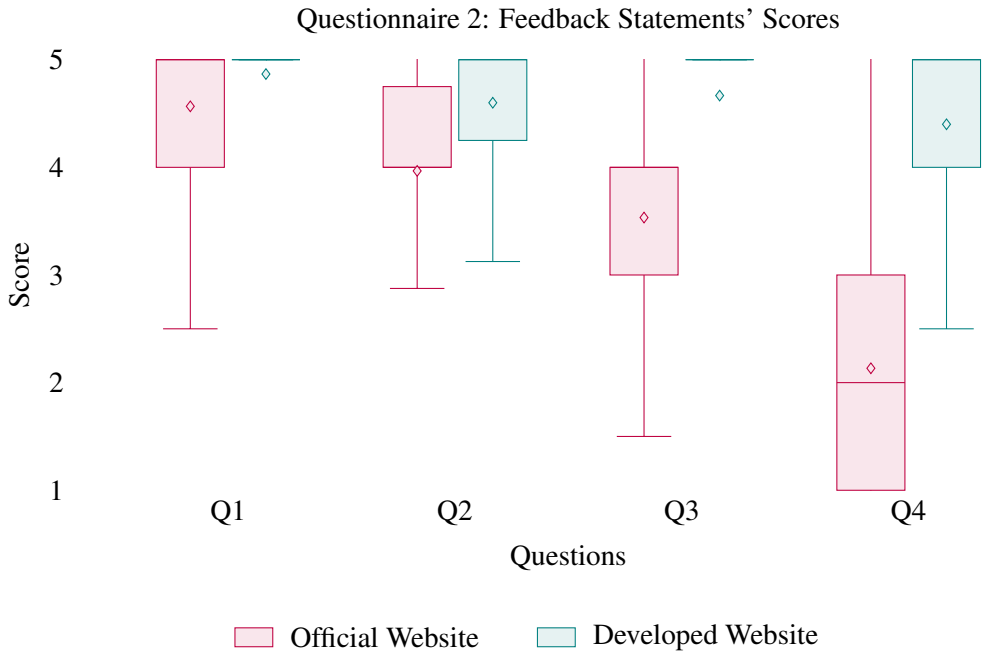


Figure 5.5: Boxplot of the feedback statements' scores of the official website and developed website for the first-league clubs' questionnaire.

4.25 to 5.00, with a median of 5.00, indicating a consistently high rating. In contrast, the official website shows more significant variability, with scores ranging from 1.00 to 3.00 and a median of 4.00.

Additionally, we aimed to assess whether the obtained results for the questions from P3 (P3.1, P3.2, P3.3 and P.34) and P5 (P5.1, P5.2, P5.3 and P5.4) followed a normal distribution to determine the appropriate statistical methods for subsequent analysis. Similar to the methodology employed in the questionnaire for low-ranking clubs, we conducted the Shapiro-Wilk test [30] on the obtained results. The test results for both datasets are presented in the Table 5.2.

Table 5.2: Shapiro-Wilk's normality test results for the first-league clubs' questionnaire feedback statements.

		Test Statistic (W)	**p-value**	**Normality Assessment**
P3	1	0.6862	9.899×10^{-7}	Significant departure from normality
	2	0.8022	7.182×10^{-5}	Significant departure from normality
	3	0.8852	3.716×10^{-3}	Significant departure from normality
	4	0.8163	1.317×10^{-4}	Significant departure from normality
P5	1	0.2776	4.588×10^{-11}	Significant departure from normality
	2	0.5455	1.679×10^{-8}	Significant departure from normality
	3	0.5422	1.542×10^{-8}	Significant departure from normality
	4	0.6499	3.145×10^{-7}	Significant departure from normality

The results of the Shapiro-Wilk test indicate a significant departure from normality for both the official website and the developed website variables. For the P3 questions, test statistics (W) range

from 0.6862 to 0.8852, with corresponding p-values ranging from 9.899×10^{-7} to 3.716×10^{-3}. Similarly, for the P5 questions, test statistics range from 0.2776 to 0.6499, with corresponding p-values ranging from 4.588×10^{-11} to 3.145×10^{-7}. These findings suggest that the data distributions deviate significantly from a normal distribution, highlighting the need to employ appropriate statistical methods for subsequent analysis.

Given the notable deviation from a normal distribution observed in the results, signifying a significant departure from normality, and acknowledging the non-parametric nature of the paired data, we have employed the Wilcoxon signed-rank [26] test to derive further insights. This statistical test has been selected due to its independence from normality assumptions and suitability for analysing paired data. By utilising the Wilcoxon signed-rank test, we aim to ascertain the presence of statistically significant differences between the paired observations.

The outcomes from applying the Wilcoxon signed-rank test are presented in Table 5.3, briefly summarising the pertinent analytical findings.

Table 5.3: Wilcoxon signed-rank test results comparing the results obtained in the questions comparing the official website and the developed website (P3 and P5).

	Sample Size (N)	**Test Statistic (W)**	**Mean Difference**	**Std. Deviation**	**p-value**
P3.1 & P5.1	11	6	-1.18	11.25	.0082
P3.2 & P5.2	16	18	-0.06	19.34	.0048
P3.3 & P5.3	22	15	-0.64	30.8	.00015
P3.4 & P5.4	26	7.5	-2.15	39.37	<.00001

The Wilcoxon signed-rank test results showed significant differences between the official website (P3.1, P3.2, P3.3, P3.4) and the developed website (P5.1, P5.2, P5.3, P5.4). The sample sizes varied between 11 and 26, with corresponding test statistics ranging from 6 to 18. The mean differences between the developed and official websites ranged from -2.15 to -0.06. The standard deviations varied from 11.25 to 39.37. The p-values ranged from less than .00001 to .0082, indicating statistically significant differences in user ratings between the two website versions.

Considering the null hypothesis ($\boldsymbol{H_0}$) posits that the official website and the developed website are equivalent. The alternative hypothesis ($\boldsymbol{H_1}$) that the developed website exhibits better usability compared to the official website and is more up-to-date, we can conclude that these findings provide strong evidence to reject $\boldsymbol{H_0}$, indicating that the $\boldsymbol{H_1}$ is supported. These findings suggest that participants perceived the developed website as more user-friendly and effective in meeting their needs. These results highlight the importance of further improving and promoting the developed website's usability to enhance user satisfaction and engagement.

5.3.2.4 Usability Assessment

In the final section of the questionnaire, the SUS template was employed to assess the website's usability more effectively. Participants were asked to rate the usability questions using a scale of 1 to 5. Similar to the analysis conducted for the low-ranking clubs, boxplots were generated

to visualize the data for each usability question. These boxplots, illustrating the distribution of responses, are depicted in Figure 5.6.

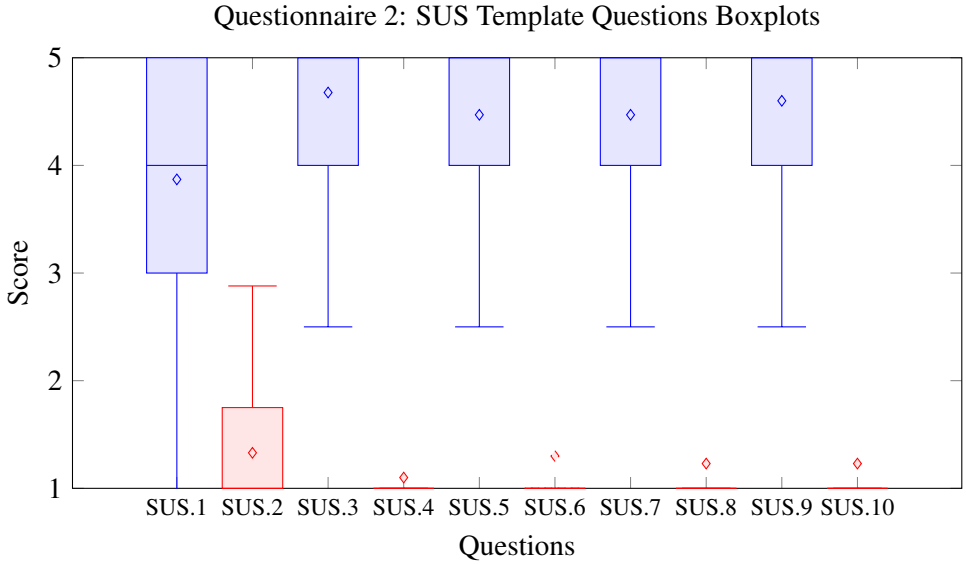


Figure 5.6: Boxplot of the results of each question of the SUS Template for the first-league clubs' questionnaire.

According to the boxplots, the responses to the SUS template questions exhibit distinct patterns when divided into odd and even-numbered questions.

- **Odd-numbered questions (SUS.1, SUS.3, SUS.5, SUS.7, and SUS.9)**: the scores cluster around the median of 4.00. The scores range from 3.00 to 5.00.
- **Even-numbered questions (SUS.2, SUS.4, SUS.6, SUS.8, and SUS.10)**: narrower range of scores, from 1.00 to 1.75. The median score for these questions is 1.00.

Among the ten questions in the SUS Template, SUS.1 exhibited the highest variability in participants' responses. The boxplot analysis reveals a broader range of scores from 3.00 to 5.00. The median score of 4.00 suggests moderate agreement regarding the website's overall usability. However, the noticeable dispersion of scores indicates a greater diversity of opinions and experiences among the participants.

Following the analysis of individual question results, we computed the SUS Score for each response in the questionnaire using the previously explained method applied to the low-ranking clubs' questionnaire. The histogram in Figure 5.7 (p. 67) depicts the distribution of the obtained SUS Scores. This graphical representation illustrates the frequency of SUS Scores, providing insights into the overall usability perception of the website among the participants.

The histogram reveals that most participants rated the website's usability with higher scores, specifically in the range of 90-100, indicating a positive perception of the website's usability among the respondents. On the other hand, fewer participants assigned scores in the lower range

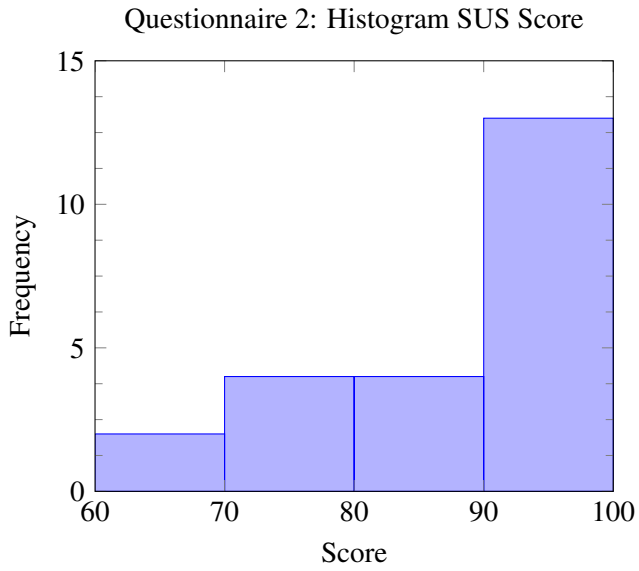


Figure 5.7: Histogram of the SUS Scores for the first-league clubs' questionnaire.

(60-70). The average of the scores obtained is 89.67, and, according to Bangor, the website usability is classified as "Excellent".

Based on these findings, the odd-numbered questions in the SUS Template, representing various aspects of website usability, received relatively higher scores and exhibited a narrower range, suggesting a more consistent and positive perception of usability among participants for these specific aspects. In contrast, the even-numbered questions had lower scores and a narrower range. These results are expected in an optimal SUS template questionnaire. The analysis of individual question results also revealed moderate agreement and a diverse range of opinions among participants regarding the overall usability of the developed website. However, when calculating the SUS Score, which combines responses from all questions, it was found that most participants rated the website's usability highly, particularly in the range of 90-100, suggesting the respondents' predominantly positive perception of the developed website's usability. The average SUS Score of 89.67 further supports the classification of the website's usability as "Excellent" according to Bangor's criteria.

5.3.2.5 Correlation Analysis

Finally, we constructed a correlation matrix encompassing all the questionnaire items to assess the potential correlations among the questionnaire responses, similar to the low-ranking club's questionnaire approach. To facilitate this analysis, we converted the qualitative data into quantitative format. Specifically, we assigned numerical values to represent the qualitative responses as follows:

- **Age conversion**: Age groups were numerically represented as follows: 18-24 (1), 25-34 (2), 35-44 (3), and 45-54 (4).

- **Gender assignment**: Gender categories were assigned numeric values: Male (1) and Female (2).
- **Device type representation**: Device types were encoded as Mobile (1) and Computer (2).
- **Importance-related questions (YouthTeams, Sports)**: Answers encoded as "Yes" (1) and "No" (2).
- **Task-related questions**: Answers encoded as "Correct" (1) and "Incorrect" (2).

Given the non-continuous nature of the data and the presence of variables that did not adhere to a normal distribution, we employed the Spearman correlation method to calculate the correlation matrix. The Spearman coefficient measures the strength and direction of the monotonic association between variables, ranging from -1 to +1.

To enhance the visualisation of the correlation matrix, we utilised the Python Seaborn library to construct a heatmap based on the correlation values. The resulting correlation matrix heatmap is presented in Figure X. Each question is identified by its respective question number. For queries with multiple subparts, it is represented as "P[*question number*].[*sub-item number*]".

Correlation values close to 1 indicate a positive correlation, suggesting a strong and direct relationship between the variables. On the other hand, correlation values near -1 indicate a negative correlation, implying a robust inverse relationship between the variables.

The findings reveal multiple significant positive correlations between the examined variables, indicating interconnections and potentially shared constructs within the study. The following positive, relevant associations highlight the relationships among the variables and provide valuable insights into their interconnected nature:

- **Sports and YouthTeams (Spearman's r = 0.47)**: positive correlation between the importance of information regarding different sports besides football on a website and the importance of information regarding youth teams on a website.
- **Device and SUS (Spearman's r = 0.45)**: positive correlation indicates that the device used impacts the website's usability.
- **Device and P3.1 (Spearman's r = 0.47)**: positive correlation showcases that the ability of the user to determine if the official website is up to date is impacted by the website used.

