



City Research Online

City, University of London Institutional Repository

Citation: Narsesian, S. (2010). Football Fans, Their Information, The Web And The Personal Home Page. (Unpublished Doctoral thesis, City University London)

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/12140/>

Link to published version:

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Football Fans, Their Information, The Web And The Personal Home Page

Shant Narsesian

February 2010

A thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

City University London

Department of Information Science

Supervised by:
Dr Tamara Eisenschitz



Table of Contents

Title Page	1
Table of Contents.....	2
List of Figures.....	6
List of Tables.....	7
Acknowledgements.....	9
Declaration	10
Abstract	11
CHAPTER I - INTRODUCTION	
1.1 Introduction	12
1.2 The Personal Home Page (PHP).....	12
1.3 Motivation for this Study	14
1.4 Aims and Objectives	17
1.5 Scope	18
1.6 Case Study	19
1.7 Limitations of this Study	19
1.8 Football Background.....	20
1.9 Football Websites.....	23
1.10 Chapter Conclusion.....	24
CHAPTER II - LITERATURE REVIEW	
2.1 Introduction	26
2.2 Search Strategy	26
2.3 Personal Home Pages.....	28
2.3.1 Definition.....	29
2.3.2 Blogs (WebLogs) and Social Networking (SN) Sites.....	32
2.4 The PHP as an Object of Study	36
2.4.1 Characteristics of PHPs and Related Studies.....	36
2.4.2 Purpose of PHPs.....	38
2.4.3 Accessibility of PHPs.....	39
2.4.4 PHP Numbers	40
2.4.5 PHP Problems	41
2.4.6 PHPs and Grey Literature.....	43
2.4.7 Virtual Communities, Communities of Practice and Knowledge Sharing.....	44
2.4.8 PHPs as Sources of Information	48
2.5 Information Seeking.....	51
2.5.1 Information Seeking Models	51
2.5.2 Environments of Information Seeking	54
2.6 Football Studies	56
2.6.1 Football Surveys	57
2.6.2 Football in the Community	60
2.6.3 Football and Legal Issues	62
2.7 Chapter Conclusion.....	67

CHAPTER III - METHODOLOGY

3.1 Introduction	69
3.2 Methodology Overview.....	69
3.3 Stage One People-Study	71
3.3.1 Football Fan Interviews.....	73
3.3.2 Football Fan Interview Questions.....	76
3.3.3 Questionnaires.....	79
3.3.4 Employed Questionnaire	81
3.3.5 Questionnaire Questions.....	84
3.4 Stage One Web-Study	86
3.4.1 Finding PHPs	88
3.4.2 PHP Inclusion	89
3.4.3 PHP Classification	90
3.4.4 Comparing PHPs.....	96
3.4.5 Relative Information Contents Ratings (RICR).....	97
3.4.6 Comparison Results	103
3.4.7 Uniqueness of Information on the Web	106
3.4.8 Accuracy of Information	109
3.5 Stage Two Overview.....	112
3.6 Stage Two Web-Study	112
3.6.1 Website Examination Procedure.....	113
3.6.2 West Ham Online Website Investigation Example.....	114
3.7 Stage Two People-Study	117
3.7.1 Web Author Interviews	117
3.7.2 Web Author Interview Questions	119
3.8 Evaluation and Replicability	121
3.9 Chapter Conclusion.....	123

CHAPTER IV - RESULTS I

- INFORMATION SEEKING TOOLS OF FOOTBALL FANS

4.1 Introduction	125
4.2 Demographics.....	126
4.2.1 Premiership Demographics.....	126
4.2.2 Football League Demographics	128
4.2.3 Comparisons Within This Study.....	130
4.3 The Objectives	131
4.3.1 Tools and Methods Used for Acquiring Football Related Information	131
4.3.2 Reasons for the Use of Particular Tools and Methods.....	140
4.3.3 The Role Played by the Web	142
4.3.4 The Role of the PHP Within the Web.....	149
4.3.5 Football Fan Satisfaction with the Availability of Football Related Information.....	152
4.4 Chapter Conclusion.....	157

CHAPTER V - RESULTS II

- PERCEPTIONS AND FACTS ABOUT PHPS

5.1 Introduction	159
5.2 The Football PHPs	160
5.3 Unique Information on Football PHPs.....	161
5.4 Archival Storage on Football PHPs	166

5.5 The Accuracy of the Information on Football PHPs	168
5.6 Factors that Affect the Availability and Quality of PHPs.....	169
5.6.1 Quantity of PHPs	170
5.6.2 Quality of PHPs	177
5.7 Football Fan Perceptions Concerning PHPs.....	183
5.8 The Causes of Football Fan Perceptions Concerning PHPs.....	190
5.9 Chapter Conclusion.....	193
CHAPTER VI - RESULTS III	
- THE CLUB COMMUNITY COMPOSITE PAGE (CCCP)	
6.1 Introduction	195
6.2 Web Author Interviews	195
6.2.1 Demographics	196
6.2.2 Web Author Motivations.....	197
6.2.3 Website Difficulties.....	198
6.2.4 Site Popularity and User Participation	200
6.2.5 Site Collaboration	201
6.2.6 Web Author Opinions And The CCCP Concept	201
6.3 Website Collaboration Investigations	202
6.3.1 Football Website Franchises.....	202
6.3.2 Supporter Groups	204
6.3.3 Fanzines.....	204
6.3.4 Forums.....	205
6.3.5 Notable Collaborative Efforts.....	205
6.3.6 PHPs to Collaboration	207
6.3.7 Summary of Online Collaborative Efforts.....	209
6.4 Crafting A Communal Site Outline	210
6.4.1 Collaborative Sites - An Ideal Tool	210
6.4.2 The Club Community Composite Page (CCCP)	213
6.5 Chapter Conclusion.....	217
CHAPTER VII - DISCUSSIONS AND CONCLUSIONS	
7.1 Introduction	219
7.2 Foundations of the Thesis - Football Fan Information Seeking	220
7.3 Personal Home Page Conclusions	223
7.3.1 Characteristics.....	225
7.3.2 Accessibility of PHPs.....	227
7.3.3 The PHP as an Information Resource	228
7.3.4 Football Fans, PHPs and Online Collaborations.....	234
7.4 Original Contributions	236
7.5 Future Work.....	239
7.5.1 Web and PHP Studies.....	240
7.5.2 Related Fields	241
7.6 Conclusion.....	243
References and Bibliography	245
APPENDIXES	
Appendix A - Interview Guides.....	261
Appendix B - Information about Interviewees.....	267

Appendix C - Calls for Respondents.....	269
Appendix D - Offline Browsers.....	271
Appendix E - Corroboration of Web-Study Methodology.....	276
Appendix F - Online Questionnaire.....	293
Appendix G - Offline Questionnaire.....	305
Appendix H - PHP Addresses (URLs).....	315
Appendix I - Website Collaboration Investigation Addresses (URLs).....	317
Appendix J - Notable Collaborative Site Addresses (URLs).....	335
Appendix K - Additional CCCP Details.....	336
Appendix L - Notable Collaborative Site Descriptions.....	340

List of Figures

Figures	Page
2.1 Wilson's model of Information Seeking Behaviour	53
3.1 PHP Identification Flowchart	92
3.2 WestHamOnline.net (WHO) front page	115
3.3 WHO FAQ page	116
4.1 Premier League demographics - socio-economic grouping	128
4.2 Football League demographics - socio-economic grouping	129
4.3 Questionnaires and interviews - socio-economic groupings	131
4.4 Notable tools or methods used by football fans for information seeking	133
4.5 Time spent talking about football (per day)	134
4.6 Heavy users of tools and methods	143
4.7 Reasons fans use the Internet (for football purposes)	146
4.8 Types of sites used by fans	150
4.9 Sites meeting the needs of fans	153
5.1 Types of information according to division	165
5.2 Archival (A)/Locally Exclusive (X) Ratings as percentage of total	167
5.3 Degree of accuracy	169
5.4 Number of PHPs per division	170
5.5 Number of PHPs per team	171
5.6 PHP proportions (to stadium capacity, average attendance and total annual attendance)	176
5.7 Geraldine's Viva Manchester United	178
5.8 Southend United Database screenshot	180
5.9 Fan Perceptions - How innovative, comprehensive, up to date and easy to use PHPs are (including online and offline questionnaires).	185
5.10 Fan Perceptions - How interactive, authoritative, independent and quick to update PHPs are (including online and offline questionnaires).	186
5.11 Fan Perceptions - How unique, archival (meaning the information found on them goes far back) and accurate the information on PHPs is (including online and offline questionnaires).	188
6.3 Screenshot of United We Stand (the Manchester United rivals.net site)	203

List of Tables

Table	Page
1 Gauges of public interest in English professional football leagues for the 2003/04 season	22
2.1 Definitions or assumed definitions of PHPs	31
2.2 Personal Home Page Providers within the "Top 50 most frequently linked" websites in 2002	41
2.3 Mode, user state and identification according to Maurer (2006)	54
2.4 The 12 sections of the FA Premier League's National Fan Survey Questionnaire (2005).	57
3.1 Interview questions	77
3.2 Questionnaire questions	84
3.3 Criteria for PHP selection (custom-made using Smith's (1997) Toolbox)	90
3.4 Inclusion and exclusion of PHPs	93
3.5 Results from QPR website search	95
3.6 Relative Information Content Ratings (RICR)	102
3.7 PHP comparisons for QPR	103
3.8 PHP comparisons for QPR (continued)	104
3.9 PHP comparisons for QPR (continued)	104
3.10 PHP comparisons for QPR (continued)	105
3.11 PHP comparisons for QPR (continued)	105
3.12 PHP comparisons for QPR (continued)	106
3.13 Locally Exclusive and Web-Wide Unique instances of information (QPR)	108
3.14 Degree of accuracy	110
3.15 Accuracy of information (QPR)	111
3.16 Stage One Web-Study: Step by step	112
4.1 The FA Premier League National Fan Survey (2005) and this study's questionnaire (online and offline combined, showing the Premier League only).	127
4.2 Football League Supporter survey (2006) and questionnaire (Football League only).	129
4.3 Demographics of the two sets of questionnaires and the interviewees	130
4.4 Percentage of respondents using tools or methods for information seeking	132
4.5 Time spent per day by respondents talking about football (percentages)	134
4.6. Contents of the Newcastle United FC Annual Report (2005/06)	138
4.7 Reasons given by heavy users of any method for the use of that method (online and offline combined).	141
4.8 The first tool or method respondents would use when looking for information	144
4.9 Main reason why fans use the Internet in relation to their football needs	145
4.10 Time spent on the web reading about their team	147
4.11 Number of people that use each type of website	150
4.12 Frequency of PHP use	151
4.13 Frequency of PHP use against time spent on the web (online)	152
4.14 Sites meeting the needs of fans (interviewees)	153
4.15 Sites meeting the needs of fans (5 point scale) against time spent on the web (online).	154
4.16 How often fans find what they are looking for (online and offline)	155
4.17 How often fans find what they are looking for (interviewees)	155
4.18 How often fans find what they are looking for against time spent on the web (online and offline).	156

Table	Page
4.19 How often fans look for a specific piece of information	156
5.1 PHP classification	161
5.2 Ratings per PHP - These are ratings (RICRs) acquired from PHP section comparisons (i.e. the numbers represent PHP sections).	164
5.3 Ratings per division	165
5.4 Instances of archival storage according to division	167
5.5 Degree of accuracy of PHPs	168
5.6 Number of PHPs, average attendance, stadium capacity and total annual attendance per team (2005/06 season, Soccer-stats.com Site 2006)	174
5.7 Features of notable PHPs	182
5.8 Notable PHPs and their RICRs	183
5.9 Don't Know replies on questionnaires (online and offline)	184
5.10 Replies from respondents when asked how Unique, Far Back and Accurate the information on PHPs was (including online and offline questionnaires).	187
6.1 Web Author Demographics	196
6.2 Motivations of Web Authors	197
6.3 Demotivations of Web Authors	198
6.4 List of Notable Collaborative Websites	206
6.5 Instances of PHPs that became non-PHPs	209
7.1 Objectives 1 to 5	220
7.2 Objectives 6 to 11	225
7.3 Objective 12	235

Acknowledgements

This document has been a long time in the making, and now that it is actually finished there are some people whom I would like to thank.

Firstly, my supervisors, Dr Stephen Morris and Dr Tamara Eisenschitz. Stephen was with me in the very beginning and was always supportive and full of ideas, even after his retirement. At the other end, Tamara was extremely helpful at a time when so much needed to be done. I am very grateful to both for everything they did for me.

In addition to my supervisors, there are numerous colleagues and friends who were extremely helpful over the years, each in their own way. These include Gilberto Cysneiros Filho, Dr Rodrigo Agerri, Dr Edu Alonso, Chris Child, Dr Marcus Pearce, Dr Panyiota Polydoratos and Abeer Farsoun.

In this instance, I am forced to mention Dr Hamid Reza Jamali Mahmuei separately, not only for the countless hours he spent listening to my PhD worries, but also for the countless hours he spent reading my work and giving advice, when he really had better things to do.

I would also like to thank my boss, Michael Iossif for being as accommodating as possible (at times when I needed all the accommodation I could get) and Olga Castillo, who was always willing to exchange this work for that, even when she would rather not have. I can only hope that I did not make their lives too difficult.

Most of all though, I must thank my family. My parents, John and Sylvia, for supporting me in anyway that I needed, whether it was love, advice or money. My brother, Andy, without whose time and ideas I would be lost. I also want to thank Vlatka for having the patience to deal with my endless conversations with Andy. It would be impossible for me to overstate how valuable the contributions of my family were over these years.

Finally, in addition to all the people I have thanked, I would like to thank God for giving me the strength to carry on during many difficult times.

Declaration

This is a declaration to grant powers of discretion to the University Librarian to allow the thesis to be copied in whole or in part without further reference to the author. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

Abstract

From the early days of the Internet to the present day, the World Wide Web has developed into one of the world's largest information resources. One of the first genres of web pages, which was also one of the first information resources, was the Personal Home Page (PHP). Over this same period of time, professional football in England has created the world's richest league and by extension an abundance of football related PHPs. This study investigates the role of the PHP as an information resource using the subject area of professional football in England.

A holistic approach is taken so as to view the PHP from a broader context, as one information resource amongst many, including non-PHPs and even offline information resources (e.g. reference books). Within this study, football fans are interviewed along with web authors, surveys are carried out (by distributing both online and offline questionnaires) and research is also carried out online, examining football related PHPs and online web collaborations.

Results suggest that whilst there are many informational benefits to be found on PHPs, such as plentiful unique information, they have low levels of use amongst football fans. The thesis concludes by proposing an avenue to the maximisation of the informational benefit of PHPs through a blueprint for a type of communal football website called the Club Community Composite Page (CCCP).

Overall, several contributions are made to the field of information science, most notably attaining an improved understanding of PHPs as unique and accurate information providers online and devising new research methods for PHP research. In particular, the method of identification of PHPs developed here will be a useful tool for future researchers of PHPs. The contributions of this thesis are likely to be of value to researchers working in relevant sub-fields of information science, such as information seeking, web genres, grey literature and virtual communities.

Chapter I

Introduction

1.1 Introduction

When the World Wide Web first came to being, it was primarily a means of making information easily available and accessible. Since then, it has developed many other uses, such as a real time communications medium (video-conferencing on the web), a sales point (e-commerce) and even the world's largest electronic playground (online gaming). However, the web's first role helped it achieve its status as one of the world's largest information resources. One aspect of this resource, is the Personal Home Page (PHP)- a website which represents the point of view of one individual (de Saint-Georges, 1997). At the same time, the popularity of professional English football has grown, with the Premier League achieving the status of the world's richest single league (BBC Website 2006d). In these terms, this thesis aims to examine the role of the PHP as an information source within the field of professional football in England.¹

1.2 The Personal Home Page (PHP)

The PHP is an entity that has attracted various kinds of interest from numerous fields. From an academic point of view, the PHP has been used from the very beginning (e.g. Koch's Personal Homepage 1993), as a portal to the individuals who created them for both work and non-work purposes. Initially the computing department's plaything,

¹ The subject area for the case study in this thesis was the top four professional leagues in England. It is acknowledged that over the years, Welsh teams have played in these divisions. However, so as to differentiate from the semi-professional Principality Welsh Premier League, this subject area is referred to as the professional English leagues (and numerous variants thereof) without the term "Wales" included.

soon it was available to all academic staff, and then students. At the time of writing (but also for the foreseeable future), millions of PHPs exist, in various forms and guises.

Once so many PHPs were created, it was inevitable that the PHP would be looked into further, examined and researched. PHP research has been ongoing now for over 10 years. Investigations have been carried out within numerous academic fields of interest, including computer science (e.g. Hoff and Mundhenk 2001), librarianship (e.g. Haines 1999), information science (e.g. Thelwall 2002), psychology (e.g. Joinson and Banyard 2002), education (e.g. Arnold and Miller 2000) and journalism (e.g. Dominick 1999) to name but a few. However, even though the interest spans many fields, there has not been a great deal of in-depth research carried out and there are still many gaps in our knowledge of the PHP. It would not have been unreasonable to expect more to have been known by this stage about PHPs as a whole, particularly their role as an information resource and the information available on them. This thesis aims to go some way in providing some potentially invaluable insight into this feature of PHPs.

In addition to this, a secondary though still potentially very beneficial and insightful aspect of this thesis is the examination of various aspects of the information seeking behaviour of football fans who follow the teams in the top four English leagues. There are, quite literally, millions of football fans in the United Kingdom alone and the English leagues have a reach way beyond their own borders. This is especially true presently, with football teams making a special effort to extend the range of their appeal with summer tours to the United States and the Far East. Traditionally, the fans of these teams, both inside and outside the UK, have been restricted to the more traditional media in order to find out about their football clubs, including most notably the newspaper, the television and the radio. Nowadays, with the explosion of the web and all that comes with it, the availability of information concerning any professional football team has grown significantly, even more so in England (with the English language being the most popular language on the Internet). This makes such a thesis interesting from more than one perspective. From an information science perspective, it is interesting to see the information seeking behaviour of people who are not seeking information for professional purposes (e.g. as librarians or physicists would)

and look at the role played by the web in providing a medium for this information. Furthermore, it is useful for those involved in the football industry to see not only how their fans behave when it comes to the acquisition of football relevant information, but also the role they play in the collection and dissemination of this information (e.g. through PHPs).

1.3 Motivation for this Study

The main aspect of this thesis is the examination of PHPs as an information resource, with a particular focus on aspects of information found on them, but also people's perceptions of PHPs and the reasons behind them. There are many reasons why the study of PHPs is an interesting one. They are said to provide insight to numerous issues, including gender issues (Arnold and Miller 2000, Flanagan and Metzger 2003), self presentation (Wynn and Katz 1997, Papacharissi 2002a), interpersonal communication (Dominick 1999), web genres (Dillon and Gushrowski 2000) and grey literature (Thompson and Guistini 2006). This thesis however, is primarily concerned with their benefits in terms of information storage and dissemination.

Traditionally, supplying information has been an expensive and time consuming business. On a broad scale, parties and institutions which have supplied information have been motivated by money (e.g. newspapers, publishing houses) or other agendas, most notably political (e.g. governments) or religious (e.g. the church). With the advent of the web and the introduction of PHPs, the ability to supply information has been given to just about anyone with a computer and an Internet connection. As a result, people who create PHPs have the ability to fill in information "gaps" in subject areas where the traditional motivations or agendas are not present. As such, any well-founded study that provides additional insight into the workings of this information resource is of value.

It also can be argued, that in a certain sense, PHPs are a resource created out of nothing. The fact that ordinary people spend time making information freely available for other people is hugely beneficial for information consumers, and particularly fascinating for information scientists as well as other information providers. In

addition to the potential of filling informational gaps, the sheer scale of the information provided alone (i.e. in terms of numbers of PHPs available) makes it worthwhile to investigate this phenomenon further.

Meanwhile, using football as a case study, creates another series of interesting aspects that would be useful to look into, either directly or otherwise. Many studies have been carried out looking into the informational behaviour of various groups of people. The titles often include broad groups such as scientists (e.g. Brown 1999), academics (e.g. Ocholla 1996), scholars and students (e.g. Bates 1996), but also more specific groups such as librarians (e.g. Schreiber and Moring 1997), mathematicians and physicists (e.g. Brown 1999), lawyers (e.g. Wilkinson 2001) and even members of parliament (e.g. Orton, Marcella and Baxter 2000). However, all these groups of people are either involved in professional occupations, or potentially involved in professional occupations (i.e. the students). In this study, the information related behaviour of the football fans is also being examined (as a result of the holistic approach being used), and this provides an insight into the informational behaviour of a group of people not normally investigated in this way. As this is an exploratory study, it will be difficult to use these results for direct comparison with studies of the "professional" groups. However, in the longer term this will help provide a new angle in the viewing of all information related studies. Indeed, to date, such studies which look at the information resources available to ordinary citizens, particularly from a football fan's point of view, have been very sparse.

Furthermore, an additional football related positive aspect of this study is its relative independence in terms of stakeholders. Such studies are often commissioned by bodies that have a vested interest in certain aspects of the field (e.g. businesses that are looking for ways to increase their revenue). In this instance, the study is purely the initiative of a football fan and without any reason to promote one group over another; the results of this study should help improve communication and interaction between all involved parties, not least between the clubs and their fans.

In academic terms, football as a field for information science studies is interesting because of its unusual information environment. This environment is a product of the various parties that are involved in football and the manner of their involvement.

Firstly, the largest group (in terms of numbers of people) of interested parties is the collection of fans who have primarily an emotional investment in football (and sometimes a monetary one too). For the fans, football is not only an interest but often a way of life, with strong social implications. Secondly, there are the individuals involved in the performance of individual clubs, such as the directors, the players, the coaches and so on. Their involvement is slightly different, as they too have an emotional investment at some level, but naturally their predominant concern is (in most cases) the state of their individual careers. Thirdly, there are related businesses (such as the television companies, clothes manufacturers), that have a financial interest either in the performance of a specific club (e.g. a club sponsor), or a financial interest in the well being of football as a whole (e.g. a football broadcaster). Finally, there are certain organisations that try to referee disputes between these parties, and whose role it is to look out for the interests of everyone involved (i.e. national governments and football associations). All of these parties take part in the absorption and dissemination of information concerning football, creating the aforementioned unusual information environment.

At the same time, with the growth of the Internet in recent years, an extra twist is now added with the relative ease of information exchange that the web has allowed between all the parties. Football fans themselves have a variety of types of sites which they use to publish information and communicate with each other (including the PHP), while those involved in football have their own sites (i.e. the club sites), as do the related businesses (e.g. British Sky Broadcasting) and the football associations. This, in combination with the fact that the English football industry is one of the largest in the world (Harding 2003, BBC Website 2006d) and that English is the primary language on the web, makes professional football in England a field which is particularly suited to such a study.

Finally, in the longer term, projects such as this will provide data which will facilitate a better understanding of what football fans use in their search and consumption of information, including aspects of the web and PHPs. In the shorter term, this thesis provides a more specific direction which can be taken in order to improve the information seeking experiences of football fans in England and more specifically the use of PHPs. With the millions of PHPs online (Pickavet 2006, Riley 2005, Wired

Magazine 2000, Lawrence and Giles 1999), and millions of individuals putting effort into the creation of sites for the benefit of others, without a systematic study, it is very difficult to say what the effects of these efforts are. If one takes the view that these sites are a "natural resource", (seeing as there was no concerted effort in initialising their creation), more studies should look into ways of harnessing the power of these PHPs.

1.4 Aims and Objectives

The main aim of this thesis was to investigate the role of the PHP as an information resource with the view to making a recommendation for improvements that would help maximise the benefits gained by the efforts put in by PHP and other web authors (as well as others active in the football Internet community). An attempt was made to understand the popularity of the PHP, what their users think of it, what they themselves offer, what they could potentially offer and what problems afflict them. In order to achieve this aim, a popular field was chosen where the web is heavily utilised and PHPs, their users (in the general public) and their potential users are readily available. A field which meets such criteria is professional football in England. Having selected a suitable field, an all round approach was taken, looking at all the ways in which fans acquire football related information and within this, the role of the PHP.

The specific objectives of this thesis were therefore:

1. To determine the tools and methods ordinary football fans use to acquire their "footballing" information.
2. To determine the reasons for which these tools and methods are used.
3. To determine the significance of the role played by the web in providing this "footballing" information.
4. To determine the significance of the role played by the PHP within the role of the web in providing this "footballing" information.
5. To determine whether, all in all, football fans are satisfied with the availability of "footballing" information.

And more specifically about PHPs:

6. To determine the number of sections that have instances of unique information available on football PHPs.
7. To determine the number of sections that have instances of archival storage on football PHPs.
8. To determine the degree of accuracy of the information available on football PHPs.
9. To determine the factors that have an effect on the availability and quality of PHPs.
10. To determine the perceptions that football fans have concerning PHPs.
11. To determine the causes behind the perceptions that football fans have of PHPs.

Having completed the research in order to meet these objectives, a recommendation was deemed suitable, and an additional objective was introduced:

12. To craft an outline for a communal site that can encourage web authors to work together in the creation of web pages for the online community.

1.5 Scope

In an attempt to meet the aims and objectives specified, this work examined certain aspects of information seeking behaviour and occasionally touched on information needs aspects (so as to give as complete a picture as possible). As this was a one-man project with coinciding constraints, it was not possible to look into all the aspects that might have been useful to this thesis (see 1.7 Limitations of this Thesis, p. 19, later in this chapter). Concerning PHPs, samples from the four top professional leagues in England were taken and indications of their representative nature have been given. The focus was on the content of the PHPs where the bulk of the information comes from non-PHPs (i.e. organisational and news sites). Non-PHPs were also looked at, though in less detail, to help ensure that key facts about football websites were not

overlooked. Clearly, more studies could have been done, including studies where PHPs are the dominant form of information (on the web), but the chosen subject area, where there are varying amounts of information provided by the various groups (i.e. the amount of information available on the highest football division is not the same as that in the lowest division), is arguably the most useful.