After carefully examining the questions, we determined that despite the presence of larger Spearman's r values in other correlations, correlations not mentioned were not significant for interpreting the results.

In contrast to the positive correlations observed, negative associations underscore potential disparities and contrasting relationships within the investigated variables. The following negative relevant correlations shed light on the divergent nature and potential conflicts within the study:

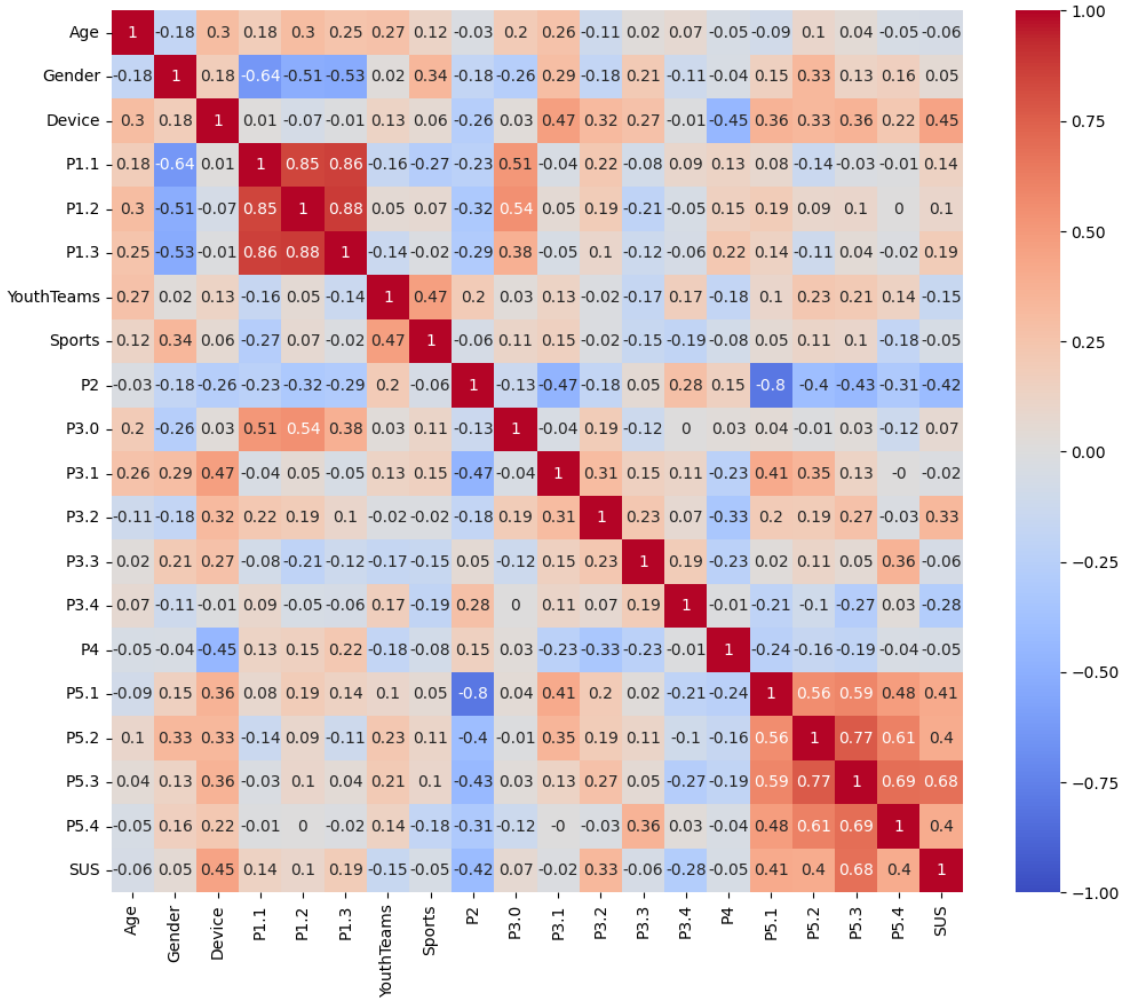


Figure 5.8: Correlation Matrix of the results obtained for the first-league clubs' questionnaire.

- **P4 and Device (Spearman's r = -0.45)**: negative correlation showcases that the device used affects the answer given in the task (participants gave wrong answers when doing the test on a mobile phone).
- **P2 and P3.1 (Spearman's r = -0.45)**: this negative correlation indicates that exists a correlation between wrong answers in the task of identifying the number of teams presented in the official website and if the official website is up to date.

Likewise, we found correlations with smaller Spearman's r values, but upon analysing the questions involved, we concluded that these correlations were irrelevant for interpreting the results.

5.4 Discussion

Based on the experimental results presented in Section 5.3 (p. 55), we will now revisit our research questions:

RQ1. *Is it possible to automatically generate websites using publicly available data that are more* ***up-to-date*** *than the club's previous ones?*

The analysis of questionnaire responses revealed that the average ratings for the updated status of the developed websites among low-ranking clubs fell within the positive range (P2.1 and P2.2 from the low-ranking clubs' questionnaire). Additionally, based on the findings obtained from questions P3.1 and P3.2 of the low-ranking clubs' questionnaires, it can be concluded that the developed websites are more up-to-date than the clubs' official websites.

In contrast, the results from first-league clubs' questionnaires indicated that the developed websites were more updated than the official ones (P3.1 and P5.1 from the first-league clubs' questionnaire).

Considering the results obtained, we can affirm that the developed websites demonstrated a higher level of currency than their official counterparts.

RQ2. *Is it possible to automatically generate websites using publicly available data that are more* ***user-friendly*** *than the club's previous one?*

Based on the results obtained from the low-ranking clubs' questionnaires regarding questions P2.3 and P2.5 (about the presence of necessary sections on the developed website and the ease of finding required information), it can be inferred that positive outcomes were achieved, indicating good usability. Additionally, the results of question P3.3 (also regarding necessary sections) further reject the hypothesis that the official website has comparable usability to the developed website, as indicated by the Monte Carlo test results, allowing us to infer again that the usability of the developed websites is better than the official websites. Additionally, the developed website received an "Excellent" rating based on the usability score derived from the responses to the SUS Template questions.

Moreover, upon analysing the results of the first-league clubs, it can be affirmed that the developed website exhibits better usability than the official website (P3.2/P5.2, P3.3/P5.3, and P3.4/P5.4). Furthermore, the developed website received an "Excellent" rating in the SUS Score from the SUS Template.

In conclusion, the developed websites demonstrate superior usability compared to the official websites.

RQ3. *Is it possible to automatically generate websites using publicly available data that require* ***minimal effort*** *for their creation?*

In order to conduct these questionnaires, eight websites were automatically generated, with four allocated to low-ranking clubs and four to first-league clubs. The website generation process requires minimal effort, as it merely involves the selection of the target club and the subsequent automated generation of the corresponding website. All generated websites exhibited flawless functionality and were consistently updated with accurate zerozero platform information.

RQ4. *Is it possible to automatically generate websites using publicly available data that allow* ***customisation****?*

As described in Subsection 4.4.2.1 (p. 41), successfully designing two available skins for the developed websites demonstrates the feasibility of creating multiple customisable skins. Despite only testing Skin 1, the absence of any encountered difficulties suggests that the process can be extended to accommodate an unlimited number of skins. This capacity for extensive customisation empowers users responsible for maintaining club websites to tailor the appearance and design elements to their specific preferences and requirements.

In general, we have high confidence that, despite specific validation threats, it is possible to automatically generate websites utilising publicly accessible data that surpass a sports club's official websites in terms of being more up-to-date and user-friendly. This achievement can be accomplished with minimal effort while maintaining the ability for extensive customisation.

5.5 Validity Threats

To rigorously assess and compare the developed website with the official websites, we deemed a questionnaire survey the most appropriate approach, considering its potential for controlled testing. Nonetheless, addressing validity threats that may compromise the study's credibility and validity is crucial. Validity threats introduce potential sources of error or bias, limiting confidence in the study's findings [18]. In quantitative research, four distinct categories of validity are recognised [13, 44]: conclusion, internal, construct, and external.

Conclusion Validity threats hinder the derivation of sturdy conclusions about the relationships between variables, while internal validity threats challenge confidence in causality by accounting for confounding influences. Construct validity threats arise from study design deficiencies or social factors that introduce biases or measurement inaccuracies, compromising the representation of underlying theoretical concepts. External validity threats restrict the generalizability of findings to broader populations or contexts, limiting the safe application or extrapolation of results beyond the specific research setting or sample [13].

In our study, we proactively addressed all these validity categories and implemented mitigation strategies whenever feasible, thus bolstering the study's scientific rigour and validity.

Conclusion Validity

Reliability of the measures Before administering the questionnaires to the participants, we conducted pilot tests to assess the effectiveness and appropriateness of the questionnaires. This pilot testing allowed us to identify any ambiguities, unclear items, or potential biases in the questionnaires. By refining and optimising the questionnaires based on the pilot test results, we enhanced the validity of the measurement instruments and reduced the likelihood of misinterpretations or measurement inaccuracies.

Statistical validity We conducted thorough statistical analyses using various techniques, including non-parametric tests like the Wilcoxon signed-rank and Monte Carlo simulations and correlation analyses. The correlation matrixes allowed us to examine the relationships between variables and assess their strength and direction. These analyses allowed us to examine the relationships between variables and assess their statistical significance accurately. Moreover, we employed graphical tools such as histograms and boxplots to visually inspect the distributions and identify potential outliers or deviations from normality. All of these techniques were useful to ensure statistical validity.

Internal Validity

History Participants of the low-ranking questionnaire also received the first-league questionnaire if they were fans of the other teams. This concurrent exposure to both questionnaires might introduce a bias in the responses, potentially impacting the results. This phenomenon can influence participants' perceptions, attitudes, and preferences, leading to altered responses compared to a scenario where participants were solely exposed to one questionnaire.

Subject Selection Despite our request for participation exclusively from fans of the respective clubs, due to the anonymous nature of data collection, we cannot ensure with certainty that participants responded to the questionnaire based on their affiliation with a specific club or even that they are fans of any club. As a result, caution should be exercised when interpreting the findings, particularly regarding the generalizability and representativeness of the responses.

Maturation To mitigate this threat, we implemented measures to ensure the duration of the questionnaires was, at most, five minutes. However, we acknowledge that some participants from the first-league questionnaire may have experienced boredom due to the similarity between the last section and the preceding sections, with the only variation being the analysed website. Given this consideration, it is essential to exercise caution when interpreting the results, despite the absence of explicit evidence indicating participant boredom. While the findings may not indicate a significant presence of boredom, the potential influence of participant fatigue or reduced engagement due to the repetitive nature of the questionnaire sections cannot be disregarded entirely.

Motivation This particular threat arises from participants' potential reluctance or apprehension to provide feedback that could influence the future direction or decisions related to their club's online presence. Despite our attempts to mitigate this validity threat by informing participants that the objective is sole to obtain feedback, we cannot guarantee the complete mitigation of this concern

Construct Validity

Mono-Operation Bias We mitigated this bias by employing a combination of objective performance metrics, such as error rates and subjective measures, including Likert scale ratings. This methodological approach enabled us to capture a more comprehensive understanding of the construct under investigation by utilising multiple indicators. By incorporating objective and subjective measures, we sought to overcome the limitations of relying solely on a single operationalisation.

External Validity

Selection Bias To mitigate this potential threat, multiple strategies were employed:

1. The study aimed to exclusively target supporters of the clubs that were the subject of investigation, ensuring that the participant sample consisted solely of individuals with a genuine interest in the respective clubs.
2. Diverse recruitment methods were employed, including contacting club supporters and company personnel willing to participate in the questionnaire. By employing multiple avenues for participant recruitment, the study aimed to minimise biases that could arise from relying on a single source or subset of the population.
3. Strict confidentiality measures were implemented to encourage participation and minimise self-selection bias, guaranteeing anonymity for participants.

However, despite these proactive measures, it cannot be guaranteed that the potential threat of selection bias was mitigated entirely. Furthermore, the requirement of confidentiality imposed by the company may have inadvertently limited the diversity of the obtained results, potentially resulting in a sample that was not demographically representative of the broader population.

User Familiarity The participants' familiarity with the club's website or comparable websites can significantly influence their perceptions of usability. When individuals are familiar with the official websites, it can engender biased expectations or pose difficulties adapting to a novel interface. We mitigated this threat by examining the potential correlation between the frequency of usage of the official club website and its usability, revealing the absence of such a relationship.

Reactive or Interaction Effect of Testing To mitigate this threat, we implemented several measures:

1. Participants were explicitly instructed to provide honest and unbiased responses, emphasising the research objective of obtaining valuable feedback.
2. We ensured the anonymity of participants' responses, creating a safe environment for open and candid feedback.

3. We conducted pilot testing to identify any potential issues or biases in the questionnaires and refined them accordingly.

Sample Size The desired sample size for our study was not achieved due to confidentiality requirements imposed by the company. It is essential to acknowledge that different results could have emerged with a larger sample size, considering the presence of outliers in some of the findings. However, no statistical evidence was found to support this claim.

Despite the diverse validity threats encountered, diligent measures were implemented to mitigate them to the best extent possible. Consequently, we maintain the assertion that the hypothesis retains its validity.

5.6 Summary

The empirical evaluation chapter focuses on examining the perception of up-to-dateness and soliciting feedback on the visual appeal and usability of the developed websites.

Section 5.1 (p. 50) provides details regarding using two questionnaires to gather data from supporters of low-ranking and first-league football clubs. These questionnaires incorporated diverse questions encompassing demographic information, football knowledge, and usability assessment inquiries. The low-ranking clubs' questionnaire additionally encompassed club affiliation-related questions and feedback statements on website usability, whereas the first-league clubs' questionnaire encompassed inquiries concerning the significance of website features and an evaluation of both the official and developed websites.

Section 5.2 (p. 54) delves into the comprehensive pilot testing conducted to ensure clarity and participant comprehension regarding each questionnaire, specifically focusing on identifying and addressing any potential areas of ambiguity. The refinement of the questionnaires based on participant feedback is outlined while emphasising the exclusion of pilot test responses from the analysis to maintain data validity.

Section 5.3 (p. 55) elucidates the results obtained through the questionnaires. The low-ranking clubs' study outcomes are delineated in Subsection 5.3.1 (p. 55), demonstrating positive evaluations where participants perceived the developed websites as more user-friendly and up-to-date than the official websites. The SUS Template employment showcased excellent usability, substantiated by the obtained mean score. Subsection 5.3.2 (p. 62) discusses the first-league clubs' questionnaire analysis, revealing the prominence of information concerning youth teams and non-football sports on club websites. Participants conferred higher ratings on the developed website, which exhibited diminished variability when contrasted with the official website, indicating favourable perceptions and preferences.

Section 5.4 (p. 69) comprises a comprehensive discussion highlighting the findings, substantiating the feasibility of automatically generating websites utilising publicly available data. The developed websites were ascertained to possess enhanced up-to-dateness and user-friendliness, particularly among first-league clubs. Statistical evidence corroborated the superior usability of

the developed websites, requiring minimal effort for a generation while permitting extensive customisation.

Despite potential threats to the study's validity, as elucidated in Section 5.5 (p. 71), diligent measures were implemented to mitigate these concerns, ensuring the scientific rigour and validity of the study. Consequently, the hypothesis retains its validity.

Chapter 6

Conclusions

6.1 Summary

High-quality sports websites are paramount in sports marketing and fan engagement, necessitating a delicate balance of aesthetics, usability, and credibility. Evaluating the quality of such websites places a high priority on fulfilment, directly impacting consumer preference and loyalty. It is crucial to design websites that effectively combine these elements while considering factors such as gender differences and hedonic/utilitarian concerns. Incorporating persuasive elements and interactive features is also essential for success in engaging with fans. However, it is important to note that web scraping, a practical technique with various applications, including journalism, requires careful consideration of legal and ethical implications.

One significant challenge in sports club websites is the digital divide between larger, higher-ranking and smaller, lower-ranking clubs. The former tend to have more robust websites with more significant resources, while the latter often struggle with outdated platforms and a lack of comprehensive information. Bridging this divide is a crucial objective, and in this Dissertation, a platform is proposed to address this issue by leveraging publicly accessible data through the *zerozero* API, which serves as a crowdsourcing platform providing a wealth of information. The goal is to automatically generate up-to-date and user-friendly websites for local sports teams, utilising the API to feed the websites with relevant and comprehensive content, including news, game schedules, and player statistics. Emphasis is placed on enhancing usability and customisation to reflect each team's unique brand identity, thereby providing fans with a more engaging and personalised browsing experience.

The club websites were developed using the Blade PHP template engine, Bootstrap 5.3.0, CSS, and JavaScript for front-end enhancement. At the same time, PHP and Apache HTTP Server were utilised for the gateway, facilitating communication with the *zerozero* API and the website generator API. The deployed architecture involved three devices, including the generator, club website device, and *zerozero* API device, enabling effective communication and data processing. The in-house website generator tool provided a user-friendly interface for website creation, cache clearance, and personalised content delivery. The solution successfully leveraged *zerozero* endpoints to retrieve and present comprehensive club information, while the front-end design incorporated intelligently designed pages and components to enhance user engagement and provide easy access to relevant content.

In order to assess the perception of up-to-dateness and gather feedback on the visual appeal and usability of the developed websites, two questionnaires were employed, targeting supporters of low-ranking and first-league football clubs, with questions covering demographic information, football knowledge, usability assessment, and feedback on website features. Comprehensive pilot testing was conducted to refine the questionnaires based on participant feedback. The results indicated positive evaluations of the developed websites, with participants perceiving them as more user-friendly and up-to-date than official websites. The SUS scores confirmed excellent usability. The discussion highlighted the feasibility of automatically generating websites using publicly available data supported by statistical evidence and measures to ensure study validity. Overall, the findings validated the hypothesis and demonstrated the improved usability and customisation of the developed websites.

6.2 Hypothesis Revisited

Based on the findings presented in Chapter 2 (p. 5), we have identified the primary issues existing in the digital presence of low-ranking sports clubs that require resolution, leading us to formulate the following hypothesis:

> *"It is feasible to develop a platform that utilises publicly accessible data to automatically generate up-to-date and user-friendly websites for local sports teams, providing them with an efficient and cost-effective solution for creating and maintaining websites that remain current with the latest news and games while ensuring good usability and customisation for brand identity."*

The project's implementation phase empirically validates the hypothesis, providing substantial evidence supporting the feasibility of constructing an automated website generation platform catered specifically for local sports teams. Through meticulous selection and integration of various technologies, the implementation phase yields empirical substantiation, affirming the practicality and achievability of the proposed solution. The systematic and deliberate approach to technology selection enhances the project's scientific merit, bolstering confidence in the effectiveness and reliability of the automated website generation platform.

Furthermore, the project's emphasis on usability and customisation during the implementation phase aligns with contemporary research in user-centred design, underscoring the commitment to creating user-friendly websites. The project acknowledges the significance of providing an engaging and tailored user experience by prioritising these aspects. The emphasis on usability and customisation further supports the project's scientific validity and practical relevance, demonstrating its alignment with user expectations and preferences.

To support the scientific validity supported by the implementation, we followed our work with an empirical evaluation to assess the perception of up-to-dateness, visual appeal, and usability of the developed websites. This evaluation employed two questionnaires to gather data from low-ranking and first-league football club supporters. These questionnaires encompassed demographic information, football knowledge, and inquiries regarding usability assessment. Additionally, a pilot testing phase was conducted to ensure the clarity of the questionnaires, and participant feedback was used to refine them. The evaluation results demonstrated positive evaluations of the developed websites among participants from low-ranking clubs, accompanied by excellent usability scores. Moreover, participants from first-league clubs exhibited favourable perceptions and preferences towards the developed websites.

Consequently, the research questions were revised to validate various aspects of the developed websites. RQ1 focused on validating the level of up-to-dateness achieved through the automatic generation of websites using publicly available data. RQ2 aimed to assess the developed websites' usability level compared to their official counterparts. RQ3 aimed to investigate whether the website generation process required minimal effort. Lastly, RQ4 aimed to evaluate the level of customisation provided by the developed websites.

The research questions concluded that the developed websites, generated using publicly available data, exhibit significant advantages over the official websites regarding up-to-dateness and user-friendliness. We also concluded that the website generation process required minimal effort and provided extensive customisation options by successfully designing multiple customisable skins.

6.3 Main Contributions

The main contributions of this work can be summarised as follows:

State-of-the-art: Our contributions in this area include:

- Exploration of the importance of high-quality sports websites and the key elements contributing to their effectiveness.
- Discussion on using external data in web development, specifically on web scraping and its legal and ethical considerations.
- Introduction of web presence models, such as the web presence pyramid, for categorising and evaluating the digital presence of small businesses and local sports teams.

- Emphasis on the need for a balanced approach to website design, considering aesthetics, usability, and credibility.
- Analysis of the digital divide between lower-ranking and well-established clubs highlights the challenges and opportunities for bridging this gap.

Portugal Clubs' Website Analysis: Our contributions in this area include:

- Comprehensive analysis of the website data collected by Mascarenhas from football clubs in Portugal, covering teams across multiple rankings.
- Examination and comparison of disparities among rankings in various aspects of the clubs' websites, including design, content, functionality, and user experience.
- Statistical analysis of the data to identify significant differences and trends based on club ranking.
- Insightful interpretation of the findings, shedding light on the implications of ranking variations and their impact on the digital presence of Portuguese football clubs.

Development of an Automated Website Generation Platform: Our work introduces a platform that utilises publicly accessible data to generate websites for local sports teams automatically. This platform offers an efficient and cost-effective solution for creating and maintaining up-to-date websites with the latest news and games while allowing for customisation.

Conducted Experiment: The work involved questionnaires with participants who evaluated the developed websites and compared them with the official websites of the clubs they support. The results obtained from these questionnaires provide valuable insights that can be analysed and applied in other contexts.

These contributions collectively advance the knowledge and understanding of sports website analysis, provide practical solutions for website generation, and contribute to improving digital presence in the sports industry.

6.4 Future Work

In the realm of website design and management, future work should focus on the following aspects:

Enhanced Customisation for Club Websites The current framework provides limited website customisation options, offering only two pre-defined skins that alter the visual appearance of the website components. To foster a sense of brand identity and individuality, future efforts should aim to collaborate with a proficient design team to create a broader range of visually appealing skins and components. Furthermore, advanced customisation features

should be implemented, allowing website managers to personalise their websites by selecting specific components for each page and customising fonts, colours, and other design elements. However, careful consideration must be given to balancing customisation flexibility and maintaining high standards of aesthetics and usability.

Website Management Dashboard The existing framework enables creating websites for each club with customisable domain names and skin selection. However, future work should focus on developing a comprehensive dashboard that empowers users to manage their websites effectively. This dashboard would offer extensive customisation capabilities, allowing users to modify various aspects of their websites, including layout, content, and interactive features. Additionally, it should facilitate the seamless integration of user-generated news articles, empowering clubs to provide timely and relevant information to their supporters.

Incorporating Revenue-Generating Pages The existing framework primarily focuses on static pages, thus limiting the revenue-generation potential for clubs. To address this issue and assist smaller clubs in their financial sustainability, future work should explore the inclusion of dynamic pages such as online stores and ticketing platforms. By providing opportunities for fans to make purchases and contribute financially, clubs can overcome monetary challenges. These revenue-generating pages can be seamlessly integrated into the website management dashboard, allowing clubs to manage and optimise these financial channels effectively.

On the other hand, in the context of validation, future work should aim to expand the following aspects:

Testing with a Larger Sample of Supporters from Smaller Clubs The empirical evaluation was limited to a specific sample size due to confidentiality constraints. To strengthen the validity and generalizability of the findings, future research should prioritise expanding the participant pool by including a larger sample of supporters from diverse smaller clubs. This broader representation would yield more comprehensive insights and enhance the study's external validity.

Evaluating the Representation of Club Identity Further investigation should be conducted to assess the website framework's efficacy in accurately representing each club's unique identity. Particular attention should be given to clubs with secondary colours, as the current framework is designed to accommodate a single primary colour. By conducting systematic analyses and gathering user feedback, the research can shed light on the extent to which the framework successfully captures and reflects the distinct brand identity of each club.

Moreover, it is worth acknowledging that there may be additional topics warranting improvement that were not addressed within this Dissertation, thus presenting avenues for future research endeavours.

References

[1] Definition of Aesthethics by Interaction Design Foundation. https://www.interaction-design.org/literature/topics/aesthetics. Accessed: 2023-2-3.

[2] Prosebot :: zerozero.pt. https://www.zerozero.pt/prosebot.php. Accessed: 2023-2-3.

[3] Twig. https://twig.symfony.com/. Accessed: 2023-06-22.

[4] zerozero.pt :: Porque todos os jogos começam assim. https://www.zerozero.pt/. Accessed: 2022-11-21.

[5] Sinan Si Alhir. *Learning Uml*. " O'Reilly Media, Inc.", 2003.

[6] Farah Alsudani and Matthew Casey. The effect of aesthetics on web credibility. *People and Computers XXIII Celebrating People and Technology*, pages 512–519, 2009.

[7] Aaron Bangor, Philip Kortum, and James Miller. Determining what individual sus scores mean: Adding an adjective rating scale. *Journal of usability studies*, 4(3):114–123, 2009.

[8] John Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

[9] Stephen Burgess. Representing small business web presence content: the web presence pyramid model. *European Journal of Information Systems*, 25(2):110–130, 2016.

[10] Stephen Burgess, Craig Parker, and Scott Bingley. The digital divide in local sporting clubs. *Information Matters*, 1(9), 2021.

[11] Stephen Burgess, Craig M. Parker, and Scott Bingley. Mapping the online presence of small local sporting clubs. *Journal of the Association for Information Science and Technology*, 72(4):433–448, 2021.

[12] Weisheng Chiu and Doyeon Won. Relationship between sport website quality and consumption intentions. *Psychological Reports*, 118(1):90–106, 2016. OA status: green_accepted.

[13] Daniela Soares Cruzes and Lotfi ben Othmane. Threats to validity in empirical software security research. *Empirical Research for Software Security: Foundations and Experience*, 2017.

[14] Marc Delisle. *Mastering phpMyAdmin 3.1 for effective MySQL management*. Packt Publishing, 2009.

[15] Rabiyatou Diouf, Edouard Ngor Sarr, Ousmane Sall, Babiga Birregah, Mamadou Bousso, and Seny Ndiaye Mbaye. Web scraping: State-of-the-art and areas of application. In *2019 IEEE International Conference on Big Data (Big Data)*. IEEE.

[16] Samuel Keith Duncan. The business of sport, sledging and the corruption of play – an interpretation through a huizingian-bourdieu lens. *Physical Culture and Sport. Studies and Research*, 80(1):15–26, 2018.

[17] Muge Elden, Sinem Yeygel Cakir, and Ugur Bakir. Website persuasiveness in sports marketing. In *Proceedings of the 2018 International Conference on Computers in Management and Business - ICCMB '18*. ACM Press.

[18] Robert Feldt and Ana Magazinius. Validity threats in empirical software engineering research-an initial survey. In *Seke*, pages 374–379, 2010.