1.6 Case Study

The chosen subject area for this research is the field of professional football in England. The benefits of this selection include having a good deal of data and also having more than just one "informational environment" (i.e. the four separate leagues). This includes the 92 professional football teams in the top four leagues (Premiership, Championship, League 1 and League 2) in England (though this also includes three teams based in Wales). The PHPs created about teams from the set of 92 were examined and the opinions of their fans were canvassed. This was felt necessary so as to be able to determine whether the position in the league had an effect on the type of sites available. Meanwhile, every year, teams move up and down the divisions as they gain promotion or suffer relegation. For this study, the league tables used were primarily those of the 2004/05 football season. For the research carried out to meet objective 12, the 2006/07 league tables were used.

1.7 Limitations of this Study

As this study was primarily the work of one man, the scope of the research was limited in certain ways. All the limitations mentioned here are present because of resource constraints, primarily time.

One limitation of this study is its size. Interviews were conducted, questionnaires distributed and websites examined. The numbers of all three could have been increased to enhance the results further. Meanwhile, though every effort was made to ensure the sampling of all three data collection methods was as representative as possible, ultimately for feasibility related reasons, at times, trade-offs were made, and these must be taken into account when reading the results.

In terms of PHPs, the case which is examined in this thesis, is one where most of the information on the web is supplied by non-PHPs (dedicated football sites and news sites e.g. football365.com, BBC.co.uk). The other side of the spectrum is the instance where PHPs provide the majority of information on a topic. Though this study provides some coverage from this viewpoint, it is not the emphasis of the work. It is possible that there are subject areas where the majority of the information available on the web is provided by PHPs and in the longer term, it is the intention of the author to carry out such studies. However, this was deemed a more suitable starting point.

On a related note, another limitation is the fact that this thesis uses only one case study. Again to provide a more complete picture of the field of PHPs as a whole, it would be useful to look at a greater number of case studies in other subject areas, so as to have a greater perspective on the issues being examined. In the longer term, more case studies should be carried out.

1.8 Football Background

At the time of writing, the footballing information environment is rich in all the necessary aspects. There are great numbers of both football fans and football PHPs. In the UK, according to one report, there are said to be 18 million football fans (Bradford and Reeve, 1997). Though there is no accurate number of football PHPs available, as already mentioned, the numbers of PHPs in general are said to be in the millions (Lawrence and Giles 1999, Riley 2005). The combination of millions of football fans, millions of PHPs and early investigations conducted as part of this research, all suggest that the number of football PHPs is sufficient.

Meanwhile, parts of the subject area (particularly the Premier League) can be considered "informationally saturated". There are several communication media which give constant daily coverage of the current state of affairs. These include newspapers (e.g. The Guardian, The Sun), football magazines (e.g. FourFourTwo, Shoot) television (e.g. Sky Sports 1), including dedicated sports news channels (e.g. Sky Sports News), websites (e.g. soccer.net, football365.com) and radio (e.g.

Radio5Live) including dedicated sports stations (e.g. TalkSport). Whereas this suggests that perhaps the non-PHP football websites might have every fact of football covered (therefore making PHPs less important in informational terms), it also means that finding evidence to the contrary is more definitive. This section provides figures concerning the top four professional football leagues in England as a whole, but also certain figures categorised by league (i.e. each division separately). Though the figures provided do not cover one particular year, they are still illustrative of the differing amounts of information available (and general interest in) each of the leagues.

Most of the media have an abundance of football coverage. On the Internet at the time of writing, there are 2197 sites listed in the Open Directory website in the professional football categories, whilst Yahoo! has 746. On television in the UK, there are a multitude of channels (both terrestrial and satellite or cable)² which have shown football matches (or clips of matches) from this field. In addition there are three (Sky Sports News, Eurosport News, Setanta Sports News) dedicated sports news channels, one of which (Sky Sports News) averages over 15 hours of football news per day.³ On the radio, there are on average, 29 programmes per week about football.⁴ In the shops, there are at least three football magazines available in ordinary newsagents in London⁵. More recently, various new forms of information delivery have also been introduced, (e.g. mobile Internet, live match broadcasts on the Internet) and such trends are set to continue as more football clubs turn into more business conscious institutions, fully concentrating on the image of their "brand".

Further evidence of the popularity of professional football in England is provided by the "financial attention" it receives. The Premiership is arguably the "richest" football league in the world, with 25% of the European football industry (Harding 2003). In

² BBC1, BBC2, ITV1 are the terrestrial channels whilst other channels digitally available such as ITV4, British Eurosport, numerous SkySports channels (1, 2, 3, Extra as well as High Definition channels), numerous Setanta Sport channels, pay per view channels and club channels (e.g. MUTV, Chelsea TV, Arsenal TV) also show relevant football matches or clips of matches.

³ Based on the names of programmes with the word football in the title or description. Average in the month of September 2008 (exact figure: 15.57 hours per day).

⁴ A search for programmes containing the word football in the title was carried out and revealed 29 (exactly 29.5) programmes a week on average in the UK. The search was carried out in September 2008 (total number of shows was 118).

⁵ Three out of the following four football magazine were found in 10 London newsagents (10 newsagents were checked in total): FourFourTwo, Shoot (Monthly), Match Magazine, World Soccer.

the list of the 20 richest⁶ football clubs in the world (2007/08 season), 6 are in the Premiership (Deloitte and Touche 2008).

Meanwhile, in terms of getting a variety of "informational environments", the chosen field has four divisions, each with a varying degree of public interest and attention. At the one end is the top division of professional football in England (the Premier League), where the concentration of information is at its peak and at the other end is League 2, which receives noticeably less coverage.

League	Live Televised Matches ⁷	Average Attendance ⁸	Websites on Dmoz ⁹	Websites on Yahoo! ¹⁰
Premier League	106	34,900	906	414
Championship	50	15,908	467	332
League 1	10	7,486	437	
League 2		5,389	387	

Table 1 Gauges of public interest in English professional football leagues for the 2003/04 season (television coverage, football match attendance and numbers of Internet websites from the Open Directory and Yahoo!).

Table 1 has certain figures which can be used as a gauge to indicate the level of public interest in the leagues, as well as some informational coverage. It is clear from the table that the Premier League gets significantly more television coverage (at least in live televised matches) than any of the remaining leagues. The number of people attending the matches at the stadiums also appears to indicate that there is noticeable variation in the divisions. The final two columns on the right show the number of

⁶ Richest in terms of annual turnover

⁷ SkySports.com Website 2004.

⁸ StatMail (www.statmail.co.uk).

⁹ Premiership figures at: DMOZ Sports > Soccer > UEFA > England > FA Premiership

Division 1 figures at: DMOZ Sports > Soccer > UEFA > England > Football League > Division 1

Division 2 figures at: DMOZ Sports > Soccer > UEFA > England > Football League > Division 2

Division 3 figures at: DMOZ Sports > Soccer > UEFA > England > Football League > Division 3

All figures checked on 13/05/04.

¹⁰ Premiership figures at: YAHOO! Directory > Regional > Countries > United Kingdom > Recreation and Sport > Sport > Football > Leagues > Premiership

Remaining division figures at: YAHOO! Directory > Regional > Countries > United Kingdom > Recreation and Sport > Sport > Football > Leagues > Nationwide Leagues

All figures checked on 27/04/04.

websites categorised per division on the Open Directory Project (Dmoz.org) and on the Yahoo! websites.

Here again, the indicators are also reinforced by football finances. In terms of nationwide television coverage in the UK, British Sky Broadcasting (BSkyB) and Setanta together paid £1.7 billion pounds for three years of Premiership matches (BBC Website 2005a), whilst only £110 million was paid for three years of coverage for the remaining three leagues (from BSkyB and ITV). Meanwhile the average player salaries (per annum) for the four divisions respectively in 2001-02 (Chaudhary 2003) were £600,000, £200,000, £70,000 and £42,000.

1.9 Football Websites

Football information generally revolves around the teams or the football matches. The most sought after information involves player transfers, player injuries, incidents that took place during matches (goals, tactics, player performance), and reaction after the match. Another aspect of football information concerns the rumours, which cannot be verified and are often incorrect (if not completely fabricated). The football websites which provide all these bits of information are very plentiful.

The most popular sites used by football fans in the Premiership are the BBC website, followed by the SkySports website followed by the FA Premier League's own site (FA Premier League 2008). Official websites play a big role for football fans with 45% using them once a day (FA Premier League 2008). At the same time, newspaper websites play a substantial role as well, with 30% of Premiership fans visiting them “usually” (FA Premier League 2008).

Naturally, every team has an official website, which helps provide the team with its public image. The official site will provide the standard information about the club (as any business website) such as information about the players, facilities, directions to the club and so on. The site also has other uses, such as an outlet for press releases and a sales tool, selling not only tickets to matches, but all official club merchandise. The majority of all official team websites are run by one company, Premium TV

(PTV). In the football league (the three professional leagues below the Premier League), all the club websites are created by Premium TV, with numerous clubs in the Premiership also signed up to the company (e.g. Aston Villa).

Meanwhile, there are also three companies that offer football website "franchises". FootyMad.net and rivals.net have a dedicated site for each professional football club in the top four leagues, whilst SportNetwork.net has dedicated sites for 24 of the 92 football clubs.¹¹

Other popular football sites include team forums or messageboards, where members discuss all aspects of the football clubs. In recent years, forums have become very popular, with larger teams having several forums recording heavy activity. Recently, they have even had a direct effect on professional football. In one instance a football player changed his mind about moving clubs when he read what the fans of this prospective club thought of the idea (BBC Website 2005b).

Local newspapers also play an important role in the dissemination of information about the local team, and by extension, their websites also play a part in providing information about the team, particularly where player transfers or rumours about them are important. Finally, other common football related sites include fan sites¹² (e.g. Celtic Supporter in London), webzines¹³ and of course, PHPs.

1.10 Chapter Conclusion

The vast number of PHPs on the Internet have, to date, been studied from numerous points of view within various fields such as psychology, information science, computer science and others. They have been useful to study as a way of gaining insight into areas such as gender issues, self presentation and interpersonal communication to name but a few. However, in the recent past, good reasons (e.g. the vast numbers of PHPs available) have emerged to believe that they are of value as information resources. For the purpose of studying the PHP as an information

¹¹ Checked on the Sport Network homepage (www.sportnetwork.net) on the 26th of July 2007.

¹² Fans of a team create a site to attract other fans of the same team, so as to arrange social events (i.e. watching a game all together in a public venue).

¹³ The online version of a fanzine.

resource, the subject area of professional English football was deemed suitable thanks to the varying informational environments it provides (a vast amount of information and PHPs for the Premier League, yet significantly less information and fewer PHPs for the lower leagues).

The next step was to establish a firm foundation from which to carry out such a study, namely the examination of all current relevant literature. This is detailed in the following chapter.

Chapter II

Literature Review

2.1 Introduction

This study took a holistic approach in the investigation of PHPs in the field of professional football in England. As a result, many fields of research needed to be covered, so as to ensure that this, a relevant and meaningful literature review, was carried out. The finding and gathering of this literature was carried out intensively at the beginning and at the end of the process which produced this thesis, but also throughout the course of study.

This chapter firstly describes the overall search strategy used. It then proceeds to examine literature looking at various aspects of the PHP (including blogs), and its relationship to grey literature, virtual communities, communities of practice, knowledge sharing and football studies (including information related football studies and football in the community). It concludes by looking at legal issues relevant to football related information on the Internet.

2.2 Search Strategy

The strategy for finding relevant literature consisted of two main parts. On the one hand were the searches for subject specific literature (e.g. searches for Personal Home Pages), whilst on the other hand, there was the reading of relevant journals in the field over a prolonged period of time.

The searches carried out in the beginning and the end of the study were carried out on a number of online and offline databases, including the Library and Information Science Abstracts (LISA), Web of Knowledge, Bath Information and Data Services (BIDS) and Research Index (CiteSeer). Searches were also conducted using both Google Scholar and "ordinary" Google on the web to catch any other traces of relevant work. The advantage of using the Google searches lay in the non-subject specific searches that were carried out, ensuring that relevant material in other fields was also picked up.

Generally, the searches carried out on most relevant subjects, such as information seeking or football, were broad and would return a multitude of results which then had to be sifted through. This ensured that crucial literature was not missed, and was a process which involved reading a very substantial number of abstracts. The main instance where sifting through large numbers of returned results was not necessary involved PHPs, where searches would often return few articles. In addition, for PHPs in particular, it was necessary to use multiple phrases so as to cover all possible literature. For example, searches for PHP relevant work used phrases such as Personal Home Pages, Personal Homepages, Personal Web Pages, Personal Webpages, Internet Homepages, Internet Home Pages, Private Home Pages, Private Homepages, Homepages and Home Pages. In addition to this, in approximately the same area, searches had to be carried out concerning blogs and social networking sites. However, in all instances, where relevant articles were found, all references within the papers were closely examined, an action that led to the discovery of many other relevant articles.

Meanwhile, whilst carrying out this research, the tables of contents of relevant journals were being browsed, including the Journal of the American Society of Information Science and Technology (JASIST), the Journal of Documentation (JDOC), Library and Information Science Research (LISR) and the Journal of Information Science (JIS), to ensure that major relevant works were not overlooked. In addition to this, the author was alert in terms of literature seeking in general to ensure that any relevant literature that was come across in any other way was also followed up on.

Finally, it is worth mentioning that what is presented here is not all the literature that was read by the author in order to conduct the research.

2.3 Personal Home Pages

The Personal Home Page (PHP) as a subject of research spans many topics. These include computing, information science, psychology and even journalism to name but a few. As a result, it has acquired a multiplicity of meanings, making any single all-embracing definition problematic. A number of attempts at defining the PHP have been made (de Saint-Georges 1997, Bates and Lu 1997, Dominick 1999, Weaver 2000, Doring 2002) though there is not a great deal of agreement. In addition, numerous PHP studies have been carried out without an explicit definition. In fact, even the term "personal home page", though clearly the most common (e.g. used in de Saint-Georges 1997, Bates and Lu 1997, Chandler 1998, Arnold and Miller 1999, Dominick 1999, Dillon and Gushrowski 2000, Papacharissi 2002a, 2002b, Rick 2007) is not universal, with other names also used to describe the same concepts, most notably the "personal web page" (e.g. Haines 1999), "private home page" (e.g. Doring 2002), "Internet home page" (e.g. Wynn and Katz 1997) or simply "homepage" (e.g. Hoff and Mundhenk 2001). The end result is a difficulty in describing the Personal Home Page (PHP) as a site. On the other hand, there are some shared ideas. Though certain types of sites are occasionally included or excluded in the definitions, there is no question that the core entity that is being described is the same- namely, sites that have been created by single individuals. Taking this into account, two relatively new (and thoroughly relevant) types of sites have also come to being. The blog¹⁴ (weblog) and the social networking site (e.g. a Facebook page or MySpace page) are two examples of Personal Home Pages that have emerged in recent years. As sites that are both (more often than not) created by and representative of individuals, any definition of the PHP should not exclude them.

¹⁴ We will use the term blog instead of weblog in this study for clarity purposes, as weblog also has a technical computing meaning.

2.3.1 Definition

As mentioned previously, a number of definitions have been attempted. These include academic definitions, (de Saint-Georges 1997, Bates and Lu 1997, Dominick 1999, Weaver 2000, Doring 2002) as well as non-academic attempts (e.g. CWIS Style Guidelines 2002). However, even within academia, the definitions vary, perhaps because of the numerous fields that study the PHP. Weaver (2000), for example, defines the PHP as "pages [which are] wholly under control of individuals, and not functioning as official library pages". Though clearly not an in depth definition, it is clear that such a definition would originate from studies in librarianship.

A more in depth definition is provided by de Saint-Georges (1997), whose definition states that a PHP is a "presentation of the self in digital (hypertextual) form, authored by one individual, and which (i) emphasizes a person (minimally, by a picture or a name); and/or (ii) a person's current activities; and/or (iii) professional experience; and/or (iv) displays a person's interest (in the body of the text and/or through hyperlinks to other sites)."

Such a definition is useful for this study in providing explicit points to look out for in terms of identifying a PHP (in a practical manner). However, there are problems with the type of sites which could be included as a PHP using this definition. For example, pages created about a person's current activities by a politically motivated group could be considered PHPs. Other definitions such as Dominick's (1999), which states that a PHP is "... a website that is typically maintained by one individual or family and contains....whatever information the author chooses to live there" are not as useful in providing clear signs to look out for when identifying PHPs. In fact, Dominick's definition also has imperfections. Again, a site created by a single person on behalf of Coca-Cola, would be considered a PHP, something which the author does not agree with. Doring's (2002) definition is perhaps the most accurate. It borrows Dominick's entire definition and adds that sites "maintained by organisations, institutions or formal groups are to be distinguished from personal home pages", something which fits in well with the concept of the PHP as a "personal" entity.

Other definitions such as Bates and Lu (1997) and the CWIS Style Guidelines (2002) can be overly specific and have no mention of the need for the PHP to be linked to the individual. Bates and Lu (1997) simply state that the PHP is the first page of a site found on a PHP provider (e.g. such as Yahoo! Geocities). This in itself is far from common as PHP definitions are generally in terms of sites or pages rather than a single page (e.g. Dominick 1999, Weaver 2000, Doring 2002). Meanwhile, the Campus-Wide Information Service (CWIS) website at Murdoch University (CWIS Style Guidelines 2002) offers a definition which is clearly from a technical, computing viewpoint:

"A HTML document prepared by an individual, that they describe as their Home Page. The document may be held on a personal computer at Murdoch, or made available from the individual's home directory area on a host computer at Murdoch, or made available from a computer accessible via the Internet on which the individual maintains a user's account. A Personal Home Page may include personal and biographical details and may offer a link to a document maintained for that individual on a CWIS Server. A Personal Home Page that is accessible via a CWIS Server will include a link to the Disclaimer Notice maintained on the CWIS Host Server. "

This is clearly not a definition from a theoretical academic point of view, but rather a practical definition to suit the purposes of a service providing institution. Definitions along these lines can be readily found on the sites of numerous PHP service providers, all over the web (e.g. NexBand Communications 2006, Rochester Institute of Technology 2007, Delaware Online 2007).

Author(s)	Assumed /Defined	Description
Buten	Assumed	Web page must either 1. Refer to itself as a PHP, or 2. Be listed under a personal name (e.g. Jack's Page)
Bates and Lu	Defined	The first screen of information that appears upon entering the URL address (drawn from the People Page Directory)
Campus-Wide Information Services (CWIS), Murdoch University	Defined	"A HTML document prepared by an individual, that they describe as their Home Page.... A Personal Home Page may include personal and biographical details and may offer a link to a document maintained for that individual on a CWIS Server..."
De Saint-Georges	Defined	A hypertextual document authored by one person which 1. Emphasises a person (minimally, by a picture or a name); and/or 2. A person's current activities; and/or 3. Professional experience; and/or 4. Displays a person's interest (in the body of the text and/or through hyperlinks to other sites)
Dillon and Gushrowski	Assumed	Web page is picked from a "home page resource" (e.g. the PeoplePlace)
Dominick	Defined	"... a website that is typically maintained by one individual or family and contains....whatever information the author chooses to live there."
Doring	Defined	1. PHP must be maintained by a single individual, regardless of affiliation with larger institutions 2. Allows for design to be delegated. 3. If the web pages represent more than a single person, it enters a specific sub-category of PHPs. 4. PHP must be distinguished from pages maintained by organisations (e.g. PepsiCo)
Haines	Assumed	Web page is picked from two directories of librarians' web pages
Papacharissi	Assumed	Web page is picked from a "personal home page provider" (e.g. AOL Hometown). Does not include pages that are "affiliated with or constructed by a commercial organisation or other institution"
Weaver	Defined	Librarian web pages which are "wholly under control of individuals, and not functioning as official library pages"

Table 2.1 Definitions or assumed definitions of PHPs

In terms of academia however, numerous studies have also been carried out on PHPs without explicitly defining them (e.g. Buten 1996, Bates and Lu 1997, Haines 1999, Dillon and Gushrowski 2000, Papacharissi 2002a, 2002b). Such studies normally just pick PHPs from personal web space providing companies such as PeoplePlace (Dillon and Gushrowski 2000, Papacharissi 2002a, 2002b, Barjak, Li and Thelwall 2007, Rick 2007), AOL Hometown (Papacharissi 2002a), the People Page Directory (Bates and Lu 1997) or other organisations where individuals might have PHPs such as universities (Buten 1996) or libraries (Haines 1999). Though this method of selection catches mostly PHPs (especially when measures are taken to ensure pages of businesses or other larger institutions are excluded), there is always the risk of misclassifying a non-PHP as a PHP. A clear definition and classification process certainly helps in ensuring that in such studies, only genuine PHPs are included, for the purpose of producing accurate results.

Hence, in this study, the definition of a PHP is the one offered by Doring (2002), using the criteria offered by de Saint Georges (1997). Though certainly an unorthodox approach (using a definition from one author and the criteria of another), it cannot be helped that de Saint Georges (whose definition has imperfections as mentioned earlier) has criteria that do indeed help identify genuine PHPs as defined by Doring, who does not offer a set of criteria to look out for.¹⁵

Hence, in order for PHPs to be included in this study, they have to either:

- (i) Claim to be a PHP
- and/or (ii) Have personal information on it (such as a CV, photograph etc.)
- and (iii) Represent a person rather than a group

2.3.2 Blogs (WebLogs) and Social Networking (SN) Sites

As mentioned earlier, and as evidenced by the definitions above, blogs and SN (social networking) sites also have a role to play in any all-encompassing view of PHPs. Though clearly not identical to each other, blogs and SN sites are very similar. In fact, the major social networking sites offer blogging tools as standard (e.g. MySpace.com,

¹⁵ Further clarifications on the process of PHP classification can be found in Chapter III (see 3.4.3 PHP Classification).

Facebook, Windows Live Spaces, Friendster, Yahoo! 360°) and most blogging sites offer the same tools offered by the social networking sites, such as the creation of personalised pages, uploading of photos and keeping lists of friends (e.g. blogger, xanga.com, blog.com). Meanwhile, in terms of information availability, because of this overlap in services, both blogs and social networking sites allow the same kind of information to be placed on their pages. For these reasons, whilst acknowledging that blogs and SN sites have their differences, this study will treat the blog and the SN sites as one.

As for definitions, the blog, much like the PHP, does not have a universally accepted definition (though most definitions of the blog fit into the PHP definition used in this thesis), there does appear less confusion about what constitutes a blog.

The online version of the Oxford English Dictionary has an entry for the “weblog”:¹⁶

"A frequently updated web site consisting of personal observations, excerpts from other sources, etc., typically run by a single person, and usually with hyperlinks to other sites; an online journal or diary."

Meanwhile, carrying out a "define: weblog" search on Google produced several results, the most comprehensive of which was from 5-Star Support (2004), a site which refers to itself as a "computer learning environment".

"(Same as blog) This is a publicly accessible personal journal for an individual. Similar to a personal diary, but shared over the web. The activity of updating a blog is 'blogging' and someone who keeps a blog is a 'blogger'. Blogs are typically updated daily using software that allows people with little or no technical background to update and maintain the blog. Postings on a blog are almost always arranged in chronological order with the most recent additions featured most prominently."

¹⁶ The OED has two definitions, the first being the technical meaning of a 'requests handled file' on a web server. Here, we mean the other PHP relevant definition.

Another definition can be found on a better known "computer learning" site. The WhatIs.com (2004) definition is:

"A weblog (sometimes shortened to blog or written as 'web log' or 'Weblog') is a website of personal or non-commercial origin that uses a dated log format that is updated on a daily or very frequent basis with new information about a particular subject or range of subjects. The information can be written by the site owner, gleaned from other Web sites or other sources, or contributed by users."

These definitions, though different, certainly have a common core. Meanwhile, in terms of social networking sites, somewhat unsurprisingly, once again, a common definition is not available. In fact, most SN site providers do not state what an SN site actually is. Friendster simply describes itself as the "best way to stay in touch with your friends and the fastest way to discover the people and things that matter to you most" (Friendster.com Site 2006). Its aim is to bring the power of social networking to every aspect of life and by doing so, make the world a smaller place (Friendster.com Site 2006). MySpace, simply offers 7 types of people for whom it is suitable for.¹⁷ A third major social networking site provider (Microsoft's own Windows Live Spaces) also simply states its aims (which are remarkably similar to MySpace and Friendster) and describes in detail the things one can do by using it,¹⁸ once again without an explicit definition of what a SN site is.

¹⁷ These include friends who want to talk online, single people who want to meet other singles, matchmakers who want to help their friends, families that want to keep in touch, business people interested in networking, class or study mates and finally, anyone looking for long lost friends.

¹⁸ In addition to the expected invitations to keep in touch with one's friends and make new friends, one can enjoy constantly updated contact lists, share photographs, write blogs, express themselves through over 100 themes, place fun gadgets on one's site and even stream your own videos (Windows Live Spaces Site 2006).

However, a definition has been offered, with Boyd and Ellison (2008) recently having proposed a comprehensive definition of the SN site. They defined the whole of an SN site (rather than an individual's page) as:

“web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system. The nature and nomenclature of these connections may vary from site to site.”

At the time of writing, it is uncertain as to whether this definition will be universally accepted, though with the definition focusing on the SN site provider rather than any individual's SN site, is only partially relevant. Regardless of this lack of universally accepted definitions (for SN sites and blogs), the aims of this thesis can be met by simply stating that a site which is potentially described as a blog or SN site can be included in the investigations carried out in this thesis as long as its description matches the definition provided for PHPs.

As for the typical characteristics and purpose of a blog, the most important feature often mentioned (also included in the definitions above), is the format of the entries, in reverse chronological order (Hourihan 2002, Nardi et al 2004b, Herring et al 2004) which is present in the three most popular types of blogs - filters, personal journals and notebooks (Blood 2002). These three types of blogs also go some way in describing their characteristics and indeed, their purpose. A filter blog is one that acts as a web filter for external content (i.e. non-personal information), whereas the personal journal is primarily concerned with the author's thoughts. The notebook is one with is concerned with both. The majority of blogs are said to be personal journal type blogs.

Concerning the purpose of blogs, a few reasons have been put forth, including to document one's life, as commentary, as catharsis, as muse and as a community forum (Nardi et al 2004b), to which others have been added such as improving writing skills, self-expression, to provide information, to pass the time and socializing (Li 2005).

2.4 The PHP as an Object of Study

As already mentioned, certain aspects of PHPs have been investigated in academic research. The majority of the research is either in terms of the general make-up and content (e.g. characteristics and design) or reflections that they have on "real life" (e.g. gender differences on the web, identity and self-presentation issues). Other studies have also been carried out, looking at the PHP in terms of its informational benefits (either looking directly at the information contained within them or as a portal to information). As often in other fields, studies occasionally fit into more than one category. All these have been looked at in the following sections of this chapter.

2.4.1 Characteristics of PHPs and Related Studies

The characteristics of PHPs have been looked at by numerous academics. One of the earliest studies on web pages (including PHPs) carried out in 1997 (Crowston and Williams), concluded that the PHP is a web page genre in its own right by looking at web page characteristics. It seems though, that this was something that had already been decided by some inside the research community (as the first PHP survey was carried out earlier by Buten in 1996) and others outside academia (Rubio 1996). Regardless, since around that time, numerous investigations have been carried out, looking specifically at the elements which commonly make up a PHP (de Saint Georges 1997, Bates and Lu 1997, Dominick 1999, Dillon and Gushrowski 2000, Papacharissi 2002a).

It appears that PHPs are made up of a number of elements such as the name of the author, email address, favourite websites, gender and photograph (Bates and Lu 1997), or title, email address, update date, table of contents and a creation date (Dillon and Gushrowski 2000), or a feedback mechanism, links to other sites, likes/dislikes (of the author), personal data and a photograph of the author (Dominick 1999). The mean number of pages which make up a PHP is 7, with 27.7% being personal, 33.8% being interests, 11.6% being on personal expression and 12.8% on families (Papacharissi 2002a). However, all these studies simply took pages from PHP service providers without an emphasis on any specific subject. Such a process invariably

excludes PHP authors who choose to buy their own domain names, who are potentially more technically advanced and have more sophisticated sites. Though possibly representative of PHPs in general, it is hard to use such figures in comparison to PHPs that focus on a specific subject (as is the case in this thesis). It must also be noted that some of the studies mentioned in this section were over 10 years old at the time of writing.

The characteristics described above were then used as a basis for studies looking at other aspects of PHPs such as gender differences (Dominick 1999, Arnold and Miller 2000, Flanagin and Metzger 2003), self-presentation and online identity (Wynn and Katz 1997, Dominick 1999, Chandler 1998, Nomura, Ishida and Yokozawa 2001), and even PHP design maps (Maruyama 1999, Ujigawa 1999).

Studies in gender differences have been carried out by examining generic PHPs created by men and women (Dominick 1999), including those created in a professional environment, for example on a university website (Arnold and Miller 2000) and even by fabricating PHPs for the specific purpose of deciphering the effects of sex (of author and user) on the perceptions of credibility of personal web pages (Flanagin and Metzger 2003). All sets of results suggest that gender differences are meaningful in cyberspace with women's PHPs trying to paint a picture which reduces the "real-life disadvantages of being a woman" (Arnold and Miller 2000), and with men and women dealing with different topics online, with more creative expression on female sites and more sports on male sites (Dominick 1999). It also turns out that both sexes find each other's sites more credible than their own (Flanagin and Metzger 2003).

Moving from gender differences to self presentation comes easily with studies that deal with both issues simultaneously (Dominick 1999), saying that strategies used in self-presentation on PHPs are the same as those used in interpersonal settings. As a tool for self presentation, the PHP is said to be well suited for giving a calculated and intricate impression (Wynn and Katz 1997, Chandler 1998). An opportunity has finally been given to the ordinary man or woman to have a publicly visible alter-ego as opposed to the personal, everyday, real-life portrayal of themselves. In fact, the PHP can even be used to supplement and complement the non-online image of

oneself, perhaps to people who are not yet fully familiar with the person in question (Doring 2002). Outside the everyday environment, one study has also looked at self-presentation in a professional setting, namely academia (Nomura, Ishida and Yokozawa 2001). There is no reason why the ability to portray oneself more should not be used professionally, where rewards are perhaps more measurable. To this effect, the results of the study showed that academics from different countries put emphasis on varying aspects of their professional qualifications and image, probably in tune with the specific country's culture as a whole.

2.4.2 Purpose of PHPs

The purpose of PHPs is something that has been looked at from various viewpoints by various studies (Buten 1996, Bates and Lu 1997, Chandler 1998, Dominick 1999, Maruyama 1999, Papacharissi 2002b).

Some have argued that the Internet has the ability to go some way in fulfilling the interpersonal needs of individuals (Morris and Ogan 1996) and indeed there are fans of the idea of the PHP supporting interpersonal communication (Dominick 1999, Papacharissi 2002b). To this end, two concrete lists have been produced. Whilst looking at the "self online", Papacharissi (2002b) gives 6 reasons for the creation of PHPs. These include passing time, entertainment, information, self expression, professional advancement and communication with friends and family, whereas according to Maruyama (1999) the possible five motives for creating a PHP are said to include "self-discovery/self-expression, mental exchange with visitors, making easy-going and delightful communication, editing and publishing information and the challenge to something new".

Of those, self expression is particularly popular among researchers with Buten (1996) Wynn and Katz (1997), Chandler (1998), Maruyama (1999) and Dominick (1999) all citing it as a reason. The publishing of information is also popular with Buten (1996), Bates and Lu (1997) and Maruyama (1999) citing it. One missing major reason seems to be understanding (Bates and Lu 1997) or improving (Buten 1996) one's web skills,

though perhaps this was included in Papacharissi's "professional advancement" category.

Finally, one study created a design map to let PHP authors design their sites according to their intentions by providing them with differing templates matching the differing needs of authors (Ujigawa 1999).

2.4.3 Accessibility of PHPs

When it comes to PHP accessibility, whereas a URL will lead to its destination, most PHPs are still hard to find (Narsesian 2004). Carrying out subject specific searches will not necessarily reveal relevant PHPs. Major search engines will often have the PHPs towards the end of any results. Of course, this depends, to a certain extent, on the subject being searched for. Invariably though, popular subjects (including professional football in England) tend to be financially profitable and by extension have numerous commercially purposed non-PHP providers of information. Subjects which are not popular are more likely to have PHPs closer to the top of any search results (often because there are not that many non-PHP providers of that information).

In an attempt to improve the accessibility of such sites, some research has been carried out in relevant areas. To date, two known attempts have been made at creating search engines which utilise primarily PHPs (Shakes, Langheinrich and Etzioni 1997, Hoff and Mundhenk 2001). The first of these created in 1996 was called Ahoy! The Homepage Finder.¹⁹ The purpose of Ahoy! was to fill the gap between manually created directories (e.g. Open Directory) and robot-generated web-indexes (e.g. Lycos). Essentially, a name of a PHP author is entered, and then Ahoy! uses search engines to find possible matches. At the same time, Ahoy! checks for the name of the PHP author in email directories and an internal database. According to the creators, the end result is that Ahoy! can find matches from other search engines, but also guess the author's PHP address. In fact, at the time Shakes, Langheinrich and Etzioni claimed that in their experiment Ahoy! worked better than Altavista. Another PHP search engine was later created with the specific intention of looking for scientific

¹⁹ It must be noted that Ahoy! is no longer functional (even though the web page still exists), and as such, the accessibility of PHPs in general at the time of writing is not affected.

papers. Hoff and Mundhenk (2001) used two systems, HomePageSearch and MOPS²⁰ in a three-step combination method for the purpose of facilitating the search for scientific papers. Unfortunately though, none of these search tools made a significant impact in the popularity or accessibility of PHPs. In fact, at the time of writing both Ahoy! and HomePageSearch have ceased operating (at least publicly).

Meanwhile, one important point is that none of the PHP search engines carried out searches using the subject. As a result, even if they had both been successfully implemented and active to this day, it would still have been questionable as to whether or not they improved access to PHPs in terms of subject specific searches.

2.4.4 PHP Numbers

It is very hard to tell exactly how many PHPs exist on the World Wide Web. Papacharissi (2002b) points out that there is no directory of personal homepages (something which still holds true at the time of writing) and none of the PHP studies come across provide definitive numbers. Earlier studies have provided some figures, and these point to a significant number. In 1996, Buten estimated that there were 600,000 PHPs in the United States alone. A few years later, in a study conducted by Lawrence and Lee Giles (1999), there were 800 million publicly indexable web pages of which 2% were personal homepages, putting the number of PHPs at just over 16 million. In December 2000, Wired magazine reported that Geocities alone had 5.5 million members.

Other indicators also point to the existence of many PHPs. The Online Computer Library Centre Web Characterisation Project (2002) has a list of the "Top 50 most frequently linked" websites in 2002. In that Top 50, there were 6 sites which are primarily PHP hosting sites, two of which were in the top 10.

²⁰ According to Hoff and Mundhenk (2001), MOPS is a search engine. However, in their research paper, they have not stated what the acronym stands for. As the search engine is no longer available online, it was not possible for the author to independently determine words in the acronym.

Name	Rank
Geocities	3
AOL (Hometown)	5
Angelfire	12
Tripod	13
Earthlink	34
XOOM.COM	35

Table 2.2 Personal Home Page Providers within the "Top 50 most frequently linked" websites in 2002

However, in the fast moving world of the web, these numbers have to be somehow refreshed. Since 2002, numbers of PHPs have not been estimated, perhaps due to difficulties caused by the mergers of many of the PHP providers into other sites (e.g. Geocities becoming part of Yahoo!). However, if we use figures of blogs or SN sites as a guideline to how many PHPs might exist, the numbers have only appeared to grow. Several estimates are in the tens of millions of sites, with one estimate putting the number of blogs at 70 million (Riley 2005), and another at 35 million (Sifry 2006). In terms of SN sites, MySpace alone is said to have 48 million users (Pickavet 2006). Of course, all the blogs and SN sites might not fit into the definition of the PHP used in this thesis. However, it would not be at all unreasonable to assume that the number of PHPs in existence is at least in the millions.

2.4.5 PHP Problems

There is no question that there are lots of negative and problematic aspects of the PHP (particularly as an information resource). For instance, as already mentioned, there is no absolute directory of PHPs (Papacharissi 2002b), and no guarantee that once a PHP is created, it will be accessible to web users who do not have its URL (Lawrence and Giles 1999). Search engines do not cover 100% of the web (Lawrence and Giles 1999), and even when the PHP is covered by the search engine, it is entirely possible that the search link retrieved is so far down the list, that people seldom (if ever) click on it. The truth is that though the web can theoretically be a tool for mass media communication, a PHP does not guarantee that one reaches more people than they would by putting an advertisement in the newspaper.

Furthermore, there are problems with the providers of PHP services (e.g. AOL Hometown, Yahoo! Geocities etc.). Firstly, there are no guarantees that the PHP provider will remain online forever (e.g. the People Place no longer exists). Like in any other business, PHP service providers are subject to market forces. Meanwhile, the accessibility of a PHP also depends, to a certain extent, on the ability of the provider to market itself (in addition to staying in business). The quality of the PHP can also depend on the provider, in terms of bandwidth (i.e. how quickly the pages from one's PHP can be downloaded and viewed), aesthetics (i.e. some providers insist on placing ads on all PHPs) and functionality (i.e. providers might not allow or facilitate the use of all available software).

In addition to the issues mentioned above, PHPs also suffer from negative perceptions, something that has been acknowledged in other studies (e.g. Dominick 1999) but can also be seen online. PHP authors themselves have been critical, going so far as to say that a PHP is a PHP simply because it has been created "in bad taste, with lots of colour clashes, a self-promoting style of writing (often entirely in the first person), with numerous grammatical mistakes, an unusually long URL, dead links, poor graphics and an amateur photograph" (Iron 2003). Though clearly not a well thought out definition, such negative perceptions could very well have played an active role in discouraging researchers from looking into the PHP in more detail.

Also in addition, the variables mentioned so far are potential problems that exist for PHP authors which are out of their control. Authors however, create a whole spate of problems (in terms of information resources) for themselves. Authors often do not finish the construction of their sites (Doring 2002, Papacharissi 2002b), often do not update the site for years (Papacharissi 2002b, Narsesian and Nicholas 2005) and could potentially create sites with misleading, if not totally incorrect information.

All these factors combine to put into place a very dynamic and uncontrolled cluster of information, which becomes very difficult to examine and facilitate the use of. However, this should not be a deterrent for researchers, particularly when the potential benefits are so great.

2.4.6 PHPs and Grey Literature

Materials treated as grey literature (also called non-conventional literature) are materials not identifiable through a traditional index, database or other publication channel. These are said to include scientific and technical reports, government reports, internal company or organisation reports, newsletters, conference proceedings, data compilations and trade literature (Auger 1998, Jeffery 2000, McKimmie and Szurmak 2002). With the advent of the web, the definition still appears to hold true, though investigations have begun to look at the effect of the web on the definition and concept of grey literature (Vilma and De-Castro 2001, Pace 2002, Thompson and Guistini 2006). One result of these investigations is the addition of certain types of pages and other electronic entities to the list of "grey matter", including pre-prints, FAQs, e-zines, listserv archives, digital libraries, audio/video over the web, portals and even blogs (Thompson and Guistini 2006), with others likening the concept of the web as a whole, with its ephemeral nature and disappearing pages, to grey literature (Pace 2002).

As an information resource, the PHP clearly bears some similarity, conceptually, to grey literature. One of the major difficulties afflicting grey literature, and perhaps the primary reason why it has been termed separately, is the fact that though in conventional literature well-established systems of bibliographic control are present, grey literature is still difficult to identify and make accessible (Auger 1998). The web as a whole, and PHPs in particular, both suffer from this lack of bibliographic control, even though Google and other search engines are said to have made big steps towards some sort of organisation of the web (Pace 2002, Thompson and Guistini 2006).

The similarities do not end there. Claims have been made that grey literature too should not be paid too much attention, stating simply that if grey literature was meant to have longevity, then it would have been published through traditional publication channels and therefore not be grey literature (Auger 1998). However, references to grey literature have been made repeatedly in academic publications over the years and this surely demonstrates value. This value has been demonstrated even more clearly in recent years with the emergence of databases devoted specifically to grey literature,

something which also shows one more similarity to PHPs, which have had search engines devoted to them (e.g. HomePageSearch, Ahoy!).

Other similarities include complaints made about grey literature concerning the lack of conformity to traditional publication standards (Auger 1998) and again here, the PHP has had similar complaints made about it (Dominick 1999, Iron 2003, Narsesian 2004). However, both grey literature (in general) and PHPs do not often conform to such standards as they were initially created with a specific audience in mind, not necessarily the same as the final actual audience. Problems arise when this unexpected audience discovers the potential value in this literature.

However, there are some differences too. Whereas grey literature is potentially unidentifiable (i.e. it might not necessarily have a unique identifier of any sort), and by extension, potentially unattainable, PHPs have their URLs which provide both a unique identifier and access. Needless to say, this does not guarantee permanence, and a PHP might disappear at any time, but if anything, this again bears resemblance to grey literature.

As for similarities with this study, ideas and concepts in grey literature do occasionally overlap with ideas presented here relating to PHPs. A suggestion made in this thesis is that the information found on PHPs could be filling in the gaps of information left by other providers, particularly in subjects where traditional suppliers of information are not plentiful. Such ideas have been presented in grey literature studies (McKimmie and Szurmak 2002) as reasons for the importance of studying the subject. Certainly, if it holds true for grey literature, it should hold true for PHPs, especially for those who would accept the PHP as grey literature.

2.4.7 Virtual Communities, Communities of Practice and Knowledge Sharing

In terms of the ideas presented in this thesis concerning PHPs and any benefits they might bring, it is important to remember that there is often an attached community element amongst the authors, particularly concerning blogs or web-rings (a web-ring

is a collection of websites created by a central site that accepts registration of one's PHP in order to group it with similar PHPs²¹). As such there are some related subject areas about groups of people who also work together for the greater good. This section looks at the PHP from a wider angle, looking at the relationships between the PHP as an information resource, various relevant communities and by extension the dissemination of information.

Virtual communities (also known as online communities) have been associated with information seeking (Rheingold 1992), football fans (e.g. Joga.com) and even with PHPs (e.g. web-rings and SN sites). As humans have always been socially active, they have had to form new social groups for various reasons, including protection and pleasure (Klang and Olsson 1999). With widespread access to the Internet, it was no surprise to see this phenomenon move to the virtual environment (Apgar 1998). As for a description, a virtual community is said to be a group of people who communicate and exchange ideas via computers and the Internet. The members of these communities may reside thousands of miles apart, and may or may never meet in person (Rheingold 1992, 2000). Differences between virtual communities and traditional communities include physical distances and unusual attributes, such as the ability to get to know someone before actually meeting them and the ease with which one can get useful resources from "weak ties" (Constant, Sproull and Kiesler 1996). Similarities include the "real" effects that a virtual community can have on one's offline life, as well as going through "long-simmering family brawls" online (Rheingold 1992). Other similarities include the ability to get all kinds of support on virtual communities, provided people know where to look, to reciprocally exchange resources and the ability to maintain strong ties (Wellman and Gulia 1999).

Regardless of these similarities and differences, it seems that these communities are first of all a social entity (Croon and Argen 1998, as cited in Klang and Olsson 1999), originally born of the need by people to discover other like-minded people, so as to participate in "wide-ranging, intellectually stimulating, professionally rewarding, sometimes painful, and often intensely emotional ongoing interchange with dozens of new friends, hundreds of colleagues, thousands of acquaintances" (Rheingold 1992).

²¹ Registrations for web-rings can be made on webring.com.

These virtual communities now play a big role in many people's lives, with individuals spending seemingly endless hours in front of the computer on such communities (Wellman and Gulia 1999), with some even said to be addicted (Barlow 1995, Kling 1996). With such endless hours spent on the virtual community, it is no surprise that information exchange, and by extension, knowledge sharing, is present.

A semi-related phenomenon to that of virtual communities, is that of Communities of Practice (COP). The reason for the use of the term semi-related, is that though COPs exist within virtual communities, they have always existed outside them as well. The phrase was first coined by Lave and Wenger (1991) who describe a COP as a community of people informally bound by what they do together (e.g. members of a band or a group of consultants). One can be in the core of such a community or on the periphery (it is also possible to not have an established role). There are three basic dimensions when defining a COP, which are seen as separate to formal organisational units which are fully dynamic and always evolving (with members coming and going). These are: what it is about, how it functions and what capability it produces.

At the peak of their activity in a COP, the members are said to engage in joint activities, creating artefacts, adapting to changing circumstances, renewing interest, commitment and relationships (Wenger 1998). Studies have now been carried out looking at the role of the COP within virtual communities (Dube, Bourhis and Jacob 2006, Carotenuto et al 1999), now sometimes called Virtual Communities of Practice (VCOP).

Members of VCOPs use a wide range of media including more traditional household communication devices such as the telephone and fax, and slightly more technologically advanced methods such as email, newsgroups, websites and databases (Dube, Bourhis and Jacob 2006). However, there is a problem of finding related research from the viewpoint of this thesis, in that COPs and VCOPs for the greatest part are looked at from business or organisational perspectives, rarely examining instances in an ordinary non-professional environment (such as that of football fans). Such a problem also exists in terms of knowledge sharing, a term also strongly linked to business, often used in conjunction with COPs and VCOPs (e.g. McLure Wasko and Faraj 2000, Sharatt and Usoro 2003, Dube, Bourhis and Jacob 2006). However,

all the studies mentioned, and indeed many others, are based on the existence of the idea that information is shared online by individuals in various environments (e.g. business) and communities (e.g. COPs), resulting in a net benefit for all those involved.

Going back to generic virtual environments, to date, empirical studies concerning the numbers of users of virtual communities have been scarce, however, with numerous studies mentioning the benefits of information found on such communities for non-business purposes (e.g. Rheingold 1992, Wellman and Gulia 1999, Kollock 1999, McLure Wasko and Faraj 2000) as well as more organised business or institutional purposes (Davenport 1996, Ellis 2001, Haimilia 2001), and taking into account the number of PHPs, there can be little doubt that the benefits created are potentially significant.

More specifically, concerning the value of virtual communities in terms of providing information resources, references have been made about what it is possible to get out of them and how quickly. For example, Rheingold (2000) states that he was able to collect 200 pages of expert advice on the issue of "Communication Systems for an Information Age", by spending a few minutes a day for six weeks. Other have talked about the usefulness of online reciprocity in terms of information exchange where a gift economy is said to exist (Kollock 1999, Rheingold 2000) and distant strangers give useful advice to other distant strangers, even when there are no strong ties present (Constant, Sproull and Kiesler 1996). A virtual community alone can be a very useful source of information and their importance practically, scholarly and politically has been stated (Wellman and Gulia 1999, Rheingold 2000).

As for the relationship between virtual communities and PHPs, though it is not one that has been looked at extensively, a link has been established in that PHPs are said to be part of a collection of tools (which also include chat tools, email, email lists, databases and other virtual reality applications) used to establish communities online (Klang and Olsson 1999). Additional evidence that there is interaction between PHPs and virtual communities comes in the form of web-rings and perhaps more strongly blogs or SN sites. Both blog writers and SN site owners can link to one another, comment on each other's sites and collaborate. Safran and Kappe (2008) analysed the

number of textual entries, the number of images, comments given, comments received, guestbook entries given and guestbook entries received on the blogspace of the *Kleine Zeitung* with the aim of analysing key factors for the “success” of weblogs. They concluded that the most important of these factors were the community activities of the authors, i.e. writing comments and guestbook entries in other blogs.

Interaction between bloggers has also been identified as one of the principal reasons for blogging (Herring et al 2004, Nardi et al 2004b). In these terms, one potentially relevant entity to this thesis is Joga.com. In 2006, NikeFootball.com and Google.com combined forces in order to create an online social networking site with football as the main theme called Joga.com (Nomensa.com Site 2006). The site offers all the services offered by ordinary social networking sites, such as the creation of personalised pages, uploading of photos and blogging tools, but also has some football related features such as a "starting eleven" section and pages about Nike sponsored football players.

2.4.8 PHPs as Sources of Information

Studying PHPs as sources of information has not been a popular custom to those in academia. Though some studies have looked at this issue directly (Narsesian 2004, Narsesian and Nicholas 2005, Paquet 2003, Thelwall and Harries 2002), most studies that could be deemed relevant have had their main focus on other subjects, for example librarianship (e.g. Haines 1999, Weaver 2000). Regardless, it is fortunate that some studies which have carried out surveys without the express intention of investigating the PHP as an information resource have produced interesting results. This in turn has produced a foundation on which it has been possible to build the idea of the PHP as a source of information.

The PHP, as already mentioned, is said to be a separate communication genre (Dillon and Gushrowski 2000), and as such is an important way to communicate ideas socially. There has also been some mention of its importance for cybercommunities and cyberenthusiasts (Rubio 1996), something which is even truer today with the popularity of blogs and SN sites (Paquet 2003). The PHP gives one the ability to

make a statement of almost any size, and keep it online for as long as they wish. Even if a fraction of all these statements made are factually informative, taking into account the vast numbers of PHPs available, one must conclude that the potential for a vast repository of information is real and present.

As for actual figures from previous studies, in terms of individuals placing information online, Buten's (1996) survey states that the third most popular reason for creating a PHP is to "distribute information to friends/people" (43% of respondents). Essentially, almost one in two people use (or intend to use) PHPs to distribute information. Other studies also support this notion with a significant percentage of PHPs created with a focus on the general interests of their authors (33.8% of respondents, Papacharissi 2002b). Another study found that 75% of PHPs contained information about either the "likes" or "dislikes" of their authors (Dominick 1999). Breaking that down, it was discovered that 44% (of the 75%) of pages contained information about hobbies, followed by music at 34%, sports at 20%, arts at 11% and books at 8%. Other indicators are also available in terms of what people write about on their PHPs and their expected audiences. According to Buten (1996), 50% of authors write about entertainment, 23% about research and 19% about sports. In terms of expected audiences, significant numbers of respondents thought that "fellow enthusiasts for a topic/hobby" would visit their PHP (61% in the generic sample, 42% in the educational sample) indicating that authors have an intention to inform fellow surfers.

Moving away from simple figures, various types of studies have implied that there is potentially valuable information found on PHPs. Studies have looked at university PHPs so as to examine the flow of information from universities to the community in general (Thelwall and Harries 2004), the factors that influence the web impact of scientists' PHPs (Barjak, Li and Thelwall, 2006), and librarians' PHPs to see how they help the role played by the library in facilitating the use of the web (Haines 1999, Weaver 2000). These already implied that something of value might be found on PHPs before the direct investigation of the PHP as an information resource (Narsesian 2004, Narsesian and Nicholas 2005) which suggested that significant amounts of unique information could be available on PHPs. The first such suggestion was touted using anecdotal evidence (Narsesian 2004), and then a preliminary study (Narsesian

and Nicholas 2005). The idea suggested was quite simple, and conceptually linked to the beneficial values of grey literature. People with access to the web might, for a variety of reasons, want to provide information on their own sites. As we have already seen, PHP authors have said they want to distribute information to friends/people (Buten 1996) and it appears that by placing your likes and dislikes on one's PHP (Dominick 1999), one ends up talking about topics of interest to others as well (Buten 1996). In addition to this, it is entirely possible with the general accessibility of the web (to the majority of the western world), that one might want to create a PHP with the express purpose of disseminating information.