[19] Julie Fisher, Annemieke Craig, and John Bentley. Moving from a web presence to e-commerce: The importance of a business—web strategy for small-business owners. *Electronic Markets*, 17(4):253–262, 2007.

[20] Assistant Secretary for Public Affairs. System usability scale (sus), Sep 2013.

[21] Samuel Greengard. Following the crowd. *Commun. ACM*, 54(2):20–22, feb 2011.

[22] Youngjin Hur, Yong Jae Ko, and Cathryn L. Claussen. Acceptance of sports websites: a conceptual model. *International Journal of Sports Marketing and Sponsorship*, 12(3):13–27, 2011.

[23] Youngjin Hur, Yong Jae Ko, and Joseph Valacich. A structural model of the relationships between sport website quality, e-satisfaction, and e-loyalty. *Journal of sport management*, 25(5):458–473, 2011.

[24] IBM. Deployment diagrams.

[25] IBM. Power analysis of one-sample spearman correlation test.

[26] IBM. Wilcoxon signed-rank test.

[27] Ankur Joshi, Saket Kale, Satish Chandel, and D. K. Pal. Likert scale: Explored and explained. *Current Journal of Applied Science and Technology*, 7(4):396–403, Feb. 2015.

[28] Simon Kemp. Digital 2023: Global overview report — datareportal – global digital insights, Feb 2023.

[29] Simon Kemp. Digital 2023: Portugal — datareportal – global digital insights, Feb 2023.

[30] Andrew P. King and Robert J. Eckersley. Chapter 7 - inferential statistics iv: Choosing a hypothesis test. In Andrew P. King and Robert J. Eckersley, editors, *Statistics for Biomedical Engineers and Scientists*, pages 147–171. Academic Press, 2019.

[31] Brandy Klug. An overview of the system usability scale in library website and system usability testing. *Weave: Journal of Library User Experience*, 1(6), 2017.

[32] Dirk P Kroese, Tim Brereton, Thomas Taimre, and Zdravko I Botev. Why the monte carlo method is so important today. *Wiley Interdisciplinary Reviews: Computational Statistics*, 6(6):386–392, 2014.

[33] Vlad Krotov, Leigh Johnson, and Leiser Silva. Tutorial: Legality and ethics of web scraping. 2020.

[34] Miguel Mascarenhas. Internship report. To Be Published.

[35] Satish Nambisan, Mike Wright, and Maryann Feldman. The digital transformation of innovation and entrepreneurship: Progress, challenges and key themes. *Research Policy*, 48(8):103773, 2019. The Digital Transformation of Innovation and Entrepreneurship.

[36] Norman J O'Reilly and John P Nadeau. Revenue generation in professional sport: A diagnostic analysis. *International Journal of Sport Management and Marketing*, 1(4):311–330, 2006.

[37] Kiemute Oyibo and Julita Vassileva. The interplay of aesthetics, usability and credibility in mobile website design and the effect of gender. *Journal on Interactive Systems*, 8(2):1, 2017.

[38] Tereza Semerádová, Petr Weinlich, Tereza Semerádová, and Petr Weinlich. Looking for the definition of website quality. *Website quality and shopping behavior: Quantitative and qualitative evidence*, pages 5–27, 2020.

[39] Apache HTTP Server. Apache http server tutorial: .htaccess files.

[40] Matt Stauffer. *Laravel: Up & running: A framework for building modern php apps*. O'Reilly Media, 2019.

[41] Lisbeth Thorlacius. The role of aesthetics in web design. *Nordicom Review*, 28(1), 2007.

[42] Luke Welling and Laura Thomson. *PHP and MySQL Web development*. Sams publishing, 2003.

[43] Wikipedia contributors. Monte carlo method — Wikipedia, the free encyclopedia, 2023. [Online; accessed 16-June-2023].

[44] Hyrum K Wright, Miryung Kim, and Dewayne E Perry. Validity concerns in software engineering research. In *Proceedings of the FSE/SDP workshop on Future of software engineering research*, pages 411–414, 2010.

Appendix A

`.htaccess` of Club Websites

The appendix on `.htaccess` file configuration is essential for routing club websites, ensuring seamless navigation and effective content delivery. It demonstrates how the file governs URL rewriting, error handling, and other aspects of the web server's behaviour. As a valuable reference, the appendix helps understand the technical configuration involved in the automated website generation process. It showcases the specific directives and rules that enable proper website functionality, search engine optimisation, and enhanced user experience.

```
#Rewrite secure
RewriteEngine On

#Errorpage
ErrorDocument 404 /notfound.php

#Main rules
RewriteRule ^$ homepage.php [QSA,L]
RewriteRule ^sport/([0-9]*)$ /sport.php?&sport=$1 [R=301,QSA,L]
RewriteRule ^competition/([0-9]*)/([0-9]*)$ /competition.php?&sport=
    ↪ $1&competition=$2 [R=301,QSA,L]
RewriteRule ^team/([0-9]*)/([0-9]*)$ /team.php?&sport=$1&teamID=$2 [
    ↪ R=301,QSA,L]
RewriteRule ^game/([0-9]*)$ /game.php?&game=$1 [R=301,QSA,L]
RewriteRule ^news/([0-9]*)$ /news.php?news=$1&zerozero=1 [R=301,QSA,
    ↪ L]
RewriteRule ^calendar$ /calendar.php [R=301,QSA,L]
RewriteRule ^contacts$ /contacts.php [R=301,QSA,L]
```

Appendix B

Website Implemented Pages

In this appendix, we present the implemented pages of the club website for both Skin 1 and Skin 2 of *Vila Futebol Clube*[1]. We will showcase the following pages to provide a comprehensive overview of the website's structure and functionality:

- Homepage: The homepage serves as the central hub of the club website, featuring highlights, the latest news, upcoming matches, and other essential information.
- Sport Page: This page focuses on a specific sport, providing detailed information about the sport, teams, players, and related news.
- Team Page: The team page offers in-depth insights into a particular club team, including player profiles, match schedules, and team statistics.
- News Article Page: This page showcases individual news articles, allowing users to access detailed articles, updates, and announcements about the club and its activities.
- Competition Page: The competition page provides information about ongoing tournaments, leagues, or competitions in which the club participates. It includes fixtures, results, standings, and other relevant details.
- Game Page: This page displays comprehensive information about specific games or matches, including match details, lineups, live updates, and post-match analysis.
- Calendar Page: The calendar page offers an organised view of the club's schedule, including upcoming matches, events, and important dates.
- Contacts Page: The contacts page provides contact information for the club, allowing users to access the club's personnel.

By showcasing these implemented pages, we aim to comprehensively demonstrate the club websites' structure and functionality, highlighting the features and information available to users.