If all these ideas are true, then it is logical to hypothesize that PHPs will offer some advantages in providing certain types of information, something which has already been done (Narsesian 2004). The strength of the PHP might lie in providing information where commercial providers or other organisations might not want to venture for any reason (Narsesian 2004), thereby filling the gap in the informational spectrum, in a way which relates to grey literature ideas (McKimmie and Szurmak 2002). Whereas established information providers (e.g. the BBC) will have up to the minute information, they might not have the information in great detail. The strength of the PHP is more likely to be in providing relatively static, in-depth information. More recently, other developments have taken place with projects such as wikipedia.org having gone some way into providing static, in-depth information (Narsesian and Nicholas 2005). Regardless, there will always be gaps left by any information provider and it is those gaps that can potentially be filled by PHPs. In addition, one more possible advantage might be the ability to contact the author. Currently, it is hard to say if it is easier to contact PHP authors, with little research carried out to date specifically on this issue. However, 58% of PHP authors have said that they receive an email once a week concerning their pages (Buten 1996). This is especially true of blogs, where authors have been said to be actively seeking others' opinions and feedback (Nardi et al 2004a).

With all this known about PHPs, it makes sense to finally examine the role of the PHP within a certain environment (e.g. professional football in England), taking into account its competitors and the needs of the information seekers. Certainly, in the online football environment, there is evidence of in-depth information on PHPs

(Narsesian and Nicholas 2005), with dedicated fans having created websites with plentiful informational content. As such, the following sections of this chapter deal with information seeking and the background to football.

2.5 Information Seeking

Information seeking as an area of research is not particularly new. Many studies have been carried out looking at information seeking behaviour from numerous viewpoints such as professional managerial (e.g. Calantone, Gudmundson and Lang 1997), librarian (e.g. Schreiber and Moring 1997), medical (e.g. Joinson and Banyard 2002), journalistic (e.g. Dominick 1999), legal (e.g. Wilkinson 2001) and so on. The football fan's information seeking viewpoint however, has not been a particularly popular theme. Fortunately though, there have been some studies looking at non-professional and non-scholarly subjects, and these are also discussed. In addition, several of the relevant studies look at information needs (as well as information seeking) and those too have been included.

2.5.1 Information Seeking Models

A central aspect of information seeking studies is information seeking models. Currently, there are several information seeking models, which can be applied in both online and offline environments (Ellis 1989, Kuhlthau 1991, Belkin 1993, Wilson 1999, Maurer 2006). Here, an overview of each model is given.

Ellis (1989) has an information seeking model, which can also be applied to hypertext environments (such as the World Wide Web). This model initially consisted of six categories which were starting, chaining, browsing, differentiating, monitoring and extracting. Later, two more categories were added called verifying and ending.

Starting is, as the name would suggest, the starting point in the model. At the beginning of the information seeking cycle, the individual needs to take a first step, which could be asking a friend or colleague a relevant question, or typing in a search term into Google. Chaining is the next step in the process, where the information

seeker will take the next natural step, using connections from the first step to get closer to the information being sought. There are two types of chaining: forward chaining and backward chaining. The first type is the process of seeking new sources that refer to the original source. The second type, backward chaining, is the act of pursuing a pointer or reference from the original source. Hence following a reference at the end of an article is backward chaining, whereas using a citation index to find new works citing the original article is forward chaining. The third category is browsing, which is the act of offhand searching in areas of potential interest, such as looking through lists of books or articles in a certain field. This is followed by differentiating, which is the selection of particularly relevant works according to certain essential criteria. The fifth category is monitoring, where the information seeker is now keeping up-to-date with the state of the art by keeping a check on certain sources. Finally, the last category is extracting, where a persistent evaluation and re-evaluation of sources produces the desired resource of information. Verifying (Ellis 1991), as the name suggests is the process of checking the information and ending (Ellis 1991) is the final step in the overall model, where the final organisation of the information resource is completed.

Kuhlthau's model (1991) also has six stages, this time with a strong link to the user's emotions. The six stages are initiation, selection, exploration, formulation, collection and presentation. The stages have some resemblance to Ellis' model and this is something that has been noticed already (Wilson 1999). Initiation is the beginning of the whole process, where the ideas are still general, and the feelings are of apprehension and uncertainty. Selection is one step ahead, the time to choose the general topic, and the feelings are of confusion and anxiety. Exploration is the process of investigating with the intent of finding a focus. The feelings here are still of confusion and now also doubt. Then comes formulation, focusing on the information now, where there is a sense of confidence and optimism. This is followed by collection where further relevant information is gathered and there is now an increased interest in the topic, with feelings of increased confidence. Finally comes the presentation stage where the overall process is complete, and the associated feelings are often ones of relief.

Wilson (1997) has also created a general model of information seeking behaviour (Figure 2.1) in order to improve the process of designing an information system. The model of information seeking behaviour though is generic, and can be applied regardless of the discipline or subject area. The information user develops a need, which initiates information seeking behaviour. This puts demands on information sources and information systems. If this leads to success, the information is used, and either by satisfaction or non-satisfaction, a new information need is created, which restarts the cycle. The model has other aspects too, of information transfer or exchange with other people depending either on information use, or information seeking behaviour.

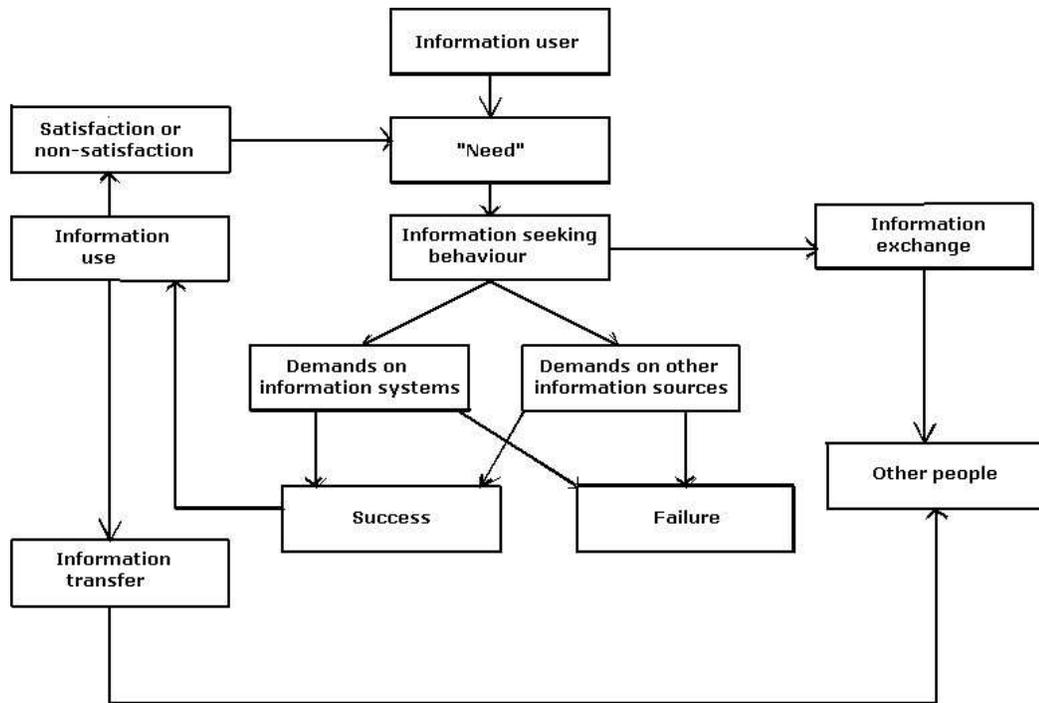


Figure 2.1 Wilson's model of Information Seeking Behaviour

Another author who writes about information seeking with the intention of aiding the design of sites on the web is Maurer (2006). A model as such has not been created and the paper does not target a specific group of people, however a form of information seeking that applies to all groups of people is discussed. The four modes are known-item, exploratory, re-finding and "don't know what you need to know". In each one of these modes, the user state is slightly different and each can be identified in various

ways (see Table 2.3). The suggestion is that designers of sites must understand that searches can be carried out in more than one manner, so as to be able to cater for as much of the website audience as possible.

Mode	User state	How to identify
Known-item	The user knows what they want, what words to describe it, and where to start	Heavy use of accurate keywords
Exploratory	The user knows what they want, may or may not be able to articulate it, may or may not know where to start	Vague phrases or repeated searches with similar keywords
Re-finding	The user is carrying out a search they have carried out before. The user knows what they want, what words to describe it, and where to start	Cookies, site registration or logs must be used
Don't know what you need to know	The user is not certain of what information is necessary (e.g. they are trying to find a solution to a problem)	N/A

Table 2.3 Mode, user state and identification according to Maurer (2006)

Ellis (1989), Kuhlthau (1991), Wilson (1997) and Maurer (2006) all look into the actual process of information seeking. Though the processes described have been studied and are relevant (as the football fans themselves go through the steps described in all the studies), this study is more concerned about the tools and methods that are used, and the reasons for their use. This section has been included to illustrate that these models have been taken into consideration throughout the thesis.

2.5.2 Environments of Information Seeking

Studies such as this one, which look at groups of non-scholarly and non-professional people in specific informational environments, though available, are far from plentiful (Rieh 2004). In fact, 95% of studies conducted on information seeking and use from 1966 to 1994 (at which time this finding was published) have focused on scholarly or professional environments (Harris and Dewdney 1994), thereby largely neglecting the ordinary citizen, and by extension the football fan. However, several such studies have been found (the majority of which were published after 1994) including two which examine the information seeking behaviour of football fans (Green 1999, Joinson and Banyard 2002) and non-football related ones (Marcella and Baxter 1999, Rieh 2004, Nicholas et al 2003), including studies in a concept known as "serious

leisure" (Stebbins 1982, Buchanan 1985, Spink and Ozmutlu 2001, Hartel 2003, Yakel 2004, Cox, Clough and Marlow 2008).

Looking at the information seeking behaviour of football fans is rare, though as mentioned two have been carried out. The first study (Green 1999) examines the kind of information services provided by football clubs and the kind of information services used by fans. Though conceptually firm, the study is relatively small in size, with few questionnaires sent out and returned. The second study (Joinson and Banyard 2002) though related to football and information seeking, is carried out from a different perspective, looking at the differences in information seeking before and after a match, to decipher the effect of a victory, draw or loss on information seeking.

Other information studies looking at non-academic and non-professional studies have looked at various populations to find out how people look for information concerning a specific subject, for example citizenship issues (Marcella and Baxter 1999). Such studies look at tools and methods available for the subject in question (e.g. Citizen's Advice Bureaus and public libraries). Studies have also been carried out looking at information hubs used by ordinary citizens, such as electronic stand-alone health kiosks (Huntington et al 2002) or the new digital interactive environment in the modern home (Nicholas et al 2003).

Meanwhile, born from the ideas of serious leisure²² (Stebbins 1982, Buchanan 1985), some related studies have been carried out within non-professional "leisure" information environments such as sex (Spink and Ozmutlu 2001), gourmet cooking (Hartel 2003), genealogy (Yakel 2004) and photography (Cox, Clough and Marlow, 2008). The relevance of this work lies in the fact that all forms of leisure require, to varying degrees, a certain amount of information seeking and use. Meanwhile, the Internet is often a resource for entertainment purposes (Pew 2005). Furthermore, leisure was said to be increasingly becoming a significant aspect of everyday life in 1996 (Argyle 1996), and the amount of research into leisure based information seeking appears to have increased since then. All these are concepts which fit in

²² Stebbins (1982) has defined serious leisure as the systematic pursuit of an amateur, hobbyist or volunteer core activity that is highly substantial, interesting, and fulfilling where in the typical case, participants find a career in acquiring and expressing a combination of its special skills, knowledge and experience.

nicely with the idea of the hardcore football fan, who finds plenty of time to spend on his favourite hobby on various media (one of which is the Internet).

To illustrate more clearly the kind of research involved in the examination of information seeking in serious leisure, one such study is looked at here. Hartel (2006) looked at the information activities, resources and spaces of gourmet cooking at home. Twenty gourmet cooking enthusiasts from various parts of the United States found by purposive and snowball sampling, took part in the study. Fieldwork was carried out in the cook's residence, where semi-structured sixty minute interviews were carried out, exploring the context of the hobby, the routines and informational elements. These interviews were recorded and transcribed. Following the interview, the cook showed and described the cooking site to the researcher who took pictures which were then attached to the cook's commentary. Using inductive and deductive analysis the interview transcripts and field notes were studied, and a visual analysis process was used on the photographs. The result is that hands-on gourmet cooking takes nine steps to complete, in what is called a "nine step episode". Information activities and resources are essential, and scattered throughout the process. The primary information activities are use and re-use, with the primary information document being the recipe.

2.6 Football Studies

The field for the case study of this work is that of the top four professional football leagues in England (which at the time of writing also include three Welsh teams). The great social and commercial interests that lie within this field imply that plentiful football studies have been carried out, and so they have. They cover numerous aspects of football, such as hooliganism (Dunning et al 1984, Williams 1988), identity (Giulianotti and Williams 1994), racism (Bains and Patel 1996, Bradbury 2001), football finances (Michie and Oughton 2004, Holt et al 2005), fanzines (Hall and Smith 1997, Haynes 1995) and social and community aspects (Jaquiss 2003, Tacon 2005). Meanwhile, the two major publications in the English footballing world are the FA Premier League National Fan Survey and the Deloitte and Touche Football Money League. In terms of relevance, this thesis is concerned with the information

behaviour of football fans and by extension the background to football (so as to have a better understanding of the subject area). This is doubly important in this instance as a football survey was carried out as part of this thesis. As such, various relevant football surveys have been looked at in some detail, on a survey by survey basis, followed by works relevant to football in the community.

2.6.1 Football Surveys

Perhaps the most notable of these is the FA Premier League National Fan Survey (2005). As part of this study, 80,000 questionnaires were sent out with 25,539 responses received. The survey includes 39 questions covering 12 sections (see Table 2.4). The topics covered include attendances, the matchday experience, the services and facilities offered by clubs, television, Internet, football spending, the football community and fan profiles. The objective of the survey is to explore attitudes and opinions of football fans, with the purpose of providing a report to the Premier League football clubs from a customer service perspective. The report reveals a number of interesting facts about fans.

Section Name
Patterns of Attendance
Attending Home League Matches
The Matchday Experience
You as a Fan
Club Service and Facilities
Television and the Internet
Football Spending
Football and the Community
The FA Premier League
Title Sponsor
Questions Specific to the Club
Personal Details

Table 2.4 The 12 sections of the FA Premier League's National Fan Survey Questionnaire (2005).

As the research method was self selection, the results do suffer from some polarisation, however, they are still indicative. According to the 2005 survey, the Premier League fan is predominantly a 44-year old male (85%), who is white (96%)

and middle class (71%). According to the report, the most important issue for fans of all clubs was safety, followed by clear sight lines (at the football grounds) and the match atmosphere.

In terms of the Internet, which the report also studied, some interesting questions were asked. The percentage of football fans who have access to the Internet is 87%, with a Great Britain (generic) adult average of 58%. The report comments that the Internet is increasingly seen as a valuable source of information for fans. Nearly two in five supporters access the Internet at least once a day for football information. Another interesting fact is that a group of alternative, non-club sites are ready to provide access to football "gossip" and "unofficial news". However, "official" news is still the more popular, with 88% of supporters with Internet access using the official sites at least once a month.

A similar study was also carried out by the Football League (2006) which consisted of the three divisions below the Premier League. The Football League sent out 450,000 emails to registered supporters, and 43,590 responses were received, giving a response rate of just under 10%, and making it the largest survey of its kind ever conducted (in 2006). Once again, the method was self selection, and so some types of fans might be underrepresented. The questionnaire contained questions concerning fan demographics, the image of the Football League, the governance of the game (i.e. the league system, the management of the teams), the refereeing, various aspects of the matchday experience (e.g. travel, ticketing, smoking), commercial and customer service issues and the involvement of the league in the community. The survey did touch on tools and methods used by football fans to acquire information, though not in any great detail. According to this survey, 93% of fans use the official websites of their clubs.

Towards the end of the completion of the work carried out for this thesis, both the Premier League (2008) and Football League (2008) had carried out and published newer fan surveys. Both of these were examined by the author and after careful consideration, it was decided to use the previous versions of the surveys for the purposes of this work for three reasons of equal importance.

Firstly, the Premier League and Football League surveys have little overlap with the surveys carried out for the purposes of this thesis (they have been included here primarily for the purposes of verifying the demographics of the respondents for this work). Secondly, the results of the previous Premier League (2005) and Football League (2006) surveys are more appropriate for the purposes of this work, as the time the research was carried out for this project is closer to time the earlier league surveys were carried out. Thirdly, the results of the later surveys (Premier League and Football League) have not greatly changed, particularly in terms of demographics.

Another notable survey was carried out by The Sunday Mirror in 2001. In this instance there were 3,982 respondents (though the original number of questionnaires distributed has not been reported), with the research method, once again, being self selection. The average age of the respondent was 36. The questions covered such topics as fan profile, football expenditure, attendances, racism and a number of contemporary (at the time), specific issues, such as David Beckham's England captaincy, and whether England would qualify for the 2002 World Cup. Some interesting points were revealed, such as the fact that 92% of fans agreed or strongly agreed that supporters and shareholders should have more say in the running of their clubs, the fact that more fans felt that their club's communication with fans was below par (than those who felt it was above par), and 76% of fans agreed or strongly agreed that top English football clubs did not do enough to sign local talent.

Perhaps the most relevant study was carried out by Green (1999) who looked at football information services. Interviews were conducted and questionnaires were sent to both football clubs and football supporters. The interviewees were selected so as to represent all aspects of football information provision and demand, including journalists, broadcasters, writers, football club officials, academics and so on. The questionnaires aimed at the football clubs were sent to all 92 league clubs, 52 of which responded (57% response rate). Questions were asked about what kind of information services the football clubs offered and the attitude of the clubs towards official and unofficial sources of information (external sources included, e.g. newspapers). Football supporters were also contacted, with 28 out of 72 questionnaires filled (39% response rate). Questions included concerned the information services used, the frequency of use, the quality of information and paying

for services. Amongst the interesting results was that almost all fans (89%) believed that there was more information available then (in 1999) than at any other time previously, and they were happy with the amount of general football information available, with only few thinking there was too much (16%) or too little (12%). Supporters of clubs from lower divisions however, felt that information about their clubs was not always enough (60%). As for the type of information sought, football fans sought primarily News followed by Results, and then Fixtures, and Match Reports. The least sought information categories were Merchandise Information, Chat/Interactive and General History. The preferences for information sources were firstly Official (46%), then Unofficial (29%) and a combination of both (25%). Generally, the majority of fans thought that the quality of information services provided by football clubs were good (43%) or average (also 43%). Green concluded by saying that English football has had some success in dealing with many of the problems of the early nineties (e.g. violence and falling attendances) and is now reaping the benefits. However, they have often made mistakes concerning certain income streams (e.g. the idea of broadcasting live matches was originally rejected) and when it comes to the Internet, they need to ensure they do not make the same mistakes again.

2.6.2 Football in the Community

In England, football has historically had its place in the community. References to football in England have been made in literature from as far back as 1183 when William Fitz Stephen, biographer of Thomas á Beckett, discussed it. Since then, football in English society has changed significantly, but it has been a central part of popular life in England from at least the 18th Century (Walvin 1975). The current football league dates back to 1888, which at the time was the first national football league in the world (Football League Official Website 2006). In fact, England was the first country in a number of aspects concerning football. It was the first to have codified the rules of the game (Williams and Wagg 1991), the first to have professionals playing the game in 1888 (Walvin 1975) and the FA Cup was the first nationally organised competition (Football Association Official Website 2006).

English football's accolades go even further, with English society and communities playing an active role in the "export" of the "national game" (Giulianotti 1999), with football travelling to many places where the English were involved in trade and commerce (Walvin 1975). The English are said to have had some influence²³ in introducing football to many European countries including Denmark, Holland, Germany, Italy and France (Walvin 1975).

Meanwhile locally, football's involvement within the community has been fairly deep, with football playing an important role in English society (Fishwick 1989). Football is said by some to embody the nation (Giulianotti 1999). As the national team gets ready to play, players wearing the national colours and fans draped in the same colours whilst holding the flag, sing the national anthem. Many have claimed that football has numerous benefits for the community (Oughton et al 2003, Collins et al 1999, The Football Task Force 1999), including changing the quality of life in English society over time (Fishwick 1989, Walvin 1975). Throughout the 20th century, the major clubs were often the focal point of their respective communities, with many seen as more important than the local council or parish (Fishwick 1989). The clubs were taken so seriously, that they created tension between locals and non-locals, something that can still be seen today. Local businesses and businessmen have often been supportive of the team in various ways, from as far back as 1929 when several businesses shut early to allow employees to watch a big FA Cup match (Swindon Evening Advertiser, 1929 cited in Walvin 1975). Meanwhile, in 1935 when Sheffield Wednesday won the FA Cup, the local mayor proclaimed that the town had turned the corner industrially as well as in footballing terms (Sheffield Telegraph, 1935 cited in Walvin 1975). Even during times of high unemployment, fans would still go to matches (Williams and Wagg 1991, Fishwick 1989). Nowadays, the value of the major clubs is clear. If a club is doing well, the local economy feels the positive effects (BBC Website 2004, 2006c). Presently, major English clubs are doing remarkably well, particularly in economic terms. One major publication which looks at football in terms of money is the annual Deloitte and Touche (2008) report, which provides information about the twenty clubs with the greatest incomes in the world

²³ For example, the first Danish club was formed in 1876 by English residents in Denmark, whereas football was introduced to Holland by Dutch students who studied in English public schools (Walvin 1975).

(the "Money League"). The report discusses various financial performance aspects of clubs such as the sources of revenue for each of the clubs and how they could improve. For each club, a financial trend is shown with commentary on the recent performance of the club. The revenue streams are broken down into broadcasting, matchday and commercial. The top three clubs in this league for 2008 were Real Madrid, Manchester United and FC Barcelona. From the standpoint of this research, it is interesting to see the strength of the English (Premier League) clubs, 6 of which are in this top 20. Their importance in economic terms to local communities cannot be understated.

However, football is important not only because of the economic aspects, but culturally, as boys are raised with the game being an everyday part of their lives. In any given year, hundreds of thousands English school boys play football, and as they grow up, it is engrained in their minds. This, as a general sentiment, is something that appears to be widely accepted nowadays (at least within those with some interest in football) as over 75% of 43,590 fans who participated in the Football League's Supporter Survey (2006) agree or strongly agree that their club makes a significant contribution to the local community.

2.6.3 Football and Legal Issues

In an area where the fields of football and the web overlap, there are certain legal issues with the potential to have a notable effect on the football website environment and as such, are noteworthy. Copyright and defamation are two subjects which are particularly relevant. Here a brief background of these issues will be given for the purpose of covering a subject matter which, though not presently crucial to the research carried out here, is often highlighted by the media and has already had some effect on football websites.

In an industry such as football, where vast sums of money are involved in the creation and sale of the products (whether these be actual matches or memorabilia), it is only logical to mention copyright. The Premier League and the Football League both hold copyrights to the broadcasting of football matches (on various media and in various

formats), as well additional peripheral data (such as fixture lists, certain photos and so on, Football DataCo Ltd Email 2008), the total value of which is in the hundreds of millions of pounds annually (BBC Website 2006a).

As a result, it is perhaps somewhat unsurprising that there have been numerous legal disputes initiated by the Premier League in order to defend these rights. In several instances, the Premier League has pursued publicans in the UK for showing live matches without having the appropriate agreements in place (e.g. for using “domestic” rather than “business” subscriptions, Morning Advertiser 2007) and for using satellite systems from abroad. The first successful prosecution concerning the use of foreign satellite systems was carried out in 2003, at which time William Clark, manager of Ryan’s Pub in Radlett, Hertfordshire was ordered to pay fines and costs totalling £1,720. By using a foreign satellite service provider rather than BSkyB (who at the time, had paid vast sums of money for the exclusive rights to show these matches in the UK), Clark was found guilty of infringing the rights of the FA Premier League. Since then, more actions against publicans have followed. A notable such action was against Karen Murphy, a pub landlady who was showing live Premier League matches in her pub using a Greek satellite system. The noteworthy aspect of the Murphy case is that at the time of writing the outcome of the case is still pending on a hearing at the European Court of Justice (Morning Advertiser 2008a), with the defendant expected to argue that the European Union should be treated as one market, and that such intra-country restrictions have no place in such a single market. Such an action demonstrates both the complexity of copyright infringement cases, but also the will and desire of all the parties involved to win. Regardless, the Premier League, unperturbed, has also taken legal action against foreign satellite suppliers AV Station and QC Leisure, with UEFA, BSkyB and Setanta also getting involved (Morning Advertiser 2008b). Certain aspects of this case have also been referred to the European Court of Justice (England and Wales High Court 2008a).

Such legal battles, concerning the broadcasting of Premier League matches on television, are likely to continue for some time. However, this has not stopped the Premier League from defending its rights on other media (including the web). In fact, at the time of writing the highest profile legal battles are taking place against websites. The Premier League has had some success already by winning the rights over a

domain name (premierlive.net) which had been used by a Ukrainian football streaming website showing live matches for a relatively low fee (The Register Website 2008). However, a much more significant legal battle started when the Premier League decided to take action against YouTube/Google in 2007 by filing a class action complaint in a federal court in New York City on the 4th of May. In this action the Premier League is claiming that YouTube/Google have "knowingly misappropriated" its intellectual property (The Telegraph Online 2007). The Premier League has asked "a federal judge for a court order to stop YouTube's unauthorized and uncompensated use of the creative works of Class members, force YouTube to adopt existing technology to prevent unauthorized content from being exploited, and for damages for past infringement". The outcome of this case is likely to set a precedent for other such cases concerning not only football, but other sports too (as boxing bodies and rugby league bodies have also joined the Premier League in this class action, Premier League Press Release 2007). It could also be said that this issue is wider than just sports, with Viacom having taken YouTube/Google to court earlier in 2007 for "brazen" copyright violations (before the Premier League class action), asking for US\$1 billion (Washington Post 2007). Though it is difficult to speculate how, it is possible that the outcome of such a case will have a noticeable effect on the footballing website environment.

Defamation on the other hand, though not as financially significant, is also noteworthy as it has also affected websites. Here, legal battles have involved players (e.g. Ashley Cole, BBC Website 2006e), managers (e.g. David Moyes, The Observer 2008), club chairmen (e.g. Rupert Lowe, The Independent 2005), websites (e.g. craigmurray.org.uk, When Saturday Comes 2007), newspapers (The Times newspaper, The Independent 2005) and even fanzines (e.g. Gull's Eye – Brighton and Hove Albion, Haynes 1995).

The interesting link between defamation and football websites comes from the fact that there have been instances where the defamation (or alleged defamation) was carried out online. One instance where an alleged act of defamation occurred was even on a PHP (a blog). Alisher Usmanov, a Russian billionaire (Forbes 2007) who is a stakeholder (at the time of writing) in Arsenal FC (BBC Website 2007) was alleged to be a "Vicious Thug, Criminal, Racketeer, Heroin Trafficker and Accused Rapist"

by Craig Murray on his blog. As a consequence of this, Murray's website was temporarily offline when the website hosts, under pressure from Usmanov's lawyers, pulled the site down (When Saturday Comes 2007). It has since been moved to foreign servers, with the content in question remaining unchanged and without legal action having been taken against Murray, who still updates the site regularly (at the time of writing).

Other notable controversial and potentially defamatory remarks were made on football forums (Independent Hereford United Forum and Owlstalk). Supporters of Sheffield Wednesday were on the receiving end of a libel suit which went to court, where a judge ruled that the site administrator (though not responsible for the remarks) had to reveal the identities of certain individuals who had anonymously posted comments (England and Wales High Court 2007). In a separate incident, one Hereford United fan (Martin Watson), a season ticket holder of 20 years, found himself suddenly banned by club chairman Graham Turner for not removing a post on his forum quickly enough. The ban was later lifted when Watson removed the offending post and issued an apology (Bulls News Website 2007).

As is apparent from the brief descriptions so far, legal issues such as copyright and defamation are not completely independent of football websites. At this stage, both issues have changed the football website landscape with copyright related issues causing the shutdown of premierlive.net (The Register Website 2008) and defamation related issues temporarily taking down a weblog (When Saturday Comes 2007), as well as causing the forced removal of threads on forums (e.g. on the Independent Hereford United Forum).

However, both issues are complex ones which affect and are affected by numerous media. Specifically, the issue of copyright infringement both on and off the web concerning football, where legal battles are being fought not only by numerous major organisations such as the Premier League, UEFA, and Google, but also in various courts (i.e. England and Wales High Court, European Court of Justice etc), is a very complex one. As such, it is difficult to predict the effects that these legal matters will have in the long term. At this stage, it is clear that the Premier League is pursuing those it believes are infringing on its copyright (England and Wales High Court

2008a, England and Wales High Court 2008b) and has had some success. At the same time, it is apparent that ISPs can be convinced to shut down websites (When Saturday Comes 2007), and that forum administrators and forum users need to pay close attention to what is published on their sites (Bulls News Website 2007) so as to avoid litigation. As already mentioned, these legal issues have already had some effect on the footballing website landscape. However on a larger scale, other attempts to control content on the Internet have had mixed success (e.g. the music industry's attempts to stop illegal downloading of music). It is also questionable as to how much effort the Premier League and other such organisations might expend on shutting down PHPs, many of which have very small audiences. Regardless, at this stage, it is still possible to hypothesise that knowledgeable and concerned web authors (and forum contributors) will in all probability ensure that copyrights are not infringed and defamatory statements are kept off their sites. This is something that could indirectly encourage web authors to pool their resources (e.g. financial resources so as to buy copyright permissions or just additional contributors to help keep a closer eye on their websites), something which is closely linked to a recommendation made later in this work (see 6.4.2 The Club Community Composite Page (CCCP), p. 213). On the other hand, it is also a possibility that such issues discourage potential forum users and PHP authors from creating sites for fear of litigation. However, it is the opinion of the author that overall, PHP creation and activity should not be so greatly affected as ordinary web authors are unlikely to be aware of all the details of such cases.

Finally, in terms of the research carried out for this thesis, though it is acknowledged that the legal issues mentioned above have already had, and will almost certainly continue to have an effect on the availability of football PHPs as well as other football sites, they did not have an effect on the chosen methodology or structure of this work. This was the case for two reasons. Firstly, it was felt that it would be more appropriate to carry out a study which looked at the present state of football PHPs (rather than attempt to predict the future), and secondly, the author had no practical means of accurately and consistently verifying whether, in any given instance, copyright infringement or defamation had taken place. As such, the results shown in the later chapters show the state of the PHPs at the time that the research was carried out.

2.7 Chapter Conclusion

Perhaps one of the most notable aspects of this literature review is the small number of available research papers specifically on the PHP as an information resource. As such, it is unsurprising to have also found that the research that has been carried out is inconclusive to a certain degree, with basic elements such as a generally accepted definition for the PHP lacking. Regardless of this however, according to existing literature on the PHP in general, a study such as this is warranted, not least because of the great numbers of PHPs available on the web. Relevant studies carried out to date certainly indicate the possibility of PHPs carrying information that would be deemed useful by others (Dominick 1999, Weaver 2000, Papacharissi 2002b, Thelwall and Harries 2004, Narsesian and Nicholas 2005), with evidence that PHPs are created for the purpose of sharing information with others (Buten 1996) or as "filter" blogs (Blood 2002), which exist to improve the accessibility of information. However, at the time of writing, this is the first study looking specifically at subject based content on PHPs.

Further arguments in support of carrying out a study such as this thesis can be also taken from other subject areas, such as that of grey literature (e.g. where literature not designed to have longevity is deemed useful enough to be studied) and virtual communities (e.g. where individuals who have never met each other are happy to provide information that is requested of them).

The selection of football as a field in which to study PHPs in such a manner, is also supported by evidence found in existing literature. A football information service study has been carried out (Green 1999) and other surveys have been carried out which show that the Internet has become increasingly important as a source of information for football fans (e.g. FA Premier League Survey 2005, 2008). In addition, studying PHPs within the context of information behaviour as a whole, and the final recommendations that have come of it, will benefit the information flow in the football community, something that has been shown by existing literature (most notably by Green 1999, albeit a not very recently in Internet terms) to be unsatisfactory for many football fans (particularly in the lower divisions).

All in all, the five contributions of this thesis, concerning an improved understanding of PHPs, a new system of PHP collection and classification (by subject area), the identification of the trend of PHP authors moving towards collaborative efforts, the detection of the need to re-evaluate the results of older studies and the formulation of a skeleton outline for communal sites for the benefit of PHP authors (all of which are discussed in the later chapters), make a notable addition to the presently available literature. In fact, taking into account the lack of extant literature in the subject area of PHPs and the high number of PHPs present on the Internet, this project can be said to be all the more valuable. Furthermore, it must also be noted that the contributions affect not only the field of PHPs, but information science themes such as information seeking behaviour, grey literature, web genre studies and virtual community studies.

The following chapter looks at the novel methodology which was employed to make these contributions.

Chapter III

Methodology

3.1 Introduction

A PhD thesis such as this needs to have a methodology chapter for two important reasons. Firstly, the creation of such a chapter plays an important role in ensuring that the objectives set are indeed being met by the processes being carried out. Secondly, the methodology chapter is needed to demonstrate that the research being carried out is being done so in a systematic manner, essentially guaranteeing replicability.

This is a project which investigated the role of the PHP as an information resource within a broader view (including non-PHPs and indeed, non-web information resources). To date, there have been many projects conducted on information seeking behaviour, the web, and various aspects of PHPs, and these were examined whilst searching for an appropriate methodology for this project. However, as no research exactly like this had previously been carried out, a combination of established and original methods was put together to achieve the necessary goals.

3.2 Methodology Overview

The employed methodology was carried out in two main stages. An initial methodology was fashioned (Stage One) in order to meet the initial 11 objectives set out in the project. Having acquired the answers to the objectives set, a smaller second section stage was devised in an attempt to move the field forward (Stage Two). Hence Stage One deals with the understanding of the role played by the PHP as an

information resource, and Stage Two deals with ideas that attempt to make more efficient use of the effort of PHP authors.

Having set out the objectives for this thesis and choosing the field of study, it was clear that the research would involve work both on and off the web, primarily based on the activities of football fans. It was necessary to gather information directly from football fans (e.g. their preferences in terms of football information) in what is referred to in this thesis as People-Study, and the web (e.g. the PHPs) in what is referred to as Web-Study. Overall a two-pronged mixed methodology is followed (Tashakkori and Teddlie 1998), with both Stage One and Stage Two including a People-Study and a Web-Study.²⁴

For the Stage One People-Study, interviews were conducted and questionnaires distributed which communicated with football fans in order to create an understanding of their information seeking behaviour. Meanwhile in parallel, the Stage One Web-Study was conducted, locating and classifying PHPs, identifying whether information found on them was unique, and giving an indication of the degree to which it was accurate. The information gathered by these two separate portions of research was then put together, painting a picture of perceptions and actuality.

With the completion of Stage One, an attempt was then made at crafting a proposal in an effort to improve the current state of affairs. It was identified that an organised communal web effort could serve a purpose, streamlining the efforts of web authors and providing a better tool to web surfers. However, additional information was deemed necessary, concerning primarily what was already available and consequently, further research was undertaken. The purpose of Stage Two was to provide a better understanding of the existing collaborative efforts between individuals on football websites and a better understanding of the web page creating motivations of web authors including the obstacles in their paths.

In order to examine the collaborative efforts on football websites, the Stage Two Web-Study looked at all the sites for all the clubs in the top (Premier League) and

²⁴ These are differentiated by using reference to the “stage” as well as the “study” (e.g. Stage One People-Study or Stage Two Web-Study).

bottom (League 2) tiers of professional football in England. In order to further understand the motivations of web page creators, telephone interviews were carried out in the Stage Two People-Study asking questions concerning their motivations and difficulties. With this additional information, it was then possible to draw closer to a feasible proposal which actually moves the field forward.

3.3 Stage One People-Study

The purpose of the Stage One People-Study was to ascertain the opinions of the football fans (i.e. the users and potential users of football web pages), concerning various aspects of football information seeking. The general approach here was to find out how the information needs of football fans are met, and where the web and PHPs fit within that frame.

For this purpose, numerous techniques were available to make sense out of informational behaviour of individuals. Questionnaires, interviews, focus groups, observations, diaries, transactional logs, citations and library issue statistics have all been used in the past (Nicholas 2000). The selection of any of these methods, in addition to providing the answers to the research questions being posed (Hannabuss 1995), needed to overcome existent practical constraints (e.g. time constraints, financial constraints etc). Effectively, these practical constraints reduced the number of available options.

Focus groups and diaries were attempted, with limited success. The focus groups involved specifying times and places for multiple individuals to gather at, for a period of time. This proved to be difficult both because of time constraints (various people could not make various times) and geographical constraints (it was difficult to agree a mutually convenient location). Diaries were also handed out in a related pilot study, though these too were not deemed a success. The reason for this was the difficulty in collecting filled in diaries.²⁵ Other methods were not used (such as transactional logs and observations) because of the nature of the information sought. In the Stage One

²⁵ Though it is impossible to say with any certainty, as there is no evidence to prove this, it is the belief of the author that volunteers simply did not fill in enough of the diary to consider returning them.

People-Study, the research was concerned with the tools and methods used by football fans more than the way in which a particular tool or method was used, thereby nullifying the need for certain data collection techniques.

Consequently, questionnaires and interviews were chosen. Both of these methods of data collection met the needs of this thesis and indeed have been commonly used when gathering information about PHPs, or other work related to this thesis, such as leisurely information seeking behaviour or football in general (e.g. Buten 1996, Rieh 2004, FA Premier League Survey 2005). Having said this, whilst carrying out the literature review, it became apparent that, though numerous studies have examined PHPs in a similar (broadly speaking) context to the that of this thesis (e.g. Buten 1996, Dominick 1999, Thelwall and Harries 2002), with attempts made at PHP search engines as well (e.g. Ahoy!, HomePageSearch), interviews with potential PHP users have never been carried out. Questionnaires, which have been used more often (e.g. Buten 1996, Green 1999, FA Premier League Survey 2005, 2008) have also never been sent to potential users asking about the PHP within a broader information seeking context. They have been used for the purposes of carrying out research in the subject area (e.g. Green 1999, FA Premier League Survey 2005, 2008) and even specifically looking at PHPs (Buten 1996, though not in the specific subject area), but they have never looked at the PHP within a broader context. It was felt that this must be done so as to verify (or disprove) previous findings concerning PHPs (few as these may be).

Thus, interviews were initially carried out with 30 football fans. These were carried out with the intention of painting a picture of the opinions of the football fans' ideas on information seeking, and provide a foundation for the questionnaires which were to follow. Questions asked during the interviews included demographic questions (e.g. name, age), team support questions (e.g. who do you support, how long), team importance questions (e.g. how much money do you spend on your team), team information questions (e.g. how do you find out about the teams, what tools do you use) and web questions (e.g. which sites do you use, why). The initial structure of the

interviews (including the original list of questions asked²⁶) can be seen in Appendix A.

Questionnaires were then produced and distributed using ideas and opinions canvassed from the initial phases of the Stage One Web-Study and the interviews. The purpose of the questionnaires was to find out in more detail the tools and methods used by football fans to meet their informational needs. The questions asked by the questionnaires inevitably had some overlap with those asked in the interviews. Again, demographic, team support, team importance, team information and web questions were asked. Naturally, additional questions had also been included and the responses to all questions were by necessity more structured. The questionnaires also went into more detail about certain aspects of information acquisition. The distribution of questionnaires was carried out on the Internet and at football matches.

With this additional data, it becomes possible to decipher to a certain degree whether the perceptions of the football fans match the results of the PHP investigations.

3.3.1 Football Fan Interviews

The study included the gathering of data through semi-structured individual face-to-face interviews with 30 football fans. Fifteen of these were season ticket holders, and another 15 were non-season ticket holders (Appendix B). The process of the selection of football fans was intended to discover the most dedicated football fans, who in turn would be very active in information seeking. For this purpose, season ticket holders were initially pursued. The reasoning here was that individuals who would spend endless hours and significant sums of money physically following their team would be more likely to be "informationally active". This logic though proved to be only partially correct. Throughout the course of the interviews, it became apparent that though some fans do spend thousands of pounds and hundreds of hours following their team, this does not necessarily mean that they are very active in terms of information seeking activity (e.g. some of them simply go for the love of the ground

²⁶ An initial list of questions was created for the interviews, and these were generally all used. Naturally, additional questions arose during the actual interviews themselves. These additional questions are not included in the appendix.

atmosphere). Once this was discovered, a decision was taken to interview fans who themselves claimed to be strong supporters, in addition to the season ticket holders in an attempt to increase the probability of finding “informationally active” fans.

The very first set of questions for the interviews was conceived during an initial "brain storming" session. Naturally, the aims and objectives of the research were used as guidelines for deciphering suitable ideas, as were the experiences from the early phases of the Stage One Web-Study. The first set of questions was then tested by conducting semi-structured open-ended pilot interviews on suitable interviewees (season ticket holders). Here, the functionality of audio recording equipment and venues was test-driven, feedback was taken into account, new possible questions were discovered and appropriate interview structural changes were made. The initial structure of questions turned out to be too rigid. As interviewees were only too happy to discuss their passion, they answered several questions at once (i.e. their reply to a single question included answers to subsequent questions). Hence, this was taken into account, and the questions were restructured. Another lesson learnt from the pilot interviews was the working of the audio equipment, which failed under certain circumstances (e.g. noise projected from a specific angle would interfere with the recording of the speaker's voice). This whole procedure was then repeated (twice) until no more question revisions were deemed necessary. At this point a final structure for the interview was settled upon, and the process of interviewing began.

The next step was searching for potential interviewees, a process which also proved problematic. Initially, attempts were made to post "calls for volunteers" on messageboards (websites) of teams in and around London. The reason for this was the geographical constraints of carrying out face-to-face interviews. The postings stated that a researcher was looking for fans (who own season tickets) to interview, that the interviews would take around 30 minutes and the time and place were flexible (see Appendix C). However, after 10 postings and one single response,²⁷ it became apparent that a new approach was necessary. To this end, searches were carried out for websites of teams in London which had names (and email addresses) of

²⁷ This response ultimately did not yield an interview.

contributors, and these fans were contacted directly.²⁸ This procedure proved to be notably more successful, with most of the respondents being both season ticket holders and interested in taking part. Thirty eight (38) emails were sent out, with 18 suitable replies received. In the end, 12 of the season ticket holders interviewed were found this way, who themselves provided contact details for the remaining three season ticket holders who were interviewed, as well as four non-season ticket holders. The remaining 11 non-season ticket holders were located through the author's own involvement with football fans.

Participants were interviewed during a period of 12 months starting in March 2005, and ending in February 2006. The emails sent to potential interviewees were informal, personalised, customised and friendly. The subject and aim of the study were briefly mentioned,²⁹ as was the confidentiality of the data and the approximate expected length of the interview. The email also offered reassurances that the interview was not a test of knowledge, rather merely a canvassing of opinions. It was made clear that the choices of date, time and location were fully flexible, and could be selected by the interviewee themselves. In fact, the content of the email was designed to meet the need to acquire consent to carry out the interview. As such, the email was designed to make the potential interviewee feel as comfortable as possible, as well as appreciated. A notable attempt to arouse the curiosity of the candidate was not made as the subject matter is already of great interest.

The interviews were then carried out at mutually agreed dates, times and places (in most cases at a neutral venue such as a quiet public house) at the convenience of the interviewees. Upon meeting the interviewee (prior to the start of the interview), a brief introduction to the nature of the research and researcher's background was given. The interviewee was then asked if they objected to the use of the audio recorder or of the hand-written notes, and the interview began. Barring any objections, the dialogue was then recorded using the digital voice recorder and hand-written notes. In length, the interviews varied between 20 minutes and 45 minutes, with an average of approximately 30 minutes.

²⁸ Fans on messageboards were not approached directly, as it is not clear from the messageboard whether or not a fan is a season ticket holder.

²⁹ The subject and aim were not revealed in any detail, so as to avoid biasing the responses in advance.

Once the interview was completed, the hand written notes were rewritten (for purposes of clarity) by the end of the day. The answers to the interview questions were then rewritten electronically whilst listening to audio recordings within three days. The results were then summarised on an MS Excel spreadsheet for ease of reference, with parts of the results also entered on to SPSS version 11 (originally, Statistical Package for the Social Sciences), which allowed for more flexibility in the management of the results.

3.3.2 Football Fan Interview Questions

For this study, two sets of interview questions were used, with very minor differences between them. Initially a set of questions was designed with season ticket (ST) holders in mind. Later, when the decision was taken to include non-ST holders (who can otherwise qualify their status as a fan), a slightly different second set of questions was designed. In fact, the only difference in the questions were concerning season tickets (or the absence of season tickets) and an additional question for non-ST holders to qualify their status of fandom. Other than these points, there are no differences in the questions asked of both groups. Both sets have the same 6 sections, with the non-ST holder questions having one additional (single question) section. It must also be noted, that this thesis does not claim or attempt to imply that a fan who is a ST holder is in any way a “greater” or “lesser” fan than the non-ST holder.³⁰

The reason for the existence of this differentiation was born of the need to qualify an individual who is to be interviewed as a football fan. Initially, a decision was made where only season ticket holders would be interviewed. As season ticket holders make such a clear effort (one which could even be monetarily measured³¹) to follow their team, it was not felt necessary to qualify them explicitly. However, so as to ensure

³⁰ This issue is beyond the scope of this thesis.

³¹ Season tickets cost upwards of £200 per annum in lower divisions and can approach £1000 in the Premiership.

that all interviewees were indeed football fans, a question was included where a chance is given to the non-ST holders to state their qualification as a football fan.³²

ST Holder Sections	Non-ST Holder Sections
Personal information	Personal information
Supporting the team	Football dedication
The importance of the team	Supporting the team
Team information	The importance of the team
Web questions	Team information
Final questions	Web questions
	Final questions

Table 3.1 Interview questions

Personal information

This section was included to acquire basic knowledge about the interviewee. The name, age, gender, occupation, place of residence and contact details were asked. By ascertaining the demographic background of the interviewee, it was then possible for comparisons to be made with other studies. The contact details were requested again here so as to provide for easy access in case follow up clarifications were needed.

Football dedication

This section was only present in the non-ST holder interviews, and was essentially the question which asked the interviewee to qualify their status as a football fan. Season ticket holders did not need to account for their fandom, as their dedication could be "measured" in monetary and temporal terms.

Supporting the team

Questions concerning the team supported were asked in this section. The name of the team was asked as well as how long, how and why a team was supported. This was also the only section which differed between the two sets of questions. For ST holders, the final questions in the section asked when and why the season tickets were purchased. For non-ST holders, the final questions asked whether they had been to any of their team's matches and

³² This differentiation is not made in the questionnaires as individuals who filled these in were either at a football match (offline questionnaires) or saw the request to fill the questionnaire in on a football forum (online questionnaires), which already implies a certain level of interest in football.

when they first went. These questions were included to provide some insight into the background of the fans, which in turn allowed for clearer interpretation of results.

The importance of the team

As the section name suggests, the questions in this section were about the importance of the team. Questions included how much money was spent (annually) on the team, how much time was spent discussing the team, why football was important, and what effect it had on other activities. Like the “supporting the team” questions above, the answers to these questions provided additional insight into the fan’s mindset.

Team information

These were the questions that started probing about the information seeking behaviour of the fans. The questions included asking about the need and frequency of the need to look for specific information concerning the fan’s team. The way the information was looked for was also asked, using which tools or methods, why and which tool or method was the most important one. The next set of questions asked about the success or failure of the searches, and the reasons why they were carried out. The questions asked if the interviewee always found what they were looking for, how likely they were to give up, what happened if they failed, when, how and why they did this. The last questions in the section asked if they had any other information needs, and if any other important questions should be asked. The Team Information questions were included to provide information concerning the general information seeking behaviour of football fans.

Web questions

This part of the interview was concerned with the interviewee's use of the web. Questions concerning what sites they primarily used (concerning their team), what types of sites they used, what the main factors were in deciding to use a site, particularly looking at quality, quantity, ease-of-use, intellectual level, viewpoint, how up-to-date a site was, and how far back it went were asked. Further questions asked if they had ever decided to stop using a site altogether

and if they used all the sites equally regularly and why. The final question in this section asked if they believed anything could be done to improve the service provided by the sites that they used. This section was included so as to provide information concerning the importance of the web and the PHPs.

Final questions

The final questions of the interview were asked just after explaining in more detail the purpose of the study. Till this point of the interview, the details of the study were not given away so as not to bias the answers of the interviewee. Once more details were given, the interviewee was asked if there were any additional comments that they would like to make, and if there was anything else the interviewee felt should have been asked. These questions were included to ensure that major gaps were not left uninspected in this section of the thesis.

3.3.3 Questionnaires

The second part of the Stage One People-Study involved the questionnaires, which were also used to gather data concerning the information seeking behaviour of football fans. In order to do that, it was important to ensure that the questionnaires produced were suitable for the data they were trying to gather. As such, the design of the questionnaire was a procedure that was not taken lightly. A thoroughly thought out and well designed questionnaire can be the difference between the success or failure of a research project. A questionnaire is said to be more than a set of questions (Oppenheim 1992), and as such, there were a number of decisions to be made. These included deciding the type of questionnaire to use, the way to motivate potential respondents, and then deciding the questions which were going to form the questionnaire.

Various types of questionnaires can be used in such projects, which can be either offline or online. Offline questionnaires include among others, telephone questionnaires, postal questionnaires and administered questionnaires (to groups or

individuals). Online questionnaires include those sent by email (either inline, as a document attachment or program attachment) and those placed on a web page.

The differences between offline and online questionnaires include:

Cost

Generally, electronic surveys are considered cheaper than postal surveys (Bachmann and Elfrink 1996, Kiesler and Sproull 1986).

Sampling

As already mentioned, this can be a problem, as only people with access to the Internet can fill out electronic surveys (Tse 1998) and they tend to be middle class (Mehta and Sivadas 1995) educated young males (Schmidt 1997).

Operational and design differences

Electronic surveys in general are produced, completed and returned faster than mail surveys (Gunter et al. 2002). Electronic surveys also have the ability to incorporate multimedia attributes (Bishop 1997).

Response rates

Online response rates are in general, poorer than offline response rates (Gunter et al. 2002).

Quality of response

It appears that the quality of online surveys is slightly better than offline surveys (Gunter et al. 2002).

Another aspect of the use of questionnaires is attaining the cooperation of potential respondents. A questionnaire can be fully functioning, and carefully designed to meet its aims, yet without respondents, the whole effort can fail to bear worthwhile results. In order to overcome this issue, several methods can be employed. These include the offer of some sort of prize, extensive advertising and re-advertising, and also the way the subject matter is delivered (if it is of an "exciting" nature).

Finally, the questions to be included and structure of the questionnaire must be decided. A questionnaire can have questions which are closed (fixed), open-ended or combined (open and closed).

As the naming suggests, closed questions are ones in which the possible replies are fixed. Open-ended questions do not have any such restriction, and as a result, the responses tend to be longer. The combined option involves offering fixed responses, but also an "other" choice (with room for additional information) which the respondent is free to fill in as they desire. Each of these choices have advantages and disadvantages. Closed questions are easier, quicker to answer, and so save time. This increases the number of questions that can be answered in any given time frame. Also, closed questions are easier to manipulate when it comes to their analysis. The main disadvantage is the lack of freedom the respondent has, as all the possible responses might not be recorded, and this in turn creates a certain bias in the final results. An additional closed question problem is that of respondents lapsing into "response sets". As they go through the questionnaire, respondents might just check the first response they find, regardless of whether it is what they believe. Meanwhile, open-ended questions can combat just that. By not constricting responses in any way, they give the respondents freedom to express their own opinions, and actually inhibit them from lapsing into a response set. As expected though, by giving respondents this freedom, the amount of time taken to fill out a questionnaire is increased. In addition, as the amount of time needed increases, the response rate drops.