[1] https://vilafc.00.pt/

Figure B.1: Homepage for Skin 1

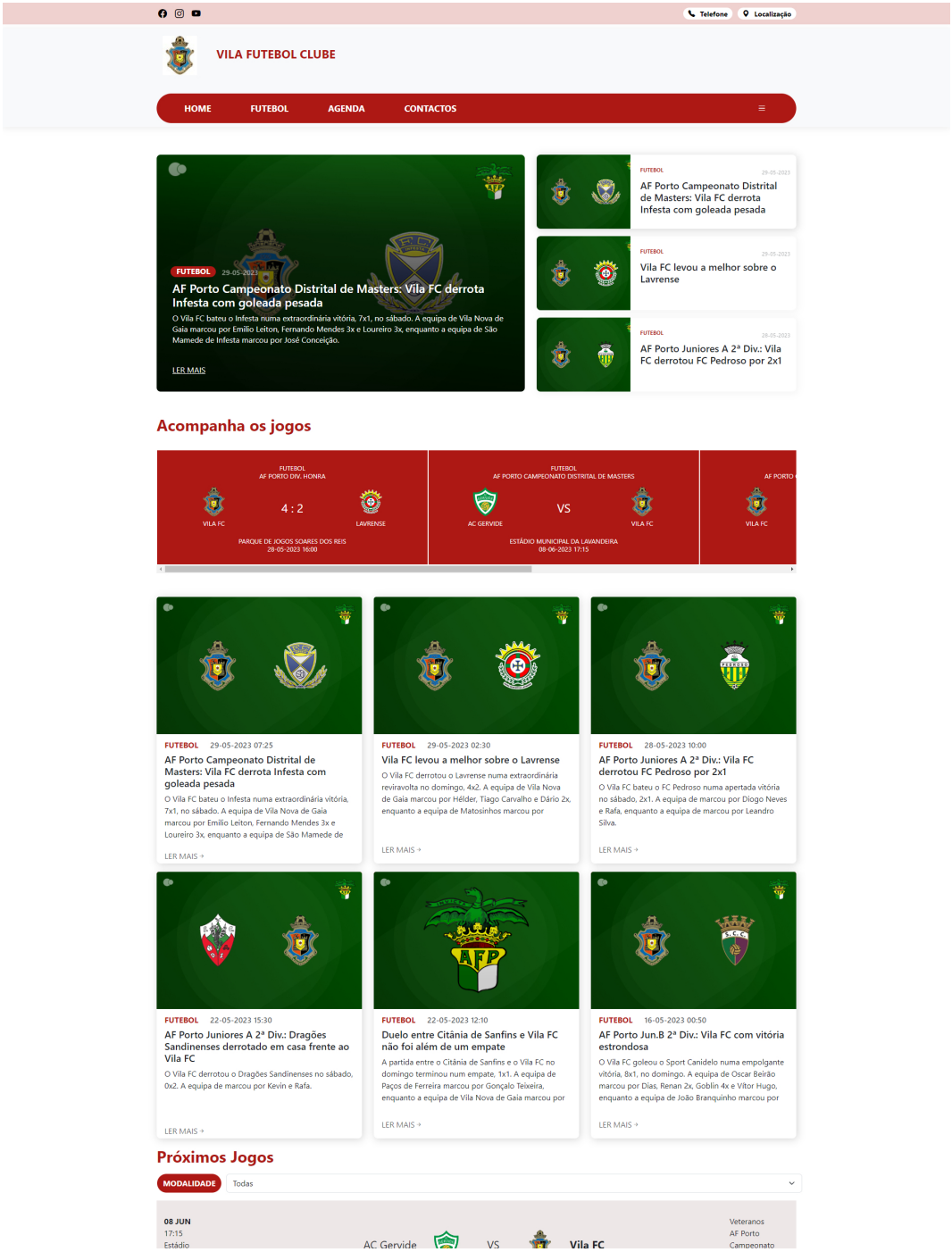


Figure B.2: Homepage for Skin 2

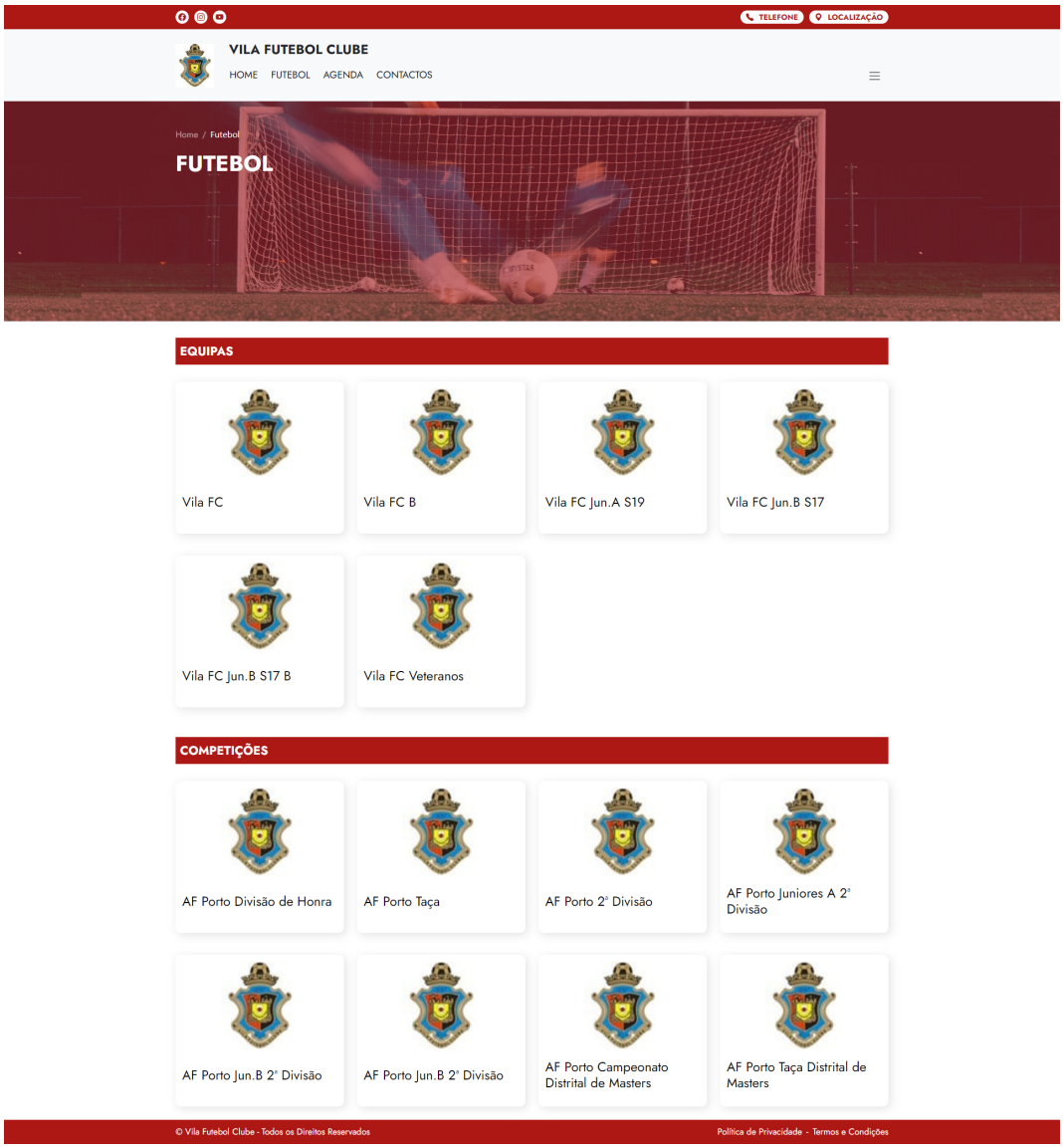


Figure B.3: Sport Page for Skin 1

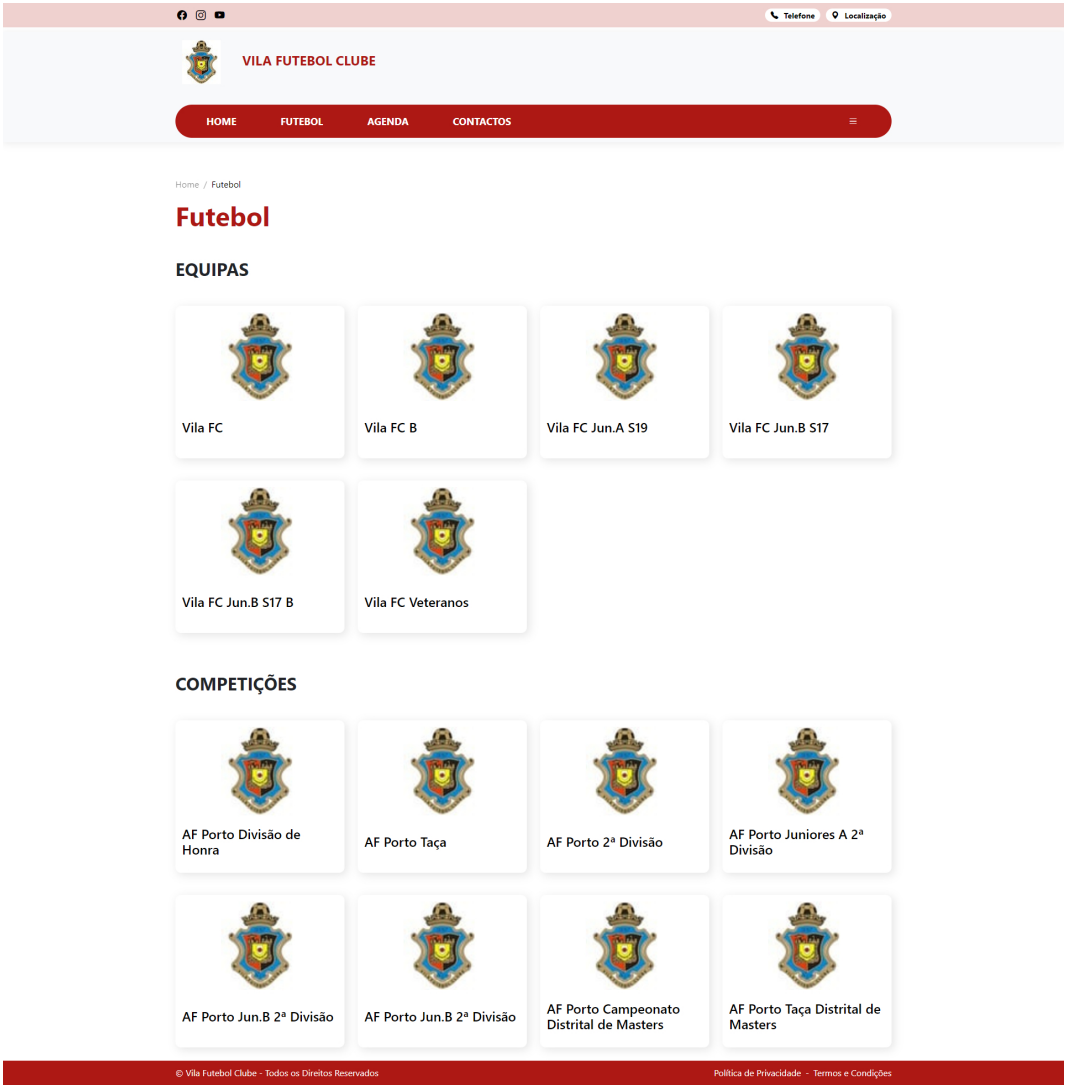


Figure B.4: Sport Page for Skin 2

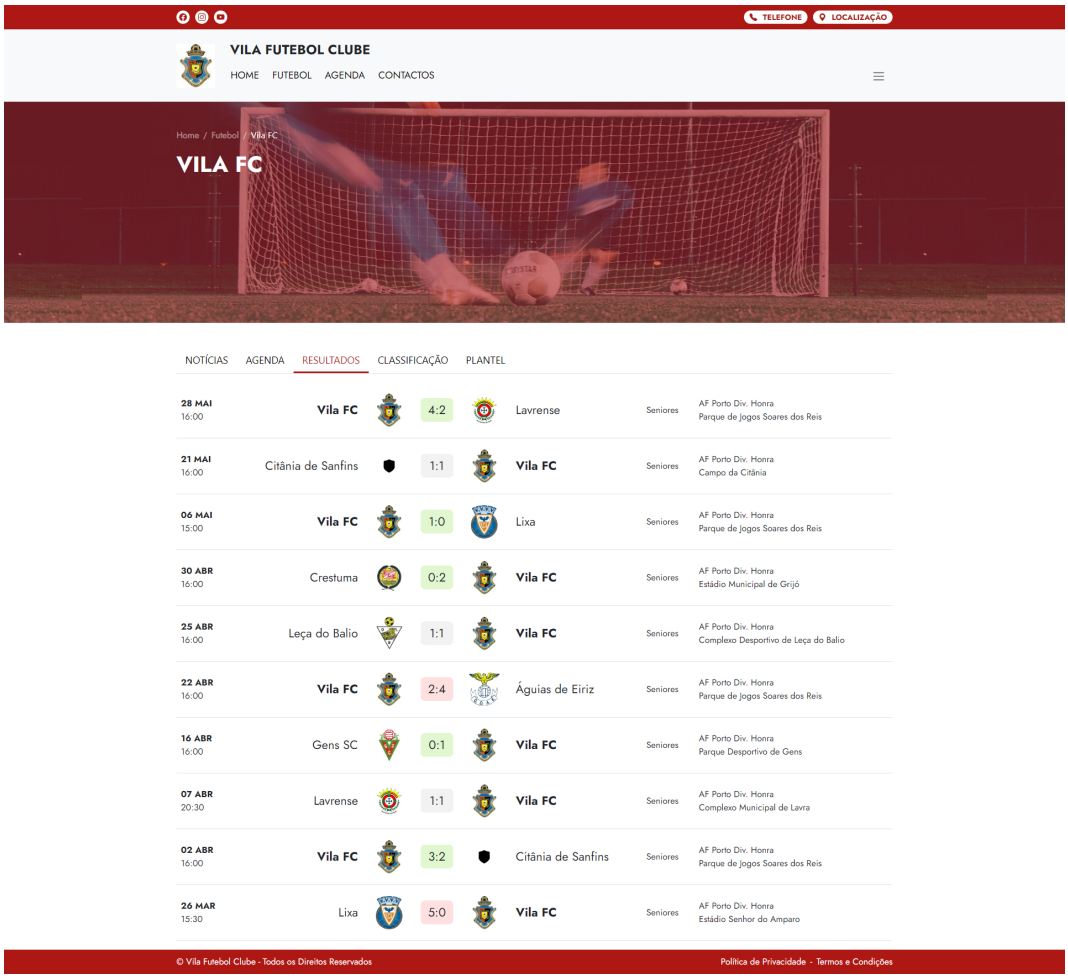


Figure B.5: Team Page for Skin 1 (in this case, showing the results of the team)

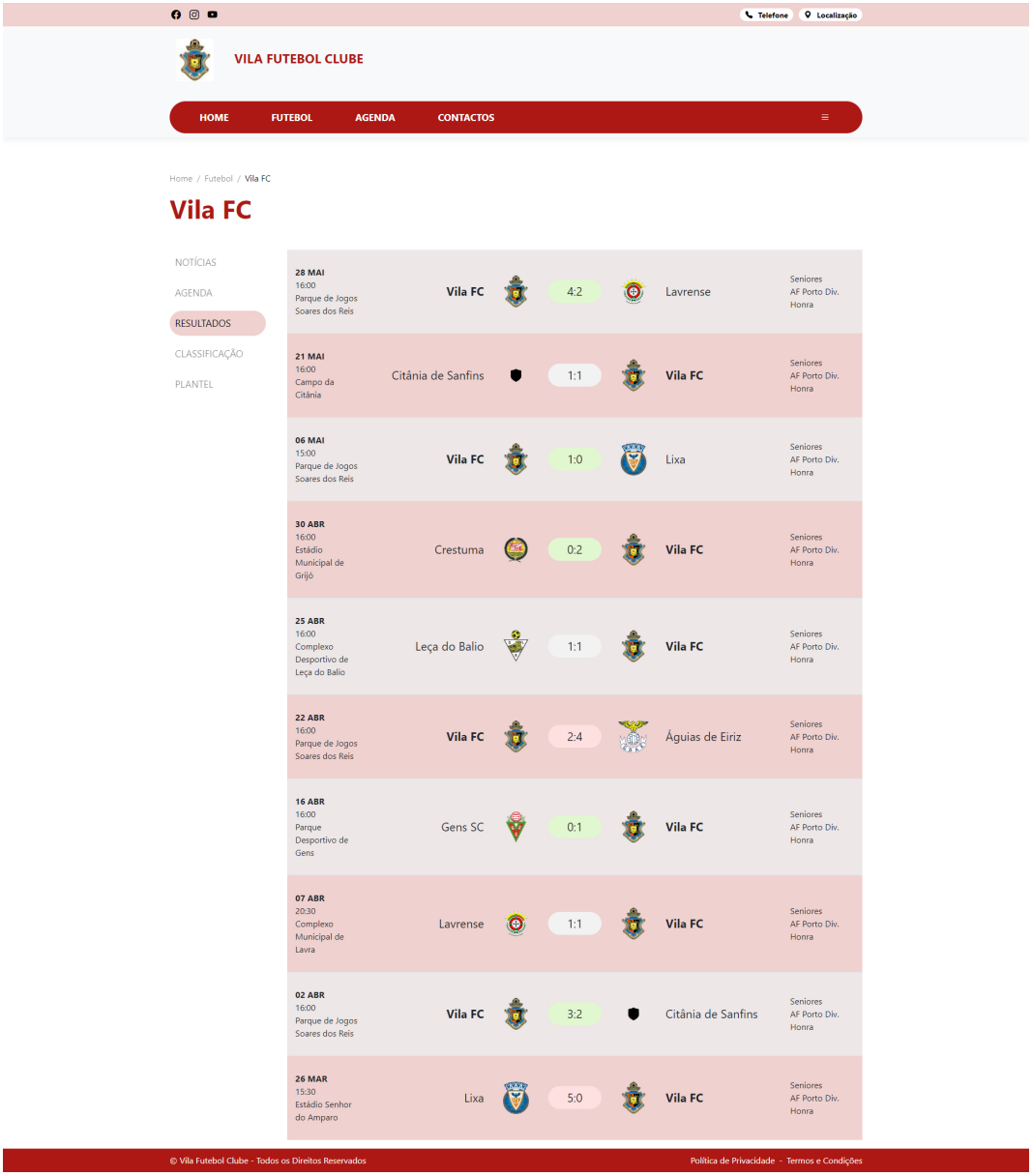


Figure B.6: Team Page for Skin 2 (in this case, showing the results of the team)

Figure B.7: News Article Page for Skin 1

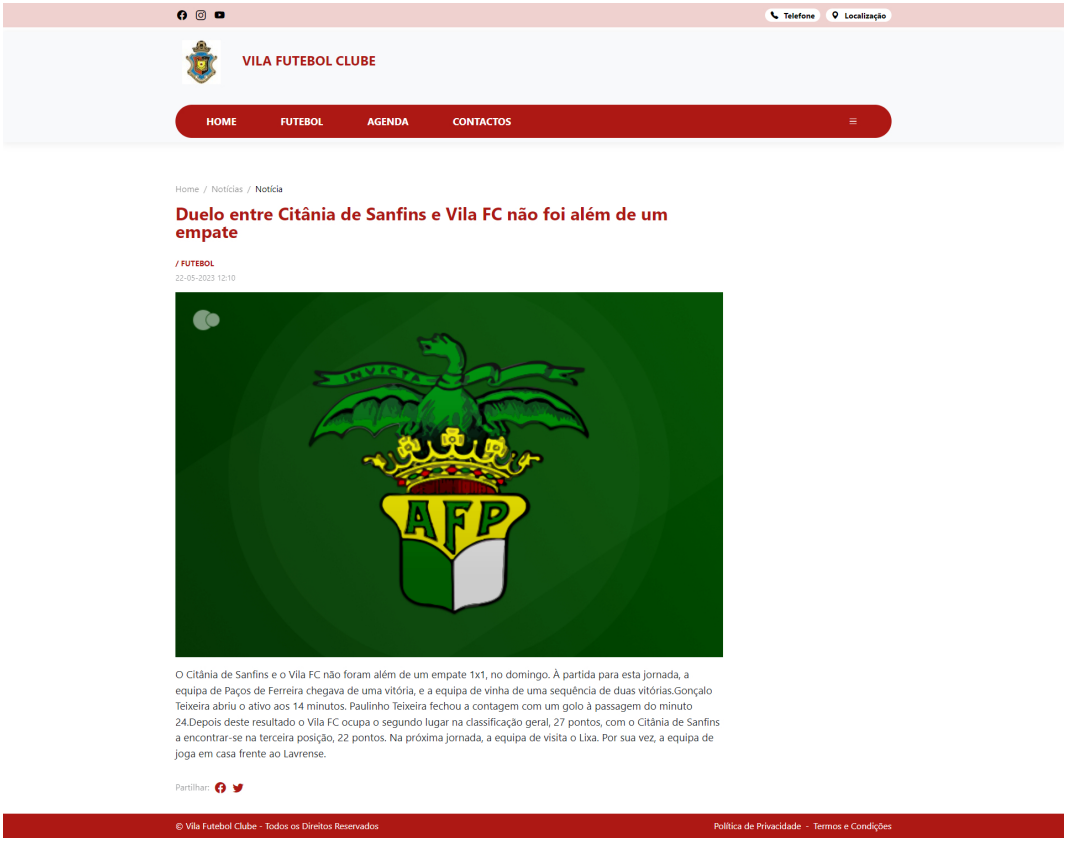


Figure B.8: News Article Page for Skin 2

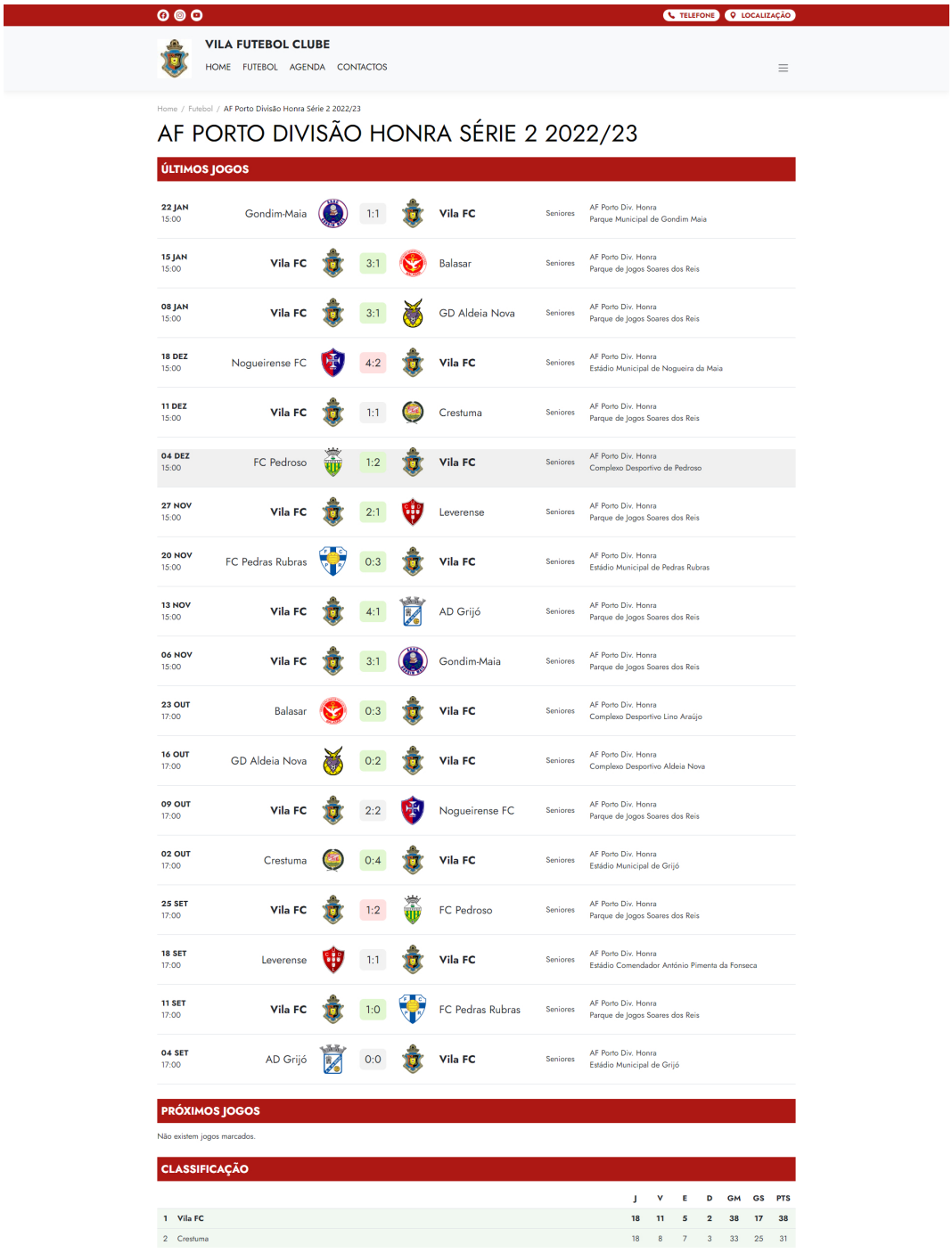


Figure B.9: Competition Page for Skin 1

Figure B.10: Competition Page for Skin 2

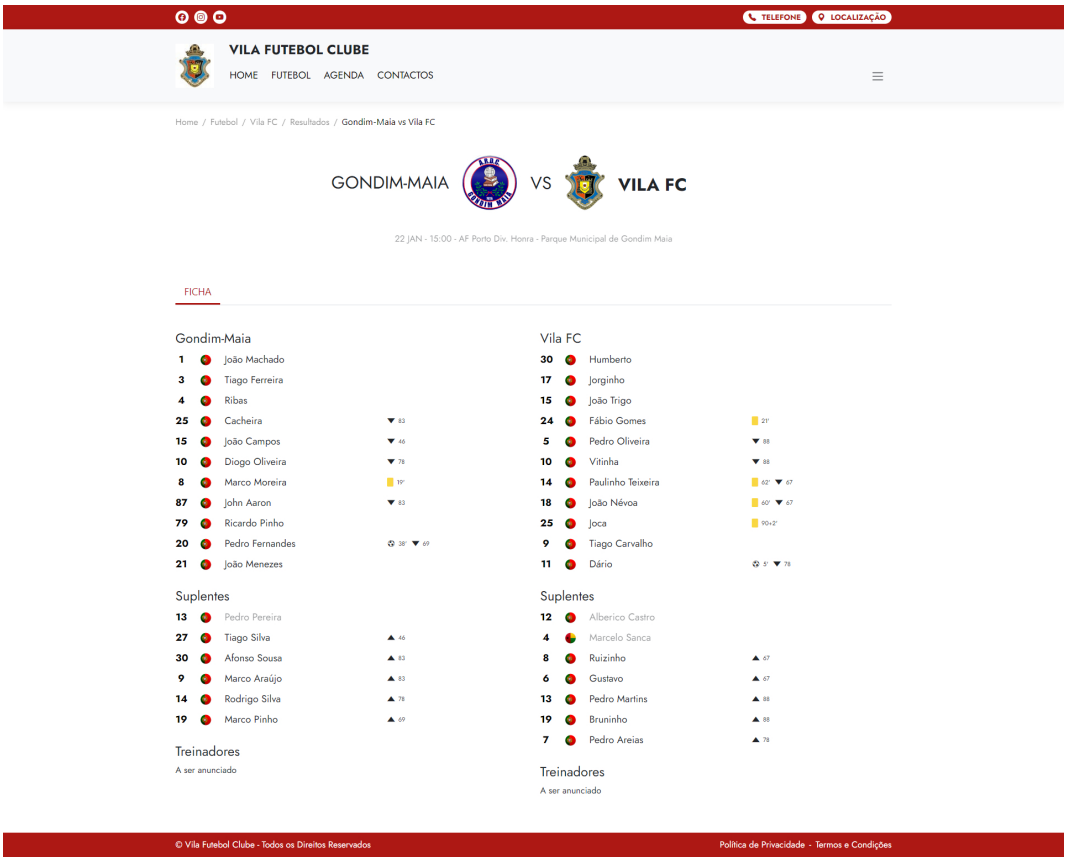


Figure B.11: Game Page for Skin 1

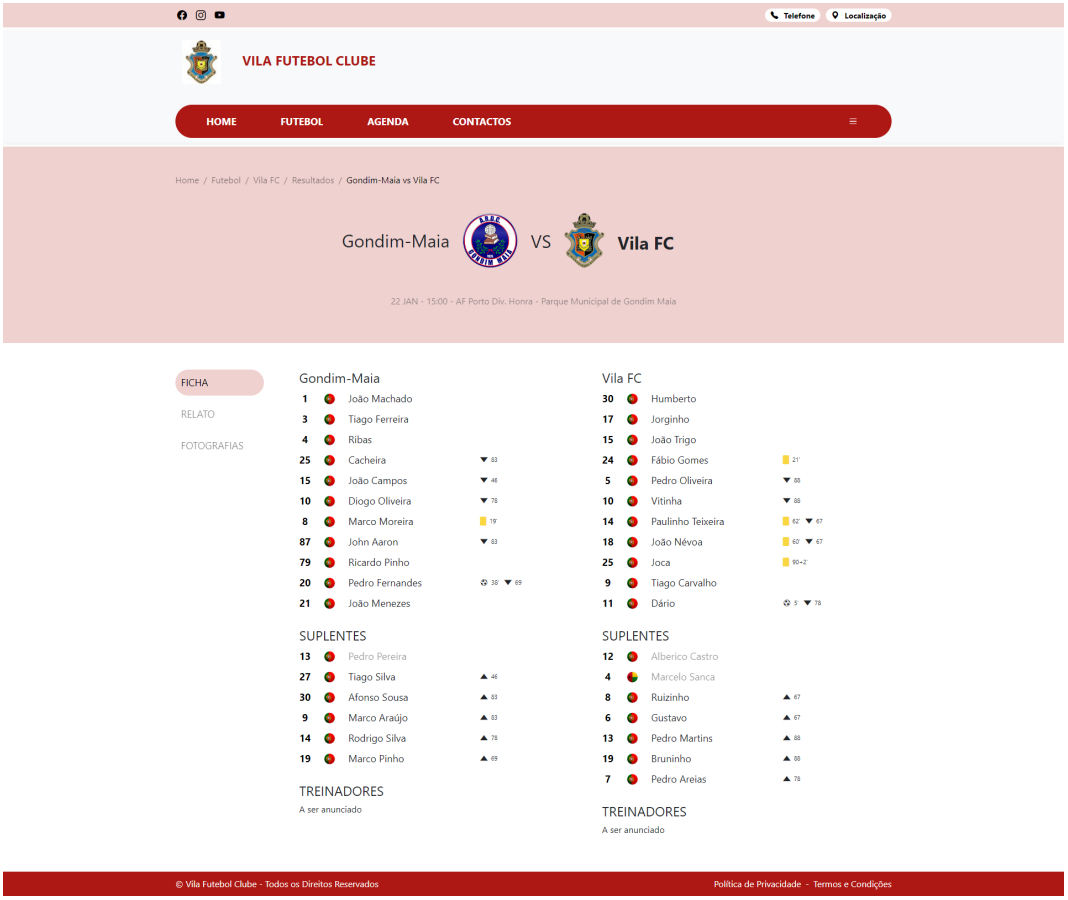


Figure B.12: Game Page for Skin 2

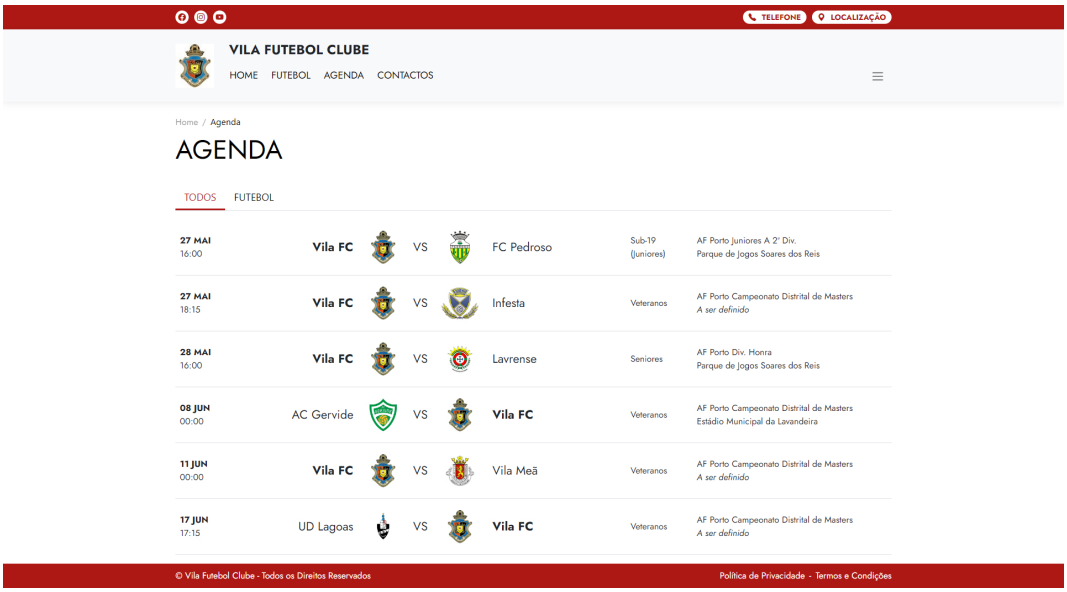


Figure B.13: Calendar Page for Skin 1

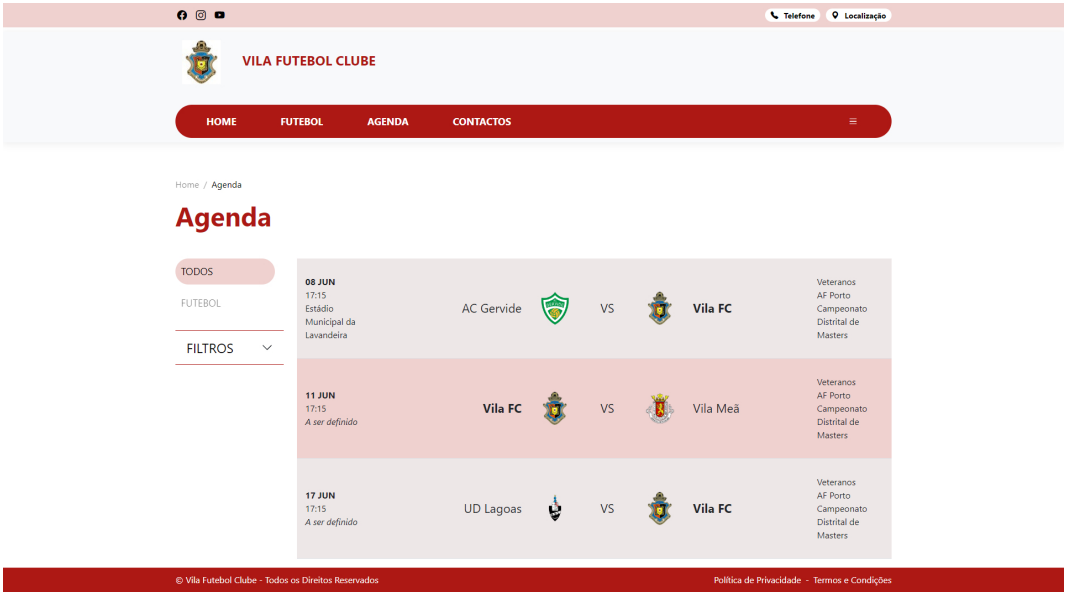


Figure B.14: Calendar Page for Skin 2

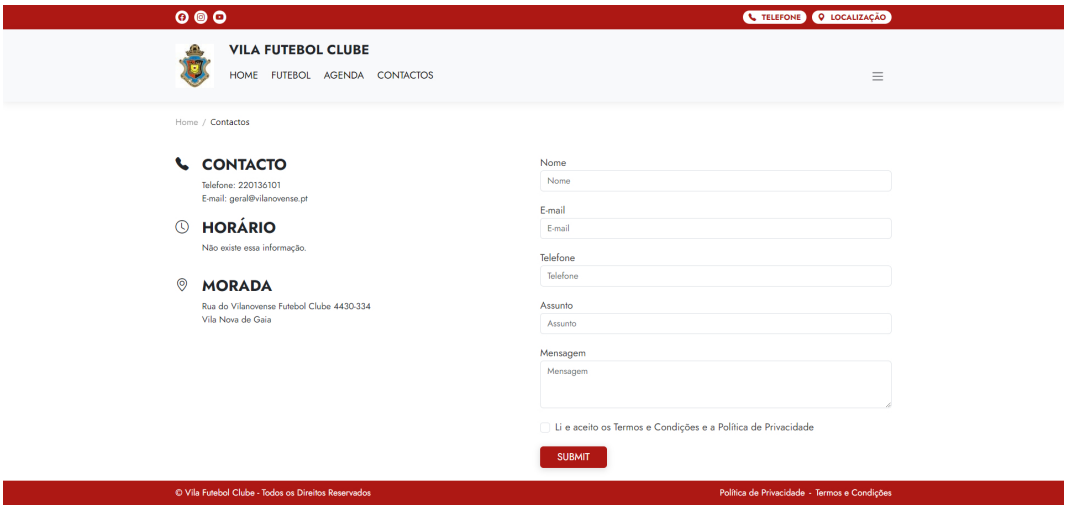


Figure B.15: Contacts Page for Skin 1

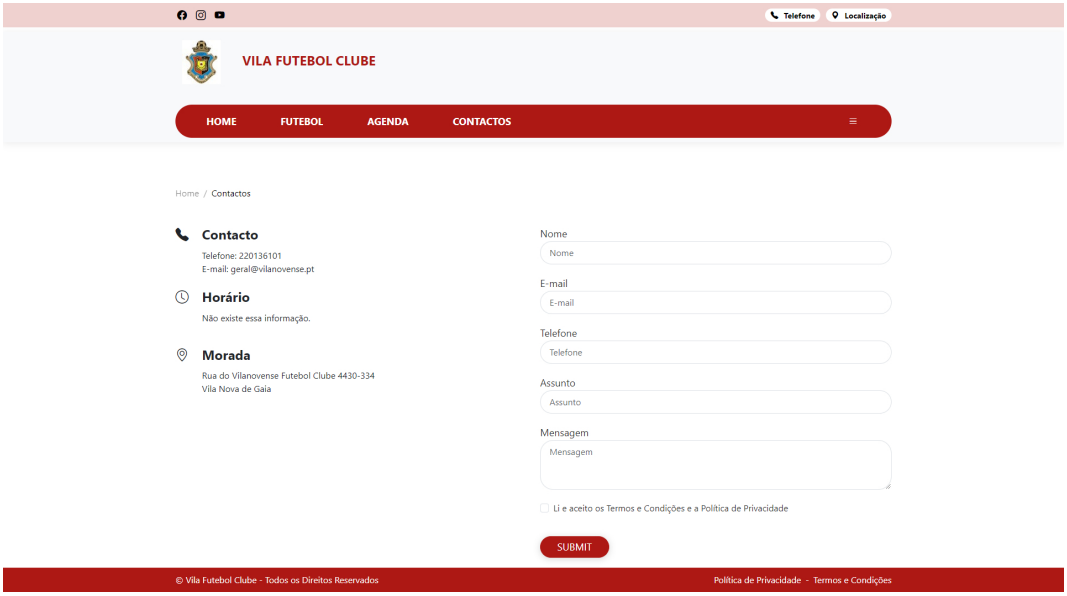


Figure B.16: Contacts Page for Skin 2

Appendix C

Questionnaires

In this appendix, we present the questionnaires used in the research study to gather valuable insights and participant feedback.

The questionnaires covered user-friendliness, up-to-dateness of the generated websites, visual appeal, ease of navigation, and overall satisfaction. Participants were also allowed to provide additional comments and suggestions to enrich the research findings further.

C.1 Questionnaire 1: Low-Ranking Football Clubs

In this section of the questionnaires appendix, we present the questionnaire designed explicitly for supporters of low-ranking football clubs. The questionnaire aims to gather valuable insights and feedback from this group of participants, providing a unique perspective on the automated website generation platform's impact on their experience and satisfaction.

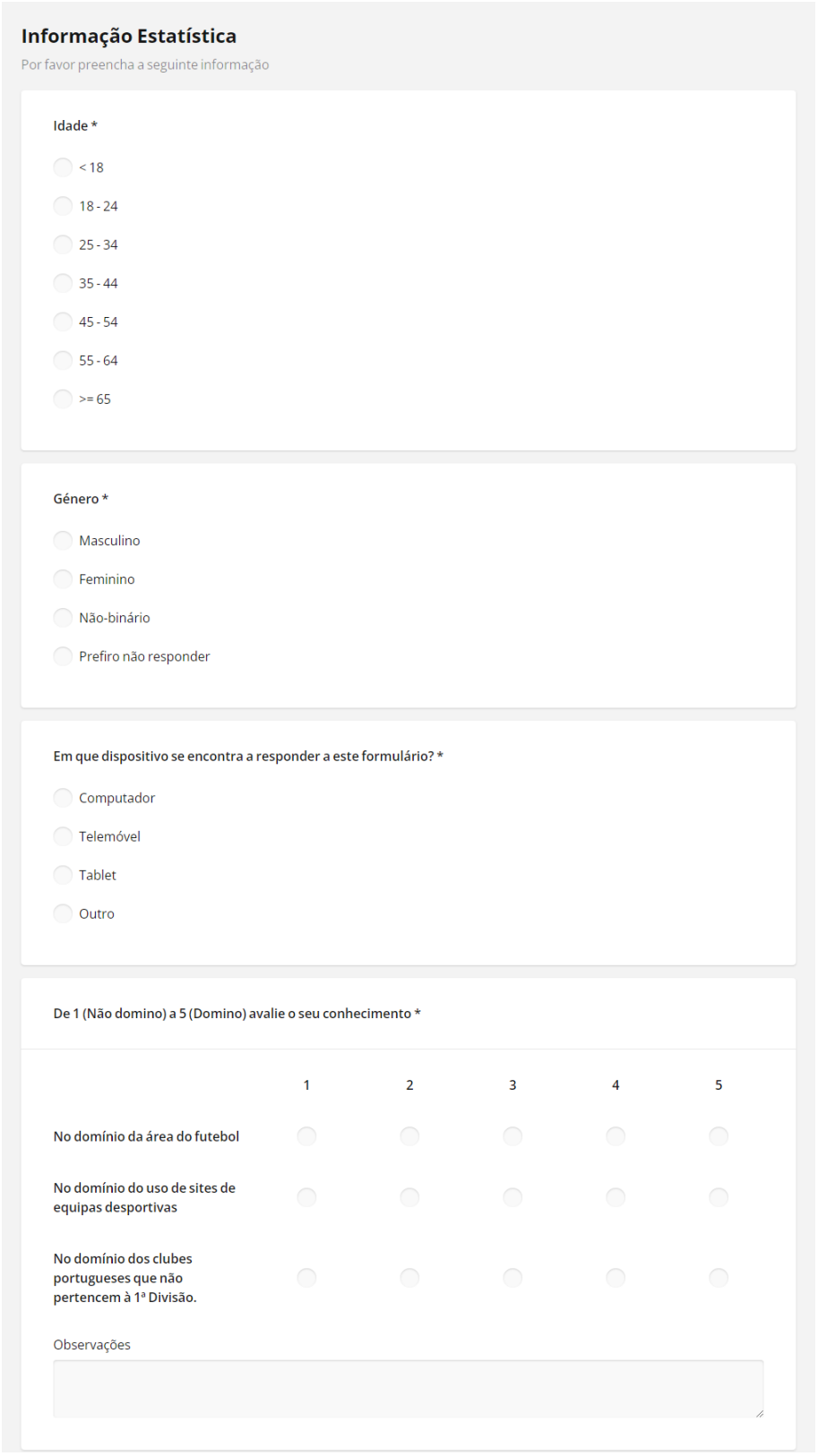