3.3.4 Employed Questionnaire

The designing of the questionnaire was a lengthy process during which numerous revisions were made. The commencement of the design process took place after the first 10 interviews had been carried out.³³ Once again, the initial set of questions was contrived by thinking of all possibilities, and then reducing the list by removing similar, overlapping or irrelevant questions. Again, the aims and objectives were the

³³ The chosen number (of ten) is not something which can be decided in advance. In this study, it was the point at which the author deemed that a pattern of responses became clearly visible. In other studies, the number of interviews conducted before the commencement of the questionnaire design takes place could be notably different yet just as valid.

main guidelines, supplemented with ideas from the Stage One Web-Study and the initial results of the interviews. Numerous other questionnaires were examined in the process, including surveys about football and PHPs (Sunday Mirror/NPower 2001, FA Premier League 2005, Buten 1996).

The online questionnaire was web-based and self-administered. As a result it was possible to get responses from all around the UK, though the respondents had to complete the questionnaires by themselves. Emphasis was also placed on ensuring that all relevant information was extracted by including the right questions (i.e. the questions that will gather the data required to meet the objectives). The layout and order were designed to ensure that the questions could be answered in as short a time as possible, and that the questions were clear. To this end, most of the questions were closed, although there were also a few open-ended ones.

The distribution of the questionnaires was carried out both electronically on the Internet and manually at football matches (in person). As such, it was necessary to have two slightly different versions of the questionnaire (an online version and a paper version). This affected only the structure of the questionnaires, and did not affect the questions themselves which were identical (e.g. on the paper version, there was an option not to answer the Internet usage questions if the participant did not have access to the Internet. For obvious reasons, this option was not present in the electronic version of the questionnaire).

Once a first draft of questions was created, it was posted on the Internet and an email was sent out to suitable candidates to test the structure, clarity and functionality of the questionnaire. Concurrently, a paper version of the questionnaire was created, and handed out to suitable candidates. Feedback was received for both versions and respective revisions were carried out. Again, the procedure was repeated (for both versions) until no more revisions were deemed necessary, at which point the final versions of the questionnaires were ready. The questions were identical on both versions, with differences only in the layout and structure of the questionnaires. To

encourage potential respondents, a chance to win a £50 cash prize³⁴ was offered to all respondents who included some form of contact details (i.e. telephone number or email address).

The electronic version of the questionnaire was created using Microsoft FrontPage (version 4.0.2) in October 2005 and was placed online. Once active, a post was made on an online football forum site (for each football team in the top four divisions, i.e. 92 posts were made in total).³⁵ In this instance, the rivals.net web franchise was used, which had a separate online forum for each football team.³⁶ Finally, the results from all the completed questionnaires were automatically saved in a text, "comma delimited" file.

The distribution of the paper questionnaires was carried out at two London football grounds (Highbury and Craven Cottage) which were chosen for feasibility reasons,³⁷ in October 2005. Permission to have questionnaires filled in outside the ground was not sought from the clubs, because such a need was not felt necessary (as the researchers were not physically on the property of the club).

Two researchers each stood near the ground (yet outside the ground, before the turnstiles), and asked every 10th fan to fill out the questionnaire. The researchers provided any help that might be required (e.g. making clarifications, providing pens etc). The completed questionnaires were collected and then manually entered onto the online questionnaire. The results were saved in a separate file, so as to keep the electronic and paper results separate.

The electronic storage of all the questionnaires (both online and offline) was done initially using comma delimited files, which were then exported to SPSS for

³⁴ With the help of an online random number generator (www.random.org), respondent #219 was awarded and paid the cash prize.

³⁵ Initially emails were sent to football websites asking to them advertise the questionnaire. Unfortunately, though this method had some success, the number of respondents was low. As a result a second method was devised which placed posts on the football forums.

³⁶ Rivals.net was chosen because (in theory at least) a single registration allows the placing of posts on all 92 forums. Though this worked for the vast majority of cases, for a handful of sites, technical difficulties meant that posts were placed on other forums.

³⁷ Namely, having matches played geographically near, at suitable dates and times, and for the convenience of a colleague who agreed to help in the distribution of questionnaires.

consolidation, ease of storage and management. MS Excel was also used in conjunction for more flexibility in the creation of charts.

3.3.5 Questionnaire Questions

The questions of both versions of the questionnaire are identical. Each questionnaire is split into 7 sections (see Appendix F and Appendix G).

Questionnaire Sections
User characteristics
Supporting the team
The importance of the team
Team information
Web questions
Personal Home Page (PHP) questions
Final questions

Table 3.2 Questionnaire questions

User characteristics

In this section basic questions were asked such as age, gender, job title and annual income. The options provided as responses were taken from the FA Premier League National Fan Survey (2005), in order to be able to match the demographics of the two surveys. This then provided the opportunity to weigh the results according to a widely accepted base.

Supporting the team

This section offered all the professional teams as options to choose from and then asked questions concerning the dedication of supporter (i.e. how long they have been supporters, are they ST holders, how many matches they have attended). This section was included so as to be able to distinguish fans who were more active supporters from others who were less active.

Importance of the team

This section asked how much money was spent annually on the team (e.g. on season tickets), and how much time was spent talking about the team. The data

gathered from this section was used to further distinguish very active fans from less active ones.

Team information

This section asked respondents to identify tools and methods they use to find information on their teams and give reasons as to why they did so. Other questions included asking about the frequency of specific information searches, how these were carried out, why, where and when. These questions were included to gather data about the information seeking behaviour of football fans.

Web questions

This section focused on Internet usage, including how much time was spent reading on the Internet, what types of sites were used, the importance of attributes when it came to selecting sites and whether all information needs were met using the sites in question. This section was included to gather data about the fans' relationship with the web.

Personal Home Page (PHP) questions

The focus in this section was the PHP in general, and the perceptions of users on PHPs. The questions were concerned with the identification on PHPs, what constituted a PHP, as well as qualities of PHPs (e.g. ease-of-use, up-to-date etc) and how often they were used. These questions were necessary to get information concerning the usage and understanding of PHPs by football fans.

Final questions

The final questions were two open ended questions which asked whether the respondent had any additional comments to make, or if there is something else that should be asked. As in the interviews, this section was necessary to ensure that major gaps in the research were not overlooked.

Hence, these questionnaires, along with the interviews described earlier, complete the Stage One People-Study. The next step in the methodology for this thesis was the Stage One Web-Study.

3.4 Stage One Web-Study

Any study looking at PHPs in detail will need to conduct a certain amount of work online (e.g. Bates and Lu 1997, Crowston and Williams 1997, Dominick 1999, Papacharissi 2002a), and this thesis was no exception. However, as an examination of PHPs for such a purpose from this subject based viewpoint has never before been carried out, it is difficult to pinpoint similarities. The benefit of the results of this Web-Study would perhaps most affect further attempts at creating PHP search engines (such as Ahoy! or HomePageSearch). With a better understanding of the information found on PHPs, more suitable search engines can be created. Furthermore, the implementation of the Stage One Web-Study to other fields should provide more insight into the information flow in other fields, and is related to virtual communities (e.g. Rheingold 2002), virtual communities of practice (e.g. Dube, Bourhis and Jacob 2006) and grey literature (e.g. Thompson and Guistini 2006).

Hence, the examination of the PHPs was carried out in the Stage One Web-Study, which found a number of PHPs and proceeded to compare various aspects of them to non-PHPs. This was done in an attempt to determine the uniqueness of the information contained, as well as how accurate the information was and how far back it went. In doing so, a number of methods were used (some of which can be considered a type of content analysis) which have only ever been used in this research and are considered an original contribution of this thesis.

Content analysis is defined as a method of making valid inferences from text (Weber 1990) or in more precise terms, “a research technique for making replicable and valid inferences from text (or other meaningful matter) to the context of their use” (Krippendorff, 2004). This part of the study (the Stage One Web-Study) can be broadly considered a type of content analysis. Although it does not aim to make inferences from the content of websites in the sense that the Krippendorff (generally regarded as the leading expert of content analysis), has defined, it has some major elements of content analysis to it, including: dealing with the content and messages of websites as a type of media, categorising the content (using the content ratings

described below in 3.4.5 Relative Information Contents Ratings (RICR)), involving judgment, and dealing with the accuracy and the validity of messages presented on websites. Although content analysis has already been applied to websites (e.g. Macias and Lewis 2003), the processes and techniques used in this part of the study are original. These include the process of finding the PHPs specific to particular football teams (i.e. specific to a subject area, something considered an original contribution of this thesis), checking the uniqueness of the information (including archival aspects) and checking the accuracy of the information.

In order to do so, the Stage One Web-Study found web pages about professional football teams in the UK, ranging from the English Premier League to League 2 (i.e. the top four divisions). Four teams from each division were randomly³⁸ chosen, for each of which, three PHPs were chosen as well as three non-PHPs (usually more commercial, CNN.com style websites). According to this method, in total 48 teams' PHPs should have been chosen (three PHPs per team, four teams from each of the four divisions).³⁹ However, for three clubs⁴⁰ in the lower divisions, only one PHP was found (instead of three). As a result, a total of 42 PHPs were checked, each against three non-PHPs. The teams were chosen randomly in an effort to get a representative sample. As there are no universal measures which define the size of a club, it was not possible to order clubs in a way to include bigger and smaller clubs.⁴¹

This process was carried out over a 9 month period from May 2005 to February 2006. The entire process is described in detail presently, with an explanatory example embedded in the descriptions, so as to further clarify the exact workings of the methodology. The illustrative example uses the Queens Park Rangers (QPR) football team. At the time this research was carried out, QPR were in the Championship (i.e. the second highest professional division).

³⁸ An online random number generator (random.org) is used to provide a set of four numbers between the starting and ending number of the teams in each division (i.e. between 1 and 20 in the Premiership, between 1 and 24 in the Championship and so on). The position of each team is that at the start of the season before any games are played (i.e. they are in alphabetical order).

³⁹ Three PHPs for each of the four teams in each of the four divisions ($3 \times 4 \times 4 = 48$).

⁴⁰ The clubs represented by just one PHP were Brentford, Rochdale and Chester City.

⁴¹ Teams are often said to be "big" or "small", however it is difficult to define these terms. Some teams are "big" in terms of stadium capacity and attendance (e.g. Sunderland) but have relatively few trophies to compare. Other teams have a very high number of trophies, but a relatively small stadium (e.g. Liverpool).

The results of the Stage One Web-Study were recorded originally on paper, and then re-written electronically, stored and managed using tables created in MS Excel.

3.4.1 Finding PHPs

The first step to be carried out in this investigation was the search for suitable websites (including PHPs) for study (i.e. sites dedicated to specific football teams). Finding PHPs, in particular, proved to be challenging. Studies have not previously looked for PHPs on specific subjects, choosing rather to use generic PHPs from service providers (e.g. Dominick 1999, Papacharissi 2002a) or PHPs of specific domains, such as academic (e.g. Thelwall and Harries 2004, Nomura, Ishida and Yokozawa 2001). The initial plan for finding such sites was to use the Open Project Directory (dmoz.org). However, as this turned out to be far from exhaustive, especially concerning PHPs, a new method needed to be devised. The second attempt at a finding a suitable procedure included carrying out searches using Google and checking the first 300 results. However, though this would catch many PHPs, often, some PHPs would be left out of the first 300 results. The method which was finally settled upon was inspired by a "web-bot" strategy, where the software looks for sites by checking all the links on every site it finds. Hence, in this method an initial search is carried out on Google, in order to find several sites (not necessarily PHPs). These sites are then checked for links to other sites on the same team, until all the links are exhausted. To ensure that as many PHPs as possible are located, additional coverage is provided by using a specific search phrase on Google, and then manually checking the first 300 entries for PHPs.⁴² The figure of 300, is a balance between that which is feasible (i.e. within temporal constraints) and that which would be ideal.⁴³ The entries are checked both for PHPs and non-PHPs. However, the non-PHP sites found have to be dedicated solely to the team in question (pages such as the QPR pages of the BBC website or of ananova.com are not included at this stage). This procedure produces two lists of sites, including a list of PHPs about a football club, and another list of

⁴² In the pilot study, the 300 Google entries were the primary search, and the link following search was intended to provide additional coverage. However, it became apparent that the link following search was considerably more effective.

⁴³ Naturally, under the best possible circumstances, every single search result would be examined. However, this is not feasible on a project such as this (which has limited resources).

non-PHPs for the same club. Such a process of finding websites, and the subsequent classification (to follow), produces the result of finding PHPs on a specific subject, something which has not been carried out before and is considered an original contribution of this thesis.

Finding QPR PHPs Illustration

For QPR, the initial searches were carried out using the Google search engine with the following search terms (quotes are included so as to treat the words as a phrase):

"queens park rangers"

The first 10 results included four sites dedicated to QPR (this included the official QPR site) and 6 semi-relevant sites (e.g. a football ground's site, a news site story on QPR and so on). All four relevant sites were non-PHPs, one of which had links to other QPR sites, including PHPs. From there on, all the QPR links were followed, until a list of QPR sites was compiled (see Table 3.5 later in this chapter).

An additional Google search was then carried out so as to provide additional coverage, using the search term:⁴⁴

+"queens park rangers" +my

The first 300 results were checked. However, all the sites dedicated to QPR had already been discovered using the "web-bot" strategy.

3.4.2 PHP Inclusion

For a PHP to be included in this study, a number of criteria must have been met. These criteria are listed in Table 3.3. Essentially, the content related requirements were that the web page be a PHP which is about a professional football club in

⁴⁴ The search phrase is +my +"[name of football team]". This ensures that the document retrieved has the word "my" in it, as well as the complete name of the football team in question. This phrase was decided (after several gruelling sessions of trial and error) so as to overcome the endless articles on the team from the numerous commercial football and news sites in the world.

England. In addition, there were three requirements related to practical issues which were come across in early trials and literature (Smith 1997, Oliver et al 1997). Here, the PHP must be "online" at the time of the searches, with the relevant sections written in English and the PHP must be free (in monetary terms) to use. The actual identification of the PHPs was done according to the definition of de Saint-Georges (1997). As such, the PHP had to either:

- (ii) Claim that it is a PHP
- and/or (ii) Have personal information on it (such as a CV, photograph etc.)
- and (iii) Represent a person rather than a group⁴⁵

Main Criterion	Description
Scope	- Must be about professional football in England
Content	- Web page must be a PHP - Web page must have at least a subsection on a specific football club
Workability	- Web page must be accessible at time of study - The relevant sections of the web page must be written in English
Cost	- Web page must be free (in monetary terms)

Table 3.3 Criteria for PHP selection (custom-made using Smith's (1997) Toolbox)

3.4.3 PHP Classification

With the selected definition and criteria, there was still ambiguity when it came to deciding what constituted a PHP in practice. To overcome this obstacle, a manual step-by-step process was devised where the researcher is given specific signs to look out for and specific classifications to make. Automatic attempts have been made to classify websites, often as parts of other studies, but only for very generic top level domain classifications (Lawrence and Giles 1999, Koehler 2002). Accurate automated classification is generally regarded as difficult (Shepherd and Watters 2004). The reason for these identification issues lies in the fact that all the information required to ascertain whether or not a website is a PHP, is not necessarily available on every

⁴⁵ If one of the following is true, the web page is considered to represent a person rather than a group

- i. The author of the web page has made references to the web page as "my web page".
- ii. The web page has a single name in the "Credits" section of the web page.
- iii. The author of the web page sends an email to the researcher claiming that the web page represents a single person.

PHP. Simply viewing a site is not necessarily enough and the solution to this problem was not readily available, hence the need for the devised non-automated step-by-step process. Web page classification methods in general have been used by numerous studies, for example for identifying the PHP as a genre in the first place (Dillon and Gushrowski 2000, Crowston and Williams 1997) or using web page classifications to explain aspects of other studies (Lawrence and Giles 1999, Koehler 2002). However, the definitions used in these have not been explicitly stated, and neither have the classification procedures used. In this sense, these studies could not be used to provide guidance. Other studies have also classified academic PHPs (Middleton et al 1999, Thelwall 2002, Adamic and Adar 2003), and these were used as a basis for ideas, but ultimately the only classification scheme to be replicated in part was Doring's (2002), who has devised a classification scheme which can be used for all PHPs.

Using Doring's (2002) classification scheme, a decision flow-chart (Figure 3.1) was created (after several trial and error exercises) to illustrate the workings of the procedure. The purpose was to be able to say whether a site was a PHP, and should be included in this investigation (by examining the site and looking for specific signs). If the necessary signs were not present (or otherwise unclear), an email was sent to the site creator (if this was possible) as a final attempt to verify it as a PHP (see Figure 3.1 for details).

The replicability of the procedure was tested along with the replicability of the entire Stage One Web-Study. The classification of sub-genres of PHPs (such as blogs and photosites) was not strictly necessary. However, as a classification process was already necessary to decipher the PHPs from non-PHPs, it was deemed potentially useful to carry this out as well.

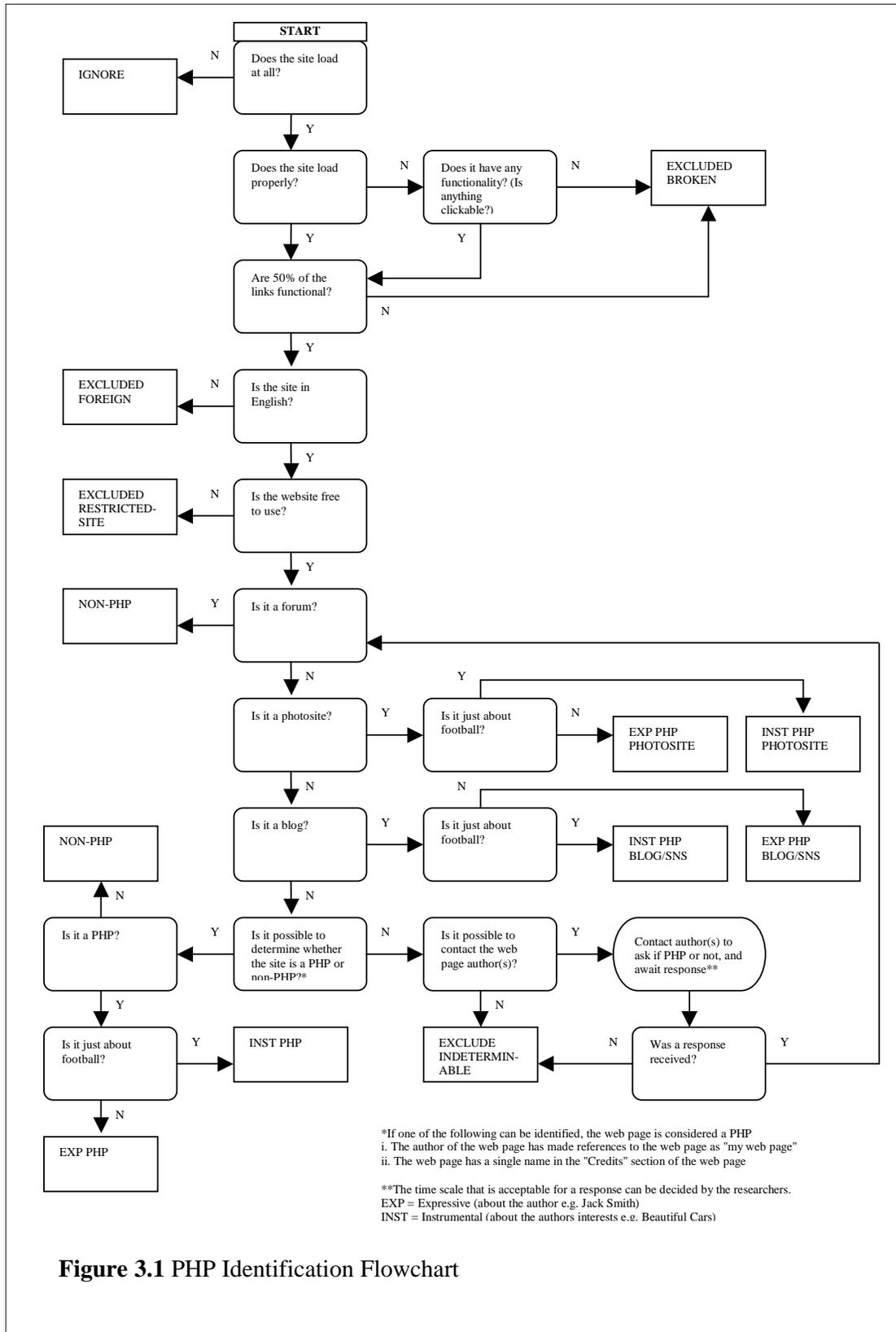


Figure 3.1 PHP Identification Flowchart

As can be seen from Figure 3.1, using this method there are a variety of possible classifications. It is apparent that a number of different types of sites are considered PHPs, namely blogs/SN sites⁴⁶ and photo-sites. There is also a further classification of PHPs as expressive or instrumental as mentioned in Doring (2002). An expressive page is one where one's own person is the topic. An instrumental page is one where the main topic is treated without explicit reference to one's own person. Naturally, there are also circumstances which lead to exclusion from a PHP classification. Three such scenarios which lead to the sites not being classified as PHPs are "foreign", "restricted" or "broken". A foreign site is one where the main language used is not English, a restricted site is a site that is protected and not readily available for all users (e.g. sites which require subscriptions), and a broken site is one where more than half the internal links on the site are not functioning. All these exclusions were necessary either for feasibility reasons (e.g. restricted sites could not be accessed) or for the sake of not producing an imbalance in the results (e.g. broken sites would put undue weight on certain aspects of the results). It must also be noted that such sites have also been excluded in previous classification schemes - e.g. Doring 2002. Table 3.4 shows sites which are included or excluded as PHPs.

Included as PHPs	PHP exceptions
Standard PHP (a generic PHP, e.g. a site on members.aol.com)	Foreign PHPs (where the primary language is not English)
Blogs/SN site (a "diary" type site, e.g. a site on blogspot.com/friendster.com)	Broken PHPs (where more than half the links are dead links)
Photo-site (a site for sharing photos, e.g. a site on photos.yahoo.com)	Restricted site (where the content is restricted e.g. pay-sites)

Table 3.4 Inclusion and exclusion of PHPs

Following the PHP identification flowchart (Figure 3.1), each site was classified. Sites included in this study are the standard PHPs, blogs/SN sites and photo-sites. Each site could be either instrumental or expressive, depending on the content. Sites which provided information about a subject in particular, as most football PHPs did, were considered instrumental. Sites which were primarily about the author, and only discussed football briefly were treated as expressive. Occasionally, a page could be

⁴⁶ Blogs are also known as weblogs. SN site is an acronym which stands for Social Networking site.

part of a site which was expressive, but the page itself was instrumental. In those instances, the classification of the site was expressive.

As mentioned in Figure 3.1, if there is insufficient information to classify a site, the author was contacted by email⁴⁷ and asked about the site. If no response was received, a reminder email was sent out seven days later. In cases where a response was still not received within 10 days of the reminder email, the site was treated as a non-PHP. This system was adopted for purposes of feasibility. Ideally, a longer waiting period should have been allowed. However, as it was important to move forward, these time scales (which were still generous) were used.

Once the searches were carried out and the lists were complete, the comparisons began. Three PHPs representing each of the selected clubs were randomly selected and checked for information. The same information was then looked for in the appropriate non-PHP sites representing the same club. The non-PHP sites included a comprehensive sport/football website (e.g. www.bbc.co.uk/football), the official home page of the football team itself (e.g. www.qpr.co.uk), and one other non-PHP home page, dedicated to the football team (e.g. www.qprnet.com).

The PHP chosen for illustrative purposes was Dave's Unofficial QPR Website (#2 in Table 3.5, www.queensparkrangersfc.com) and the non-PHPs chosen were the Official QPR site (#10, www.qpr.co.uk), QPRnet.com (#17, www.qprnet.com) and the BBC web page (<http://news.bbc.co.uk/sport1/hi/football/teams/q/qpr/default.stm>) on QPR. At this stage, the selection of PHP was made for the purpose of showing a complete illustrative example. The selection of non-PHPs was solely to depict a fair representation of the "footballing web".

Dave's Unofficial QPR site has been put together by D.A. Barton, and generally comes across as a non-PHP.⁴⁸ Its main page has six headings. These are Latest, Players, History, Interactive, Links and About (in terms of football sites, these headings or their equivalents are common in dedicated non-PHP football sites such as

⁴⁷ The email address is usually available on the site. If there is no way to contact the author, and it is unclear as to what type of site it is, the site is treated as a non-PHP.

⁴⁸ With the tools available to the general public nowadays, it is quite common for skilled web authors to put together sites which appear to be professionally built.

qpr-mad.co.uk). Each of the headings is split further into subheadings. Each of these subheadings was examined to see whether the information within it could be found on any of the three non-PHP sites which were being examined (see Table 3.7 onwards).

The three non-PHPs were chosen in an attempt to find a representative sample of sites that football fans were likely to visit. Surveys have already pointed at the BBC and the official sites being extremely popular (e.g. FA Premier League Survey 2005), and so these were chosen. In addition, a third single active non-PHP dedicated solely to the team in question, was also chosen, as an additional site. A popular type of site that would have also provided insight would have been a forum. However, as forums do not have the layout of a traditional site, searching through them proved to be unfeasible, and as a result, these were not included in the comparisons. The sites found and classified for the explanatory example can be seen in Table 3.5.

Number	Name	Type
1	Blueandwhite.net	INST PHP BLOG/SNS
2	Dave's Unofficial QPR Website	INST PHP
3	Home and away with the mighty Rs	INST PHP PHOTO
4	Home of the Swedish Hoops	Non-PHP*
5	Hoops Online	Non-PHP
6	In the Loft (QPR Mad)	Non-PHP
7	My Cousin Plays for QPR!	EXP PHP
8	Nigel Unlimited	INST PHP
9	North London Rs	Non-PHP
10	Official QPR site	Non-PHP
11	QPR 1ST [supporters trust]	Non-PHP
12	QPR Ireland	Non-PHP*
13	QPR loyal supporters association	Non-PHP
14	QPR Rule!	INST PHP
15	QPR Russia	EXP PHP PHOTO
16	QPR.ORG.UK	INST PHP
17	QPRnet.com	Non-PHP
18	Rangers into Europe (well Cardiff)	INST PHP PHOTO
19	Rangers till I die	Non-PHP
21	Richard Langleys PHP	Non-PHP*
22	Supporters club – Jarrow branch	Non-PHP*
23	The alliance in the loft	Non-PHP
24	The alternative page for qpr football club	INST PHP
25	Untitled- (Commentary by Cure guitarist about QPR)	EXP PHP
26	Westbourne Rs	Non-PHP

*Sites whose type is undetermined are eventually treated as non-PHPs.