By targeting supporters of low-ranking clubs, we aim to uncover valuable insights into the platform's ability to bridge the digital divide, providing equal opportunities and enhancing communication for clubs of varying sizes and rankings. The questionnaire enables participants to share their perceptions, opinions, and suggestions, contributing to a comprehensive understanding of the platform's effectiveness and potential areas for improvement.

C.2 Questionnaire 2: First-League Football Clubs

In this section of the questionnaires appendix, we focus on supporters of first-league football clubs and present a questionnaire specifically designed for this group. This questionnaire aims to gather valuable insights and feedback from these participants, offering a unique perspective on how the automated website generation platform has influenced their experience and satisfaction.

By targeting supporters of first-league clubs, we seek to gain in-depth insights into how the platform performs in catering to the requirements of clubs with higher visibility and fan bases. The

questionnaire allows participants to share their perceptions, opinions, and suggestions, contributing to a comprehensive understanding of the platform's effectiveness and identifying potential areas for further improvement.



Informação Estatística

Por favor preencha a seguinte informação

Idade *

- < 18
- 18 - 24
- 25 - 34
- 35 - 44
- 45 - 54
- 55 - 64
- >= 65

Género *

- Masculino
- Feminino
- Não-binário
- Prefiro não responder

Em que dispositivo se encontra a responder a este formulário? *

- Computador
- Telemóvel
- Tablet
- Outro

De 1 (Não domino) a 5 (Domino) avalie o seu conhecimento *

	1	2	3	4	5
No domínio da área do futebol					
No domínio do uso de sites de equipas desportivas					
No domínio dos clubes portugueses que não pertencem à 1ª Divisão.					

Observações

Figure C.1: First page of the Questionnaire 1

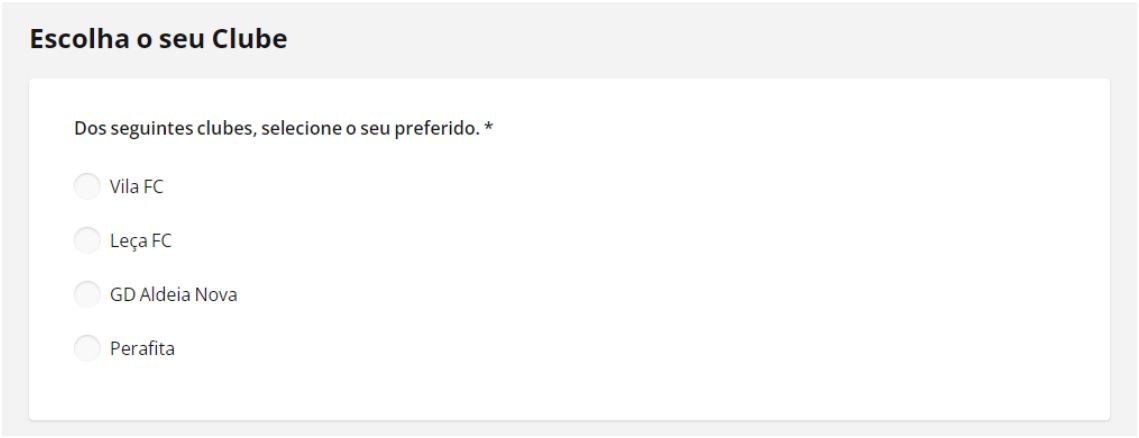


Figure C.2: Second page of the Questionnaire 1

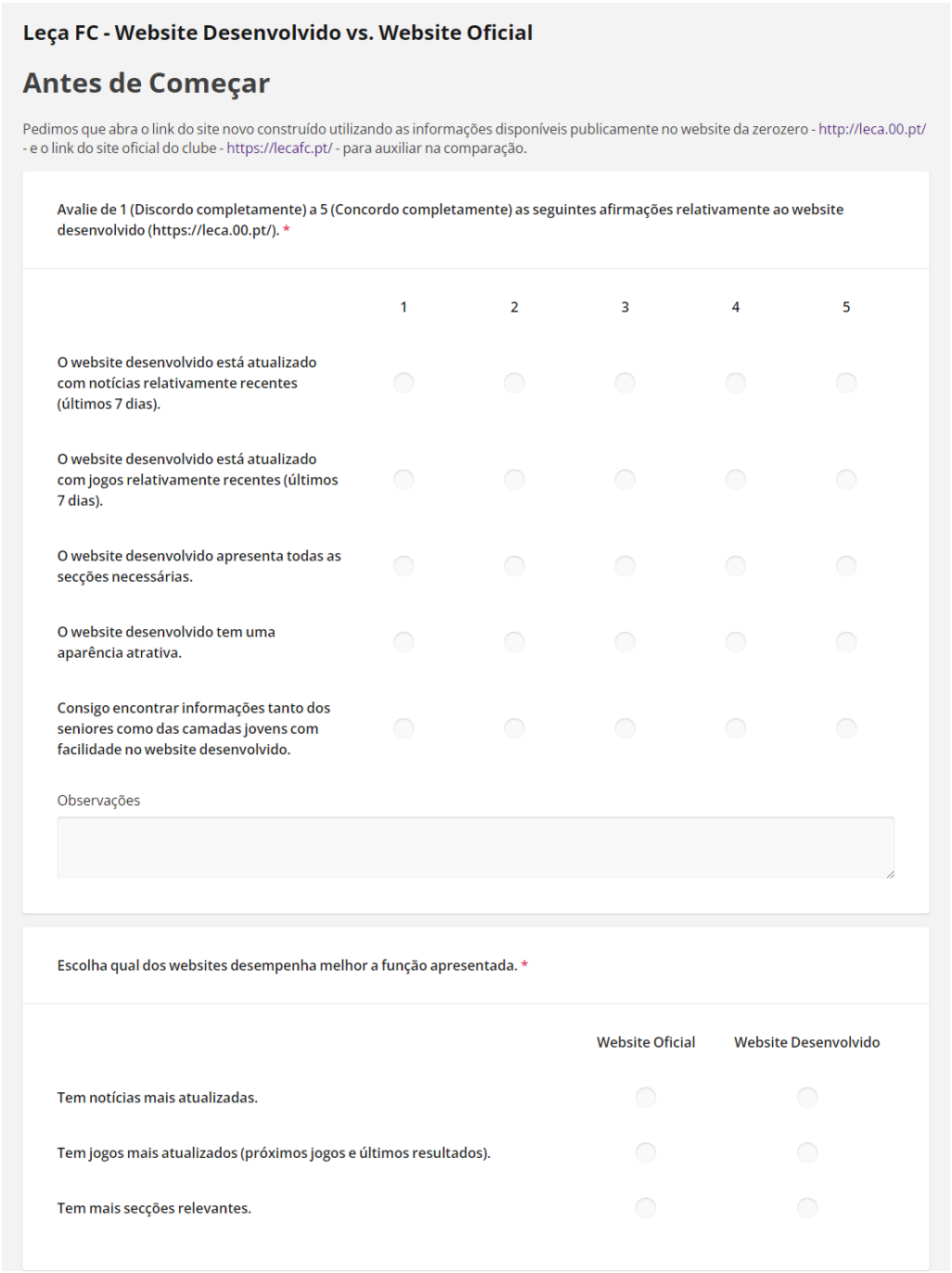
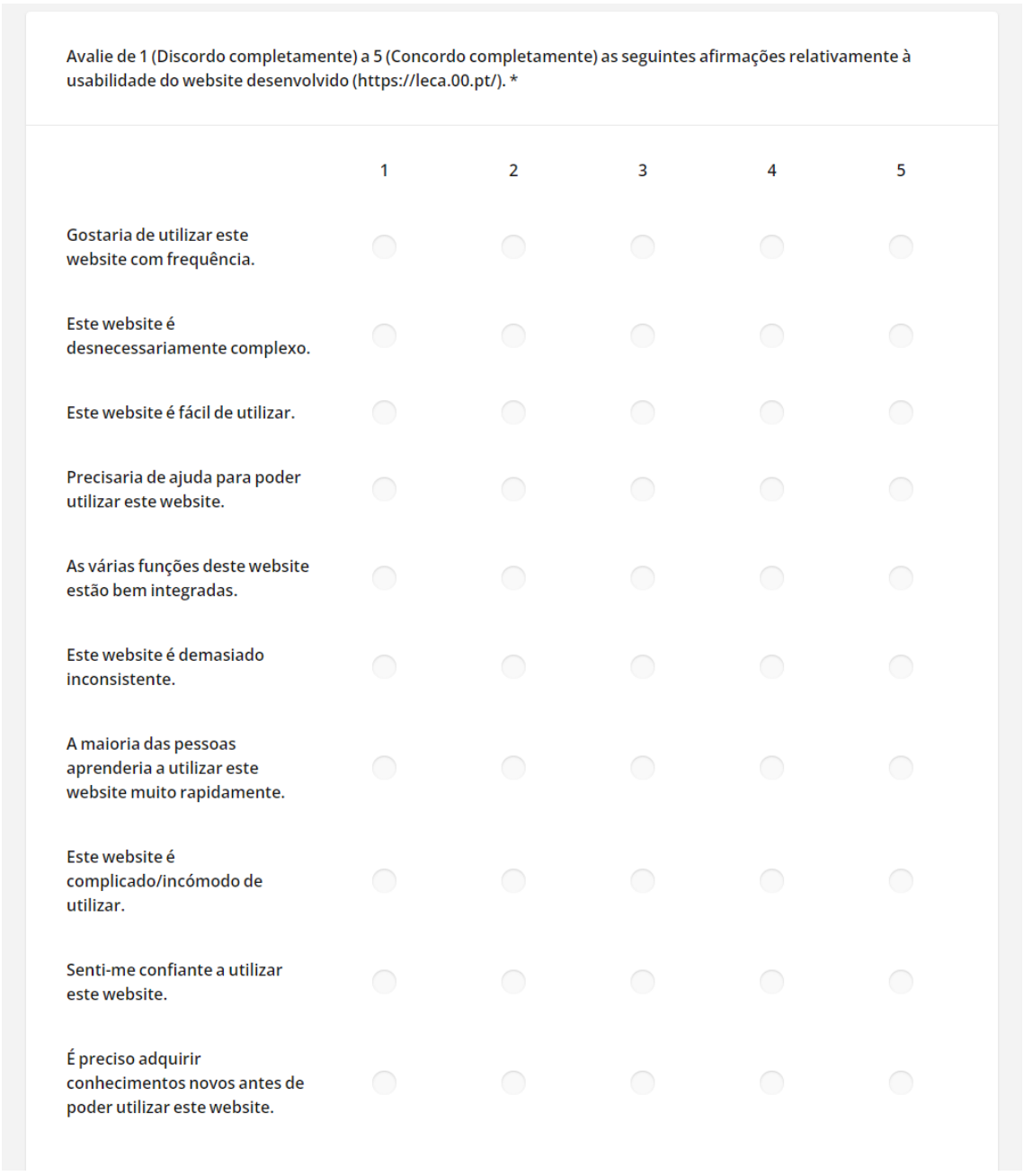


Figure C.3: First part of the third page of Questionnaire 1, after answering in the previous page "Leça FC"



Avalie de 1 (Discordo completamente) a 5 (Concordo completamente) as seguintes afirmações relativamente à usabilidade do website desenvolvido (https://leca.00.pt/). *

	1	2	3	4	5
Gostaria de utilizar este website com frequência.					
Este website é desnecessariamente complexo.					
Este website é fácil de utilizar.					
Precisaria de ajuda para poder utilizar este website.					
As várias funções deste website estão bem integradas.					
Este website é demasiado inconsistente.					
A maioria das pessoas aprenderia a utilizar este website muito rapidamente.					
Este website é complicado/incómodo de utilizar.					
Senti-me confiante a utilizar este website.					
É preciso adquirir conhecimentos novos antes de poder utilizar este website.					

Figure C.4: Second part of the third page of Questionnaire 1, after answering in the previous page "Leça FC"

Figure C.5: First part of the first page of the Questionnaire 2

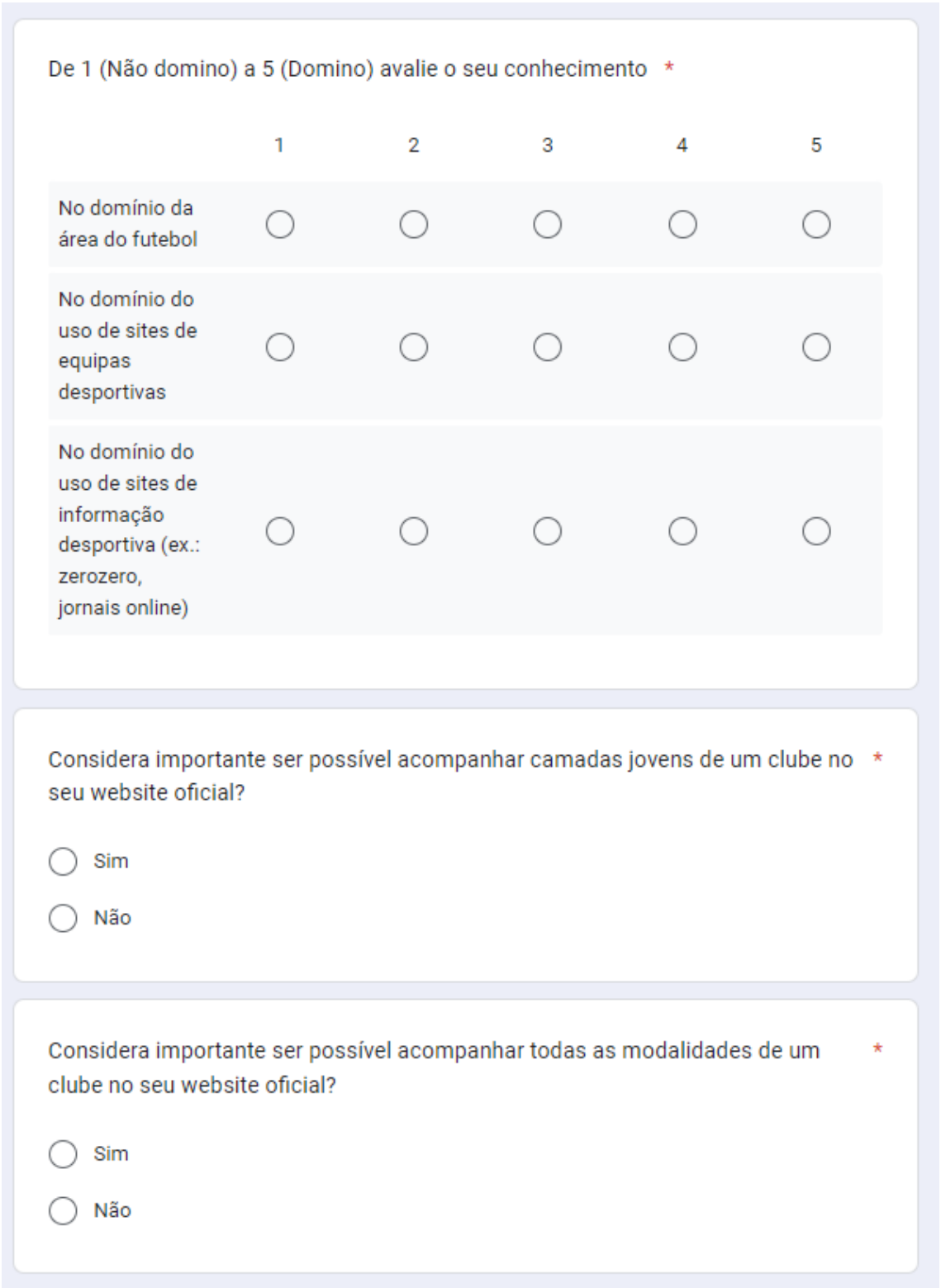


Figure C.6: Second part of the first page of the Questionnaire 2

Website Oficial

Nesta secção deve responder às perguntas que se seguem avaliando o website oficial do clube FC Porto: https://www.fcporto.pt/pt

Tarefa
Abra o link do website oficial - https://www.fcporto.pt/pt - para realizar as próximas tarefas.

No website oficial, quantas equipas de futebol podem ser vistas? *

- 3
- 5
- 10
- 12
- Outra:

Avalie de 1 (Discordo completamente) a 5 (Concordo completamente) as seguintes afirmações relativamente ao website oficial do FC Porto. *

	1	2	3	4	5
Utilizo o website oficial com regularidade.					
A informação disponibilizada no website oficial encontra-se atualizada.					
É fácil obter informações (notícias, jogos, resultados, planteis) sobre outras modalidades que não o futebol.					
É fácil obter informações (notícias, jogos, resultados, planteis) sobre outros escalões que não o senior (no caso do futebol).					
É fácil obter informações (notícias, jogos, resultados, planteis) sobre outros escalões que não o senior em modalidades que não sejam o futebol.					

Figure C.7: Second page of the Questionnaire 2

Appendix D

Monte Carlo Simulation

The appendix includes the Monte Carlo simulation code used in the research study. This Python code performs a series of simulations to calculate the proportion of preference for the "*Desenvolvido*" outcome compared to the "*Oficial*" outcome. The code accurately determines the proportion of preference by randomly assigning preferences and counting the occurrences of "*Desenvolvido*" preference. The Monte Carlo simulation is a valuable tool for evaluating the likelihood of preference and contributes to the analysis of research findings.

```
import random
from decimal import Decimal, getcontext

# Set the precision of decimal calculations
getcontext().prec = 28

def monte_carlo_simulation(results):
    # Define the number of simulations
    num_simulations = 1000000
    count = 0
    num_results = len(results)
    for _ in range(num_simulations):
        simulated_preferences = [random.choice(['Oficial', '
            ↪ Desenvolvido']) for _ in range(num_results)]
        if simulated_preferences.count('Desenvolvido') > results.count
            ↪ ('Desenvolvido'):
            count += 1
    probability = probability = Decimal(count) / Decimal(
        ↪ num_simulations)
    return probability
```