Table 3.5 Results from QPR website search

3.4.4 Comparing PHPs

Once the pages to be examined had been selected, sections (and subsections) of the PHPs were compared, and then given Relative Information Content Ratings (RICR). In order to give an indication of the differences between the various sites, it was necessary to come up with a process of checking and recording these differences. After initial trials where numerous sites were compared, initial RICRs were decided upon. These were assigned to sections of sites in trials, with new circumstances causing minor changes in the way the RICRs were assigned. After 20 website sections were compared with no further revisions necessary, the process was deemed suitable for a pilot study, the results of which were subsequently published in a peer reviewed journal (Narsesian and Nicholas 2005).

There are two points which must be noted in order for the explanation of the comparisons to be clear. Firstly, for the purposes of these comparisons, all information was assumed to be of equal value (whether on a PHP or non-PHP). Different pieces of information generally are of different value to different people. As an example we can say that the outcome of a cup final match between Liverpool FC and Fulham FC might be considered "essential information" to a fan of Fulham FC. However, the same information can be of very little value to a whole array of other people (e.g. a French teacher in Cameroon with no interest in football). The value of individual pieces of information is beyond the scope of this project. Secondly, though the names of the sections and subsections on a website were useful in guiding the authors to where the information on the website was, they were not important in determining whether information on a site was available or not (e.g. Hall of Fame (DUW) is compared to Loftus Legends (QPROW)). All websites which were compared were thoroughly checked.

Once the selections were carried out, all the instances of information found only on the PHPs was looked for on the rest of the web. This was done using terms the author believed to be relevant on a search engine (i.e. Google). If this failed, a follow up search was conducted using the Dogpile meta-search engine (see 3.4.7 Uniqueness of Information on the Web, p. 106).

A sample of non-unique information was then taken from the PHPs (approximately 10% of the total number of instances checked) and checked for accuracy against other web sources. Ten percent was chosen, once again for feasibility reasons. The process of carrying out these comparisons is very time consuming, and as a result, it was not possible to carry out more in depth checks. However, it is recognised that 10% is not sufficient to claim with certainty that the results are valid. This is not an absolute measure, but is used to provide an indication of the degree to which the information on the site is accurate.

Once again, this method of website comparisons was devised specifically for this thesis and is considered part of one of the original contributions of this work. Previous examinations of PHPs have been either observational (e.g. Bates and Lu 1997, Dominick 1999, Papacharissi 2002a) or focused on improving design (e.g. Maruyama 1999, Flanagin and Metzger 2003). Even studies looking at the PHP in terms of information resources never compared PHPs to non-PHPs, creating the necessity to devise such a system, almost all aspects of which are unique to the work carried out in this thesis.

3.4.5 Relative Information Contents Ratings (RICR)

There are five different RICRs. These are Locally Exclusive (X), Web-Wide Unique (XX), Archival (A), Equinferior (\leq) and Not Applicable (N/A). A Locally Exclusive rating indicates that the section contains information not available on any of the sites compared to, with Web-Wide Unique indicating that the section has information not available elsewhere on the web. The Archival rating is given to sections that contain "old" information, the Equinferior rating to sections where the information is available in the sites being compared to and finally the Not Applicable rating is given to sections that do not contain information or contain irrelevant information. These RICRs apply to one section of a website against other websites and they help determine what PHPs have to offer in terms of information resources in comparison to non-PHPs. These are all discussed in more detail presently.

The Locally Exclusive (X) Rating

This is one of the more significant ratings for the PHPs in this methodology. When the PHP has a piece of information, or has an information subsection, which is not found on any of the non-PHPs it is being compared to, this is classified as Locally Exclusive. The symbol for Locally Exclusive is the letter X.

There are three instances when this can occur. The first and obvious instance is where there is some information which is available on the PHP, and not available on the non-PHP. An example can be found in the Match Reports/Current subsection of Dave's Unofficial QPR Website (DUW), which in turn can be found under Latest. The match reports contain a detailed commentary of what happened at the match in question in an essay or report layout. The DUW match reports contain, in addition to the essay style commentary, the ratings of every player, and the man-of-the-match (the player who had the greatest impact on the match).

Of the three sites DUW is being compared to in this example, all have some sort of match report. The Official Website of Queens Park Rangers Football Club (QPROW) has a skeleton report which does not contain an essay style report, but has instead only certain details about the game, such as the score, the players and any cautions (see Table 3.7 PHP comparisons). QPRnet.com (QNET) has match reports in a style very similar to DUW, with a detailed essay as well as a man-of-the-match. Finally, the BBC Sport -QPR Section (BBC) has a detailed essay, but also a live text commentary which details every instance of the match. This live commentary can still be viewed after the match has ended.

The Locally Exclusive rating is given in this instance to the Match Reports/Current subsection of DUW because of the player ratings that the site provides. This is a specific piece of information which is only available on DUW, and not on any of the non-PHPs, and thus qualifies for the Locally Exclusive rating.

The second instance is when the researcher believes that the information provided by the PHP is, for some reason, superior to that provided by any of the non-PHPs. In this

instance, the explanation behind the decision will be given in the relevant table. An example of this case is the Club History section within History on DUW. This section covers the history of the QPR club from 1885 to 2001. QPROW has a section covering the QPR's history from 1887 to 1997. QNET and BBC have no section dedicated to the history of QPR. As DUW covers the history of the club over a period not covered by any of the non-PHPs (i.e. 1885-1887 and 1997-2001), the Locally Exclusive rating is given to this section.

The third and most subtle instance is where there is a certain piece of information on a PHP which simply goes further back in time. As an example, we can use a comparison between DUW and QPROW. DUW has a subsection within Latest News called Newspapers.⁴⁹ As expected, this subsection contains a roundup of stories found in the papers about QPR on days when there are such stories. Meanwhile QPROW has its own section called What the Papers Say which provides the same service. In this case, the two sites do not have a notable difference in terms of the service provided. What is notable though, is the fact that DUW started the service before QPROW and hence the information goes further back (23.3.03 on DUW as opposed to 16.5.03 on QPROW). Therefore the rating given is Locally Exclusive.

Within this methodology, the Locally Exclusive category might imply that there is some bias towards PHPs. For an instance to be placed in this category, it means that a piece of information is available on the PHP, but not on the non-PHPs. The reverse cannot be true. This study occasionally discovered instances where there was information uniquely available on non-PHPs, however, this instance is not shown explicitly. The reasons for this are both practical and logical. It would be unfeasible to do the required systematic investigation to discover all the instances where information on non-PHPs is unique and at the same time, discovering these instances is not part of the aims and objectives of this thesis.

⁴⁹ Though the link to the subsection is called Newspapers, the actual heading on this part of the website is titled 'What the Papers Say'.

The Web-Wide Unique (XX) Rating

A Web-Wide Unique rating is given to a section of a PHP that has information which is completely unique web-wide. In order to get a Web-Wide Unique rating, a section must already have been assigned a Locally Exclusive (X) rating. A search is then carried out to ascertain whether the information found in the section in question is available anywhere else on the web. If the search does not return any results, then the rating is converted from a Locally Exclusive to a Web-Wide Unique rating. Any section can be assigned either a Web-Wide Unique rating or a Locally Exclusive rating (not both). The symbol for the Web-Wide Unique rating is XX.

Clearly, the Web-Wide Unique rating gives the clearest indication (of all the RICRs) as to the contribution that PHPs can make in informational terms. Keeping track of the instances where information available on PHPs is not available elsewhere on the web is certainly worthwhile, and for that reason, this is a rating in its own right.

An example of the Web-Wide Unique rating is the Players subsection of the DUW site. This section of the site provides average monthly ratings for each player. This subsection was originally given a Locally Exclusive rating. Following that, a search was carried out for the average QPR player ratings on the Internet and this information was not found. As a result, this subsection was given the Web-Wide Unique rating.

The Archival (A) Rating

The Archival rating is given to an instance where the section on the PHP contains information that is "older" (meaning it concerns earlier years). Again, for a section to be assigned an Archival rating, it must have first had a Locally Exclusive rating. However, the Archival rating is given in addition to the Locally Exclusive rating (in the PHP Comparison tables, e.g. in Table 3.7, the Archival rating is always in brackets next to the Locally Exclusive rating).⁵⁰

⁵⁰ Whereas a section can have either an X or XX rating, this is not true of the A rating. A section can have an X rating and also an A rating. The A rating is then used separately (i.e. in the results chapters,

The Archival rating is useful to have so as to be able to demonstrate instances where the PHPs are providing a historical snapshot of the web. For certain studies, such as social history, the snapshots provided by PHPs could potentially be used to shed light on matters that other sources might not be able to. For this reason, it was deemed to be worth keeping track of Archival instances separately from the Locally Exclusive instances.

To further illustrate the point of the way the Archival rating is given, the instance which demonstrates the assignment of the Locally Exclusive rating above will also be used here. The comparison is between DUW and QPROW. The DUW (as already mentioned) has a subsection within Latest News called Newspapers which was given the Archival rating. This subsection contains a roundup of stories found in the papers about QPR on days when there are such stories. Meanwhile QPROW has its own section called What the Papers Say which provides the same service. The Archival rating is given because the DUW started this feature before QPROW.

The Equinferior (\leq) Rating

An Equinferior rating can be given to an instance where the researcher believes the information found on the PHP is either equal to or inferior to that on the non-PHPs. The symbol for Equinferior is \leq .

Clearly, there will be instances where the information supplied by the PHP will not be as thorough or in depth as the information on the non-PHP. Though this is not the focus of the study, it was felt that this too should be kept track of, so as to have a clearer view of the picture as a whole.

An example of the Equinferior rating is the Latest subsection within the Latest News section within the Latest area of DUW. This subsection covers QPR news, and the available news dates back to 17.8.01. All the non-PHPs also have sections on QPR

if the total number of sections is 100, then the X, XX, \leq and N/As must add up to 100. The A rating will not be included here).

news. QPROW has news from 16.6.02, QNET since 9.3.00 and the BBC since 29.9.00. Since DUW does not cover a period of time covered by the BBC (or QNET), the rating given is Equinferior.

The Not Applicable (N/A) Rating

In any instance where the comparison of the sections should not be carried out, the Not Applicable RICR is given. Websites often have complete sections which are in no way related to their usefulness as an information resource on a specific topic (in this case football). When this occurs, the Not Applicable rating is given. This rating is also given if the information within the PHP section is non-existent (this occurs when a PHP has a section on a dedicated topic, but the section is incomplete or unavailable). The symbol for the Not Applicable rating is N/A.

A good example to demonstrate such a case is in the Credits subsection within Links on DUW. This subsection is about people who have at certain times helped the author/creator of DUW with the site. The non-PHPs being compared to DUW may or may not have equivalent sections as part of their sites, however, as this has no bearing on the football information aspect of the site, the comparisons are never made. Thus the rating given is Not Applicable.

Relative Information Content Rating	RICR Symbol	Description
Web-Wide Unique	XX	Provides a resource unavailable elsewhere on the web
Locally Exclusive	X	Provides a resource that none of the non-PHPs (being compared to) provide. This can be in the form of information that is completely novel or simply more in depth
Archival	A	Provides unique information about older events (i.e. events that occurred in earlier years)
Equinferior	≤	The PHP provides an inferior or equal service
Not Applicable	N/A	The comparison is not valid either because the information in the section is missing or because it is irrelevant (i.e. unrelated to football)

Table 3.6 Relative Information Content Ratings (RICR)

3.4.6 Comparison Results

This section contains the tables of results from the comparisons made, using the RICRs defined above.

Section: LATEST	RICR	Dave's Unofficial QPR site	Official QPR site	QPRnet.com	BBC Sport- QPR section
Latest News -Latest	≤	QPR news since 17.8.01	QPR news since 16.6.02	QPR news since 9.3.00	[QPR news since 29.6.00]*
Latest News -Newspapers	X (A)	What the papers say since 23.3.03	What the papers say since 16.5.03	N/A	N/A
Match Reports -Current	X	Very thorough with player ratings and Man-of-the- Match	Skeleton report**	Very thorough, with Man-of- the-Match	Very thorough, with live commentary
Match Reports -Archive	X (A)	Archive of reports since the 1998-99 season, player ratings, Man-of-the- Match	N/A	Archive of reports since 2001-02 season, with Man-of-the- Match	[Match reports since 12.8.00]
League Table	X (A)	Five league tables in total, since the 1998- 99 season	League tables since 2001-02 season	Current league table only	Current league table only
Fixtures/ Results -First Team	X (A)	1996-97 season onwards	2001-02 season to present	N/A	April 2000 to present
Fixtures/ Results -Reserves	N/A	N/A	Not checked		
Fixture/ Results -Youth Team	N/A	N/A	Not checked		
Columns	≤	1 column	N/A	5 columns	Not QPR specific

*Information in square brackets is placed there because "it is not in the right place". Sometimes the information on the websites is not in the right section (e.g. under the latest news section, there are the news archives, but not the latest news), or is not easily available (e.g. BBC has match reports dating back to 12.8.00, but these can only be found via their site search engine, there is no section on Match Reports). If this is the case, it is placed inside square brackets. It must be noted that the information must be in the same "format" to be accepted as being available (e.g. a table of goalscorers can be put together by reading all the match reports, however the table of goalscorers is treated as separate information).

**Skeleton reports do not have an essay style report, rather just listed facts including the score, scorers, venue, attendance, cards shown (cautions) and substitutions made.

Table 3.7 PHP comparisons for QPR

Section: PLAYERS	RICR	Dave's Unofficial QPR site	Official QPR site	QPRnet.com	BBC Sport- QPR section
Players Stats	X	Apps, goals, average rating, Man-of-the- Match	Red cards, yellow cards, clean sheets/ goals, appearances (by competition)	Red cards, yellow cards, goals, appearances (by competition) [Man-of-the- Match]	Red cards, yellow cards, goals (by competition)
Players Profiles	X	Thorough biography	Brief stats* and basic biography	Brief stats** and basic biography	N/A
Top Scorers	X (A)	List of scorers since 1991-92 season	Top scorers only, 2001-02, 2002-03	N/A	List of scorers, current season only
Monthly Awards	X	By decision: Player-of-the- Month, Goal-of- the-Month, Young-Player- of-the-Month since 1999-00 season	N/A	By vote: Player-of-the- Month (since 2001-02), Goal- of-the-Month, Match-of-the- Month (since 2002-03)	N/A
Yearly Awards	X (A)	By vote: Player-of-the- Year since 1999-00 season	N/A	By vote: Player-of-the- Season (since 2001-02), GOTS, Match- of-the-Season (since 2002-03)	N/A

* Official QPR stats: Birthplace, DOB, Position, Signed, Honours, Height, Weight

**QPRnet stats: Squad Number, Position, Signed, DOB, Nationality, Previous Clubs, International Honours.

Table 3.8 PHP comparisons for QPR (continued)

Section: HISTORY	RICR	Dave's Unofficial QPR site	Official QPR site	QPRnet.com	BBC Sport- QPR section
Hall of Fame	X	17 players with biography	[only 2 players at a time]	N/A	N/A
Club History	X	1885-2001	1887-1997	N/A	N/A
Results	X (A)	1996-97 season onwards	2001-02 season onwards	N/A	April 2000 onwards

Table 3.9 PHP comparisons for QPR (continued)

Section: INTERACTIVE	RICR	Dave's Unofficial QPR site	Official QPR site	QPRnet.com	BBC Sport- QPR section
Images	≤	19 sets of pictures (for 16 players and 3 matches)	N/A	26 sets of pictures	Not QPR specific
Wallpaper	≤	17 wallpapers (of 8 players)	N/A	25 wallpapers (of 15 players)	N/A
QPR CD-ROM [QPR merchandise]	X	CD of the 1982 FA Cup final	QPR Club Shop merchandise*	No QPR specific merchandise	No QPR specific merchandise
Quiz	X	Small 5 question quiz with prize	N/A	Yes**	N/A
Polls	≤	15 polls	No separate section	Monthly and yearly awards	No separate section

*Even though the site sold all the merchandise available in the club shop, the 1982 FA Cup final was not available (on any media).

** The site has a Quiz (in the Interactive section) subsection, however, at time of writing, no Quiz had ever been held/available.

Table 3.10 PHP comparisons for QPR (continued)

Section: LINKS	RICR	Dave's Unofficial QPR site	Official QPR site	QPRnet.com	BBC Sport- QPR section
QPR Links	X	21 links	N/A	15 links	1
Football Links	X	17 links	N/A	11 links	No separate section
Other Links	N/A	Not Relevant			

Table 3.11 PHP comparisons for QPR (continued)

Section: ABOUT	RICR	Dave's Unofficial QPR site	Official site	QPRnet.com	BBC Sport- QPR section
Web Site History	N/A	Not Relevant			
Stadium Guide	X	Directions, stadium layout and history, ticket prices, pubs, parking	Directions, stadium layout, ticket prices	Directions, ticket prices, match day programs, fanzine	N/A
Contact the Club	X	Address, phone numbers, email addresses	Address, phone numbers	Address, phone numbers, names of staff	N/A
Credits	N/A	Not Relevant			

Table 3.12 PHP comparisons for QPR (continued)

3.4.7 Uniqueness of Information on the Web

Once the comparisons were completed, the next stage involved further investigations to determine whether the instances of Locally Exclusive (X) information were available elsewhere on the web. Once again, this was something which had not been previously attempted and makes up part of an original contribution of this thesis. This procedure is a possible two-step process (the procedure ends when the information is located, so the number of steps taken depends on how quickly the information is located) that consists of collecting all the instances of Locally Exclusive RICRs given and carrying out searches on the web to find them. Firstly, a search was carried out on Google to find the specific instance of the information. If this yielded no positive results, the same search was carried out on a meta-search engine called Dogpile (www.dogpile.com). This procedure was carried out with Dave's Unofficial QPR Website, and the results are in Table 3.13. To illustrate further, the details of the first search carried out will be discussed presently.

This search was of the player ratings for QPR players for the current season. The search terms used were:

+QPR +"player ratings"

Google returned 378 results and Dogpile returned 54 results. Of these results, the first 300 from Google were checked, and the first 30 from Dogpile. Again, the numbers of 300 and 30 used were a balance between what was ideal (all the results) and what was feasible within the given time frame. Having carried these searches out, the player ratings required were not found. There were three sets of results which were promising. The first was of player ratings in a football match that involved QPR, but the discovery was on a non-QPR site (Swansea City FC), and the player ratings were for non-QPR players. The second result was of a match with ratings of football players who were fans of QPR, playing in an Internet Football League. Finally, the third result was ratings for QPR players in a fictional game (the results from a computer game, called Championship Manager). Just to conclude, player ratings for QPR players at the time of the search did not appear to be available anywhere on the web. The full results for Web-Wide Unique instances of information found on DUW's site are available in Table 3.13.

This is perhaps the section of the process which is the least rigid and most open to variable results. The problem of not finding information which is available on the web, has afflicted the web since the early days of search engines (Lawrence and Giles 1999). One might argue that where the author has failed in finding certain data or information, others might succeed. However, the counter argument is that the author, as a Computing graduate and with over 15 years of web searching experience, should be reasonably likely to find the required information. In addition, it could be argued that if a PhD candidate in Information Science cannot find certain information on the web, it is inaccessible enough for it to lose at least some of its value.

As for replicability, carrying out web searches for a specific piece of information can be done in several ways, and two individuals might not carry out exactly the same searches to find the same information. However, this section of the procedure was also checked for replicability (see Appendix E), and was deemed a success. Though it might be somewhat surprising, it appears that successful replication of this process is possible.

DUW Section	Description of Instance	Availability	Final RICR
Latest - Latest News - Newspapers	What the papers say, since 23.3.03	No	XX
Latest - Match reports -Current	Player ratings	No	XX
Latest - Match reports -Archive	Player ratings	No	XX
Latest - League table	League tables since 1998-99 season	Yes	X
Latest – Fixtures/ Results – First Team	Results since 1996-97 season	Yes	X
Players – Player Stats	Average ratings	No	XX
Players – Player Profiles	Thorough biography of various players	No	XX
Players – Top Scorers	List of scorers since 1991-92 season	No	XX
Players -Monthly awards	Young Player of the Month	Yes	X
Players – Yearly Awards	Player of the year since 1999-00	No	XX
History – Hall of Fame	Biographies of 17 players	No	XX
History – Club History	Club history from 1885-2001	Yes	X
History – Results	Results since 1996-97	Yes	X
Interactive -QPR CD-ROM	CD of the 1982 FA Cup final	No	XX
Interactive -Quiz	QPR quiz	Yes	X
Links – QPR Links	21 QPR links	Yes	X
Links – Football links	17 links	Yes	X
About –Stadium Guide	Where to find parking	Yes	X
About –Contact the Club	Club email address	No	XX

Table 3.13 Locally Exclusive and Web-Wide Unique instances of information (QPR)

The table above shows that 52.6% (10 out of 19) of the instances of Locally Exclusive information were not found using search and meta-search engines, making them Web-Wide Unique.

3.4.8 Accuracy of Information

A sample of non-unique information was then taken from the PHPs (approximately 10% of the total number of instances checked) and checked for accuracy against other web sources. Ten percent was chosen once again, for feasibility reasons. The process of carrying out these comparisons is very time consuming, and as a result, it was not possible to carry out more in depth checks. However, it is recognised that 10% is not sufficient to claim with certainty that the results are valid. This was not an absolute measure, but was used to provide an indication of the degree to which the information on the site is accurate.

The actual method for verification has been devised specifically for this research. Previous studies of PHPs have never examined their contents in such terms. As a result, once again, the method devised here is part of a process which makes one of the original contributions of this thesis (see 7.4 Original Contributions, p. 236).

In order to verify the accuracy of the information, the instance of the information was checked against the same section in another website. Practically speaking, this means that another website must have had the same information about the same issue (i.e. if on site X, it was written that Wayne Rooney is 1.78m tall, then site Y must have been found saying that Wayne Rooney is 1.78m tall). The website chosen to verify the accuracy of the information would have been known to contain the information either because of the PHP comparisons, or the checking of the availability on the web.

Meanwhile, as a section of a website might contain any number of facts or pieces of information, it would not have been feasible, given the size of this project, to check every individual fact. As such, only 7 "facts" were checked from any section.⁵¹ Once these 7 instances of information were examined, a "degree of accuracy" rating was given. There are only three degrees of accuracy. These are Low, Medium and High. The rating given depends on the number of accurate or inaccurate "facts" checked (see Table 3.14 Degree of Accuracy).

⁵¹ If there are less than 7 facts in a section, they will all be checked.

Degree of Accuracy	Number of errors (out of 7)	Description
Low	4 or more	The information is either completely inaccurate, or inaccurate enough to be unusable. For the information to be deemed unusable, it must be less than 50% accurate.
Medium	2 or 3	The information is more than 50% accurate but not in the High category.
High	1 or fewer	The information checked is all accurate, or almost all accurate. Information is deemed to be "almost all accurate" if only one error is found.

Table 3.14 Degree of accuracy

Other practical considerations came into play when making comparisons on various types of information found on web pages. Generally speaking, there were three common comparisons to be made. Clearly, the most common one was information in textual form. In this instance, the same information was looked for on the non-PHP. If the same information was available on another site, it was considered accurate.⁵² The second most common checks made involved web hyperlinks to other sites. In this instance, if the link was a functioning link, it was considered accurate. Finally, the third most common were labelled photographs. In order to classify the accuracy of the label as accurate, the exact same photograph must have been found with a matching label.

In the rare instances of finding other non-textual files (e.g. video or audio), the checks were not carried out for feasibility reasons. Instead, the next suitable section was checked.

Concerning the sections that were checked in general, choosing the sections involved dividing the total number of sections by ten. The resultant number was then used as the "skip" (i.e. if the number is five, then every fifth section was checked). If this nth section could not be checked⁵³ (e.g. the information on it was unavailable), the next

⁵² Clearly, it is possible that some information on the web as a whole is inaccurate. This study does not deal with that aspect of the accuracy.

⁵³ There are three main instances where sections cannot be checked. Firstly, the section of the site might not be relevant. Secondly, the section of the site might not be available or external (it could either be a dead link, or a link to another site). Thirdly, the information in the section of the site might be completely unique to web.

valid section was checked⁵⁴ and this did not affect the position of the skip. The end result was that (approximately) 10% of sections, all in all, were checked, thereby completing the Stage One Web-Study.

DUW Section	Description of Instance	Degree of accuracy	Website used for verification
Latest- League Table	League table	High	BBC Sport
Players- Player Profiles	Additional Information on Players	High	Official QPR website
History- Results	Previous results of football matches	High	Official QPR website
Links- QPR Links	Links to other QPR sites	Medium	N/A

Table 3.15 Accuracy of information (QPR)

Once again, within all the PHP and other relevant studies which were examined (see 2.4 The PHP as an Object of Study, p. 36), a similar methodology for checking the accuracy of information on PHPs was not found (so as to be compared against). However, like the rest of the Stage One Web-Study, this section too has been peer reviewed and published (Narsesian and Nicholas 2005). Overall, it is the belief of the author that this new methodology for the analysis of the content of PHPs (i.e. the entire Stage One Web-Study) is an important contribution of this thesis. With the number of PHPs (in various forms) available on the Internet, there can be no doubt that potentially valuable information can be found on them. As a result, a methodology such as this that facilitates both the finding of PHPs on a subject basis and the inspection of the information found on PHPs, must be considered valuable for the further examination of these webpages from this different viewpoint. They have already been considered valuable enough to provide insight into self presentation (Wynn and Katz 1997, Papacharissi 2002a), web genres (Dillon and Gushrowski 2000) and grey literature (Thompson and Guistini 2006) to name but a few issues. At this point, there is no reason to believe they cannot be useful as an information resource as well.

⁵⁴If sections one, four and 7 were to be checked, and section four was not available, sections one, five and 7 would be checked.

Step	Action
1	Find PHPs and non-PHPs on a club
2	Check availability of all information on PHPs against availability of information on selected non-PHPs
3	Check availability of information on rest of web
4	Check accuracy of information found on PHPs against other sources

Table 3.16 Stage One Web-Study: Step by step

3.5 Stage Two Overview

Having met the objectives set initially, an attempt was then made at crafting an outline for a communal site that would encourage web authors to work together in the creation of web pages for the online community. Stage Two was carried out as a check to ensure that the blueprint designed was feasible. Stage Two itself also has two parts (again a Web-Study and a People-Study). Firstly, in the Stage Two Web-Study (concerning the communal site), every site dedicated to a Premiership and League 2 club was examined in order to have a solid understanding of the types of sites already in existence and so as to ensure that any proposal is indeed moving the field forward. Secondly, for the Stage Two People-Study, interviews were carried out with web authors, to ensure that their ideas would be taken into account before finalising any recommendations.

3.6 Stage Two Web-Study

These website investigations were carried out so as to have more information available concerning already active online website collaborations. Using this information, additional ideas that might be applicable to the concept of the communal site could be utilised. These online website investigations were carried out over a four month period, from June 2007 to September 2007. As already mentioned, every site dedicated to a club in either of the two professional leagues chosen (the Premier League and League 2) was examined (a total of 834 sites - see Appendix I). The league table for the 2006/07 season was used, giving a total of 44 clubs⁵⁵ covering the top end of the professional leagues (i.e. the Premier League) where there is the

⁵⁵ The 20 clubs in the Premier League and the 24 in League Two make a total of 44 clubs.

greatest amount of information available (and by extension the highest number of sites) and the bottom end of the professional leagues (i.e. League 2) where the least amount of information is available. Once again here, comparisons with previous studies are difficult to make, however, some studies have looked at a certain genre of websites in order to define them, and the examinations here bear similarities (broadly speaking) to such studies (e.g. Crowston and Williams 1997, Ujigawa 1999). However, once again, these did not look at websites from a subject based viewpoint.

3.6.1 Website Examination Procedure

The procedure for finding sites was the same as the procedure followed in the Stage One Web-Study (see 3.4.1 Finding PHPs, p. 88). In this instance, all sites dedicated to a club were examined, looking specifically at the way the site functioned, trying to decipher whether or not the creation and maintenance of the site was a collaborative, publicly open effort.

To this end, two aspects were looked out for. When examining the site, searches for the "About" and "Help" sections were carried out, in a process which was looking for information on ways to contribute to the site (e.g. write a match report). If such a mechanism could not be found within the About or Help sections, every link on the site was clicked (excluding certain sections of sites such as advertisements, polls and multiple identical sections, e.g. one match report was clicked on, not every match report), the purpose always being to find a way to edit the site. The reason for the examination of the About and Help sections was so as to determine whether or not the site had an unusual structure or functionality. In such a scenario, one would expect the novel features to be advertised and described in some detail (particularly if a novel mechanism was being used), probably in the About or Help sections, if not an area with even higher visibility.

The idea was that the site be open to a large group of people for instant contributions and updates. Hence, an overall process for identifying such a mechanism could not be set in stone, simply because it was not possible to predict what the site might look like. What was clear though, was that for a site to be open to the general public, it

must be reasonably obvious that there was an open mechanism in place. Though this essentially placed the responsibility of being found on the site itself, the procedure did examine practically every publicly clickable link which must be considered sufficient.⁵⁶

Once it was determined that a site did indeed have a contribution mechanism and that it was a collaborative effort, more information was sought on the site. This was done by reading the relevant sections of the site, registering (and logging on) if necessary, and sending emails to the site administrators to determine the details of the collaborative aspects (where these were not clear).

The results of the website investigations were recorded on to an MS Excel spreadsheet.

3.6.2 West Ham Online Website Investigation Example

For purposes of clarity, a website investigation example has been included here. The searches located a number of sites (see Appendix J), for all of which the following procedures were carried out. This example uses the WestHamOnline.net site (WHO), which had notable collaborative aspects.

The first step in the procedure looks for the About or Help sections. On the WHO site, however, there are no such sections. The front page though, does give the impression of a regular West Ham website (see Figure 3.2).

⁵⁶ Of course, every effort was made on the side of the researcher to spot such a mechanism. However, the idea was to find a site that was open to football fans in a manner that encouraged the fans to make an instant contribution. By logical extension, it is vital that such a mechanism be easily identifiable for it to work. Using this argument, though every effort was made on the researcher's part to identify such mechanisms, if one which exists was not found, it was probably the system on the site as a whole not fulfilling its intentions.

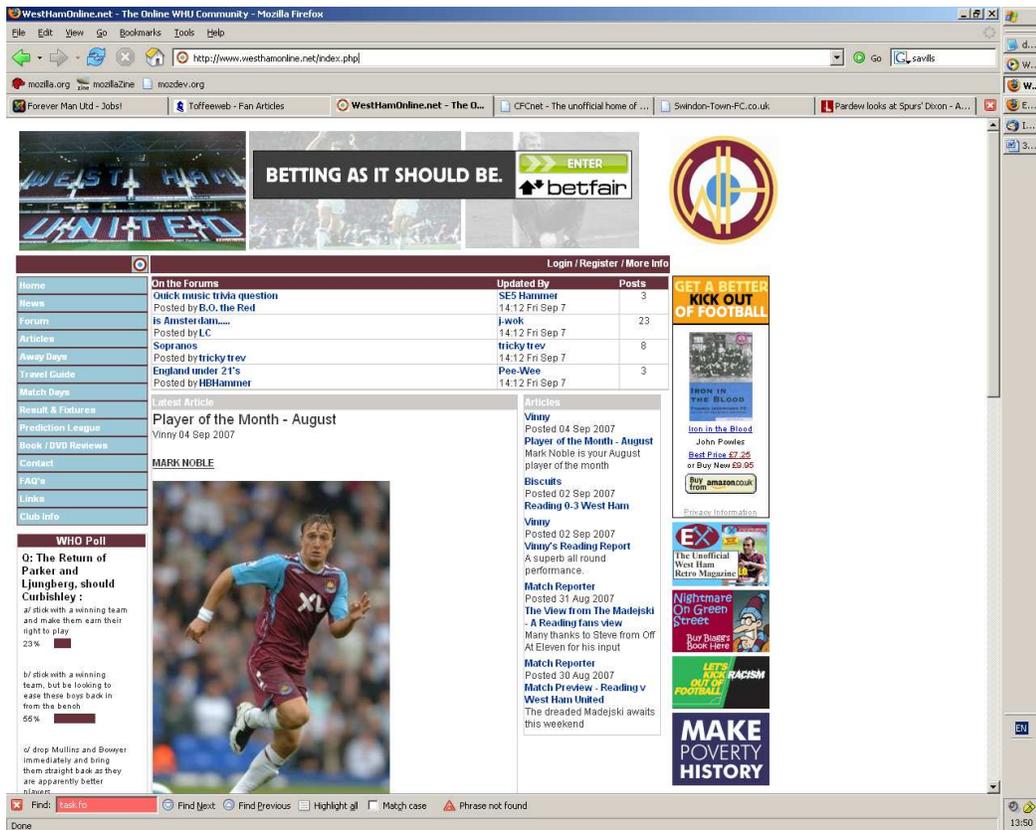


Figure 3.2 WestHamOnline.net (WHO) front page

Furthermore, in the top centre, immediately under the banner and advertisements, there seem to be stories posted by numerous individuals (as if on a forum), indicating that the site might be a collaborative effort (actually, the site does have all the attributes of a forum). The procedure dictates that having found no About or Help section, every relevant available link is clicked on (for details see 3.6.1 Website Examination Procedure, p. 113). On this site, the links include (clockwise from the top left hand menu) the entire menu on the left hand side (not including the poll), one link from the "On the Forums" section, one link from the "Updated By" section, the three links immediately to the top right of the Updated By section (Login/Register/More Info) and then an article from the "Articles" section.

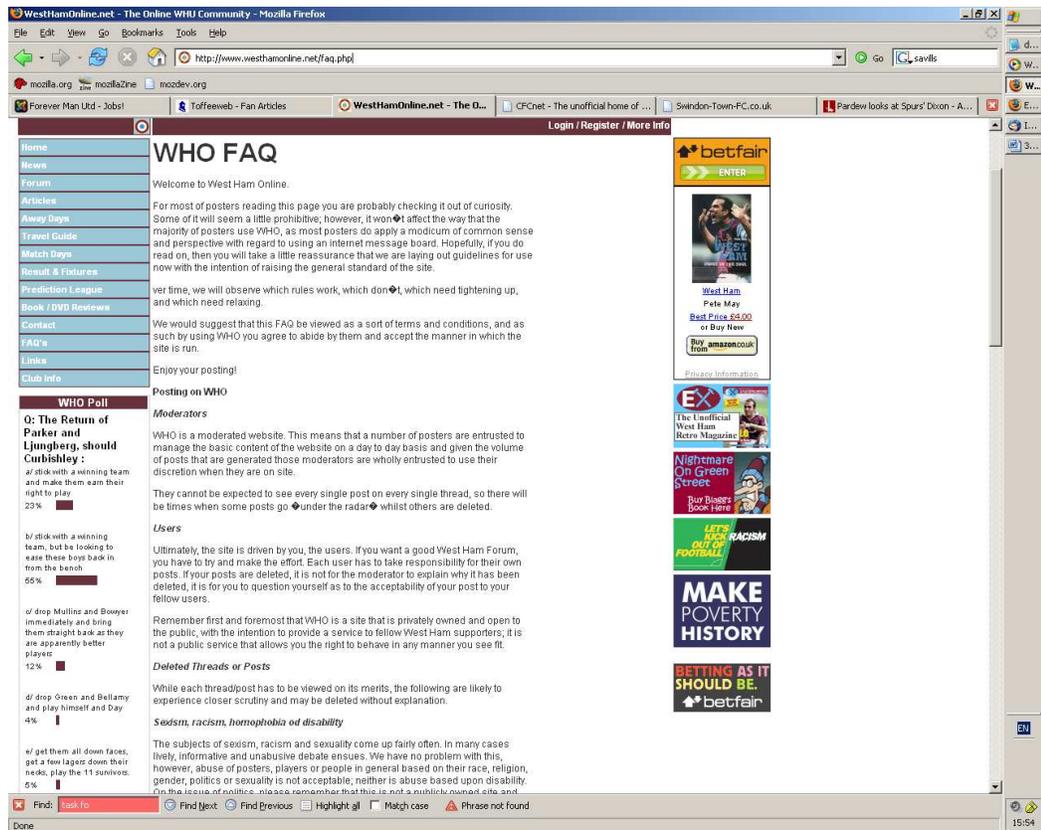


Figure 3.3 WHO FAQ page

Having clicked on all these links, certain discoveries were made. The most valuable information concerning the WHO was found in the FAQ section. Here it is clearly stated that the site is driven by the users.

"Ultimately, the site is driven by you, the users."

However, there are also references to the site as a forum, making it unclear as to whether or not it was possible to post an article directly to the front page.

For this purpose, the contact details of the site administrators were searched for, and emails were sent out to WHO, asking about whether it was possible to post articles directly to the front page, and also enquiring as to whether it was possible to add further sections to the site (e.g. expanding the menu on the left hand side). Replies

were received, stating that individuals with "Contributor" status could place an article automatically on the front page, though only a handful of people had the authority to add sections to the site.

3.7 Stage Two People-Study

The Stage Two People-Study was carried out in order to incorporate the opinions of web authors into this work. It was felt necessary to take into account these opinions for the purposes of having a basis from this perspective as well to make any recommendations (i.e. concerning the communal website recommendation). Previously, studies have solicited the opinions of web authors (e.g. Buten 1996, Maruyama 1999), however, this Stage Two People-Study contributes to the field by doing so from a subject based viewpoint. Hence with the additional information resulting from the use of this Stage Two People-Study, it was possible to make a recommendation which also satisfies, to a certain extent, the needs of the web authors.

Concerning the motivations of web authors, in terms of feasibility and the information necessary (for reasons mentioned in Stage One), the choice was between questionnaires and interviews. However, as web authors are available in significantly lower numbers than ordinary football fans, and the information retrieved from questionnaires is not as in depth as that from interviews, it was decided that interviews would be more suitable (in this instance, over the telephone). The need to carry out telephone interviews rather than face-to-face interviews was borne from the difficulty in locating willing web authors who lived in and around the London area.

3.7.1 Web Author Interviews

The interview questions for the web authors were designed much in the same way as the initial football fan interviews. Once more the interviews were semi-structured, though this time they were on the telephone, carried out with authors of football websites. For this purpose, telephone conversation recording equipment was purchased, as was a telephone headset, to allow for the recording of the interview both electronically on audio files, but also in the form of hand written notes.

Once more the first set of questions was contrived by thinking of all possibilities, and then reducing the list by removing similar, overlapping or irrelevant questions. In this instance, the guidelines for the questions were the outcome of the results of the earlier research. The structure and content of the questions, as well as the functionality of the telephone conversation recording equipment and telephone headset, were then tested by actually conducting semi-structured open-ended pilot interviews with suitable candidates. Once again, feedback was taken into account, new possible questions were discovered and changes were made. The whole procedure was again carried out thrice, after which time no more revisions were deemed necessary.

When it came to finding potential interviewees, initially, attempts were made to contact web authors of teams in London, so as to carry out face-to-face interviews. Emails were sent to addresses found on the websites of the authors, requesting an interview. However, it soon became apparent that most web authors of teams in London do not live in London. Furthermore, many of those contacted stated that they would rather carry out the interview via email or on the telephone. As a result of this, it was decided that the interviews be carried out via telephone. A total of 188 emails were sent out to potential interviewees, of which 100 were replied to positively. However, it proved very difficult to actually arrange the interviews, with candidates constantly postponing possible dates. In the end, a total of 11 web authors were interviewed over a period of two months, from July 2007 to August 2007. Here, it is acknowledged that the conversion rate is quite low, and the idea of using email interviews as well as the telephone interviews was considered. However, as the purpose of the interviews was “exploratory” (so as to sample opinions primarily on the viability of the communal site), it was not felt necessary to carry this out.

As in the earlier interviews, friendly, informal, personalised and customised emails (with interviewees addressed using their first names) were sent to potential interviewees asking them to participate voluntarily in the research. The aim of the email was to maximise the possibility of the potential interviewee being comfortable with the idea and willing to participate. Once again, the aim of the research was briefly mentioned, along with the confidentiality of the data and the approximate

anticipated length of the interview. It was made clear that the request was for a telephone interview, and that the date and time were flexible.

The interviews were then carried out at the mutually agreed times. Prior to the commencement of the interview, a brief introduction on the nature of the research was given, and the interviewee was asked if they objected to the use of the recording equipment or the taking of hand written notes (this was possible in conjunction with the use of the headset). The interview was then carried out and recorded, using both the recording equipment and hand written notes. In length, the web author interviews were between 25 and 55 minutes long, with the average length of the interview at approximately 35 minutes long.

The recording and storage of the interview data was carried out exactly as it was for the Stage One People-Study, using hand written notes which we rewritten electronically, stored and managed using MS Excel and SPSS (see 3.3.1 Football Fan Interviews, p. 73).

3.7.2 Web Author Interview Questions

The motivations of web authors were looked into in order to decipher whether or not they would fit in with the ideas behind the communal site that was being designed. In order to craft an outline for a communal site that could encourage web authors to create web pages for the online community, it was felt important to consult existing web authors.

All of these questions were deemed necessary in order to have an understanding of potential contributors to any communal site, with a view to a recommendation that can benefit both online website users and contributors.

Questions concerning the motivations of web authors were asked so as to ensure that any proposal would take into account their motivational needs. In the same way, questions concerning their difficulties were asked so as to ensure that any proposal would not unduly create additional difficulties. Website popularity, user participation

and possible improvement questions were all asked with a view to gathering ideas that would ensure that the site would be of interest to football fans, and an improvement on what existed before.

The earlier Stage One People-Study conducted already involved certain individuals who would partake in various online activities as contributors (e.g. forum users), however it was felt potentially useful to look into individuals who were explicitly web authors. Such activity was felt missing in the literature of the field in general. Though questionnaires have been sent out to PHP authors (e.g. Buten 1996), at the time of writing the author was not able to locate a study in which web authors in a specific subject area were interviewed.

Below is a section by section break down of the questions asked.

Personal information

This section just acquired basic knowledge about the interviewee. The name, age, gender, occupation, place of residence and contact details were asked. Though not strictly necessary at this stage, these questions reveal the interviewee demographics.

Website motivations

Questions concerning website creation were asked here. Namely, why the author created a site, what motivates them to keep maintaining the site, and what could reduce this motivation. This section was included so as to provide vital insight into phenomena that might encourage the creation of sites.

Website difficulties

As the section name suggests, the questions here were about the obstacles that the web authors had to overcome. Questions concerned what resources the web authors lacked, the causes for website closures and competition from other sites. This section was included, this time to provide vital insight into phenomena that might discourage the creation of sites.

Possible improvements

This section asked the web authors what they felt could help facilitate the improvement of their sites, whether they believed collaboration could benefit their site and if so, what the benefits might be. These questions attempted to discover what authors themselves feel might be useful in facilitating an improvement for them, and whether they themselves are interested in further collaboration as a means to achieve their goals.

Site popularity and user participation

What makes the site popular, what encourages reader participation and if special attention was paid to these aspects (so as to ensure that the popularity of the sites was maintained) were questions asked in this section. The data gathered from these questions helped provide focus onto aspects that should be paid particular attention to.

Final questions

Here the interviewee was asked whether they felt any relevant questions were missing, so as to ensure that vital questions were not left unasked.

Having carried out both the website investigations and further interviews with web authors, Stage Two is complete.

3.8 Evaluation and Replicability

Finally, in order for a study of this kind to be effective, it is crucial to undertake some evaluation and ensure that any processes can be replicated.

In truth, the evaluation of such a process as a whole is difficult. Normally, such processes can be compared to other related processes in other studies, though with no related studies carrying out such processes, this path could not be taken (see 2.4 The PHP as an Object of Study, p. 36). However, the exact procedure for the Stage One Web-Study was tested for meeting the needs of the objectives and replicability in an initial pilot study, the results of which have been published in a peer-reviewed journal

(Narsesian and Nicholas 2005). Thus, the problem of a lack of literature and methodologies to compare with is to a certain extent dealt with through the publication of the process.

In terms of sampling, there are two points to consider. Firstly, as it was not possible (see 3.4 Stage One Web-Study, p. 86) to define teams by “size”, the teams for which PHPs were chosen, were themselves chosen “randomly” after having been found based on subject. This in itself is hard to compare with other studies, as no other study has (to date) chosen PHPs based on subject. This random selection was done by assigning numbers to teams (in alphabetical order as they would be placed in the league before the commencement of the season, where the first team is numbered one) and then using a random number generator to generate numbers (random.org). Secondly, a number of types of PHPs have been disregarded. Namely, those with too many broken links, those in foreign languages and those to which access was restricted. Though it can be argued that these are a legitimate subset of PHPs and should be included, for both feasibility and theoretical, logical reasons, they have not been included. In terms of feasibility, for restricted and foreign sites, it simply was not possible to include the sites excluded. As for the “broken” PHPs, there must be a point at which a website is considered a website (in any definition). For sites where the majority of the pages are not functional, it becomes questionable as to whether the average Internet user would themselves consider such a site to be a website. Since this thesis look at sites from the point of view of an average football fan, it was thought logical to choose a point at which fans themselves might disregard a site. Finally, another important point concerning the sampling of sites is that other studies have also excluded PHPs from parts of their studies for such reasons, because they had restricted access (Doring 2002), or because they were incomplete (Dominick 1999). Foreign language PHPs have not been mentioned in such contexts because the sampling methods used by other studies have not chosen PHPs by subject (thereby not having to contend with foreign language PHPs).

In terms of replicability, two issues have been addressed. The first of these is the saving of the versions of the sites that were checked. After various attempts to use software to save every site (being examined) failed (see Appendix E), it was decided to manually save to disk two of the sites checked (this was carried out during the pilot

study). These sites were then used to resolve the second consistency issue. As significant sections of this study are conducted manually, there is an inherent risk that different coders might interpret sites differently. Towards that end, these sections of the pilot study (that were saved manually) were used to carry out the study again, once by the original coder with a time lag (6 months), and once by a completely separate coder. The coders tested this procedure from start to end, using these saved sites. The results can be found in Appendix E.

3.9 Chapter Conclusion

Having examined the available literature in the previous chapter, which looked not only at PHP studies, but also other relevant fields, it was determined that the most thorough approach would be one that looked at the PHP as one of many tools and methods for the acquisition of information. Towards that end it was deemed suitable to look specifically at PHPs, but also to take into consideration the thoughts and ideas of football fans and web authors. However, for the examination of the PHPs, no comparable study was available from which to take central ideas, a situation which resulted in the creation of the Stage One Web-Study, a wholly originally approach to PHP content examination, which is considered a contribution of this thesis. The use of this methodology resulted in the meeting of the aims of this project, which is the first such project to look at PHPs from a subject based viewpoint, furthering the works of Dominick (1999), Doring (2002) and Papacharissi (2002a, 2002b) among others. Furthermore, it is the belief of the author that further original contributions can be made from the re-application of this methodology in other fields.

In addition to the Stage One Web-Study, interviews were carried out and questionnaires were sent out (in both Stage One and Stage Two) to canvass the opinions of football fans and web authors. Stage Two also examined collaborative websites so as to aid in the formulation of a recommendation, another of the project's original contributions, something which was felt to be necessary to fully meet the aims of this thesis. The end result is an overall methodology that provides a holistic view of the information seeking behaviour of football fans of the top four professional English leagues with a particular emphasis on PHPs. The following three chapters

look at the results of the implementation of this methodology, starting with the first five objectives.

Chapter IV

Results I *- Information Seeking Tools of Football Fans*

4.1 Introduction

Earlier in this work, a number of aims and objectives (see 1.4 Aims and Objectives, p. 17) were mentioned. The first set of objectives concern the informational behaviour of football fans. Broadly speaking, these objectives look at "how and why" football fans look for information concerning their teams in the ways that they do. This chapter discusses that first set of objectives:

1. To determine the tools and methods ordinary football fans use to acquire their "footballing" information.
2. To determine the reasons for which these tools and methods are used.
3. To determine the significance of the role played by the web in providing this "footballing" information.
4. To determine the significance of the role played by the PHP within the role of the web in providing this "footballing" information.
5. To determine whether, all in all, football fans are satisfied with the availability of "footballing" information on the web.

The results to these objectives were derived from the interviews carried out in the Stage One Web-Study (30 in total - see Appendix A and Appendix B), the responses to the online questionnaires (369 responses - see Appendix F) and offline questionnaires (51 responses - Appendix G). For the greater part, the questionnaire results have been shown separately, so as to be able to note any potential bias created

by having distributed most questionnaires online. However, where appropriate the questionnaire results have been merged (i.e. where the online distribution of questionnaires does not affect the results).

4.2 Demographics

As mentioned earlier in this work (see 2.6.1 Football Surveys, p. 57), other studies have carried out their own surveys looking at the same fan base. Here, the demographics from the data collection carried out for this thesis and those surveys are compared to give an indication as to how closely the population samples match. In each instance, the relevant parts of the data from this thesis are compared so as to compare like with like. Two separate sets of data are taken from the results here, and compared to two external surveys- the FA Premier League National Fan Survey (2005) and the Football League Supporters Survey (2006).⁵⁷

4.2.1 Premiership Demographics

For this study, in terms of age, the category with the greatest number of Premiership respondents is the 25-34 age group (32.1%). In fact, 75.8% of respondents were fans aged 16 to 44. Fans aged 45 and over provided the remaining responses. The FA Premier League (2005) survey points to fans being more middle aged. For this study, as the majority of the responses for the Premiership came from the online questionnaire, it is quite likely that the age of the average respondent was lowered, thereby explaining this age difference in the demographics. Perhaps the enthusiasm of youth, in combination with the more free time youths have in front of the computer, might account for the reason responses are so heavily clustered towards the younger age groups. This very same age group would also explain why the average income⁵⁸ of fans in this survey (£26.9k) is so much lower than the FA Premier League's figure (£36.5k). Table 4.1 shows the demographic details of both surveys.

⁵⁷ It must be noted that the research methods used by the Premier and Football Leagues were self selection as well, and as such, they too are not perfectly representative of their fans.

⁵⁸ Respondents were asked to fill in their income so as to allow a demographic comparison with major surveys.

Age⁵⁹	FA Premier League National Fan Survey (n=25,539)	Questionnaire - Premier League only (n=112)
16-24	7%	22.3%
25-34	17%	32.1%
35-44	29%	21.4%
45-54	25%	12.5%
55-64	14%	8%
65+	7%	3.6%
Gender		
Female	15%	13.5%
Male	85%	86.5%
Socio-Economic Grouping		
Upper Middle (AB)	46%	40.2%
Lower Middle (C1)	25%	31.3%
Skilled Working (C2)	19%	7.1%
Working Class/Subsistence (DE)	10%	21.4%
Income		
Income	£36.5k	£26.9k

Table 4.1 The FA Premier League National Fan Survey (2005) and this study's questionnaire (online and offline combined, showing the Premier League only).

The socio-economic⁶⁰ (see Figure 4.1) and the gender groups, match the FA Premier League figures more closely. The exceptions are the figures for the C2 and DE classes in the socio-economic group (C2 being at 19% against 7.1% and DE being at 10% against 21.4%). The reason for this is most likely due to the particular sample, which happens to be demographically weighted in this way (see 3.3.3 Questionnaires, p. 79).

⁵⁹ One respondent did not fill in their age group.

⁶⁰ The socio-economic groupings were worked out by looking at guidelines provided by the Office for National Statistics.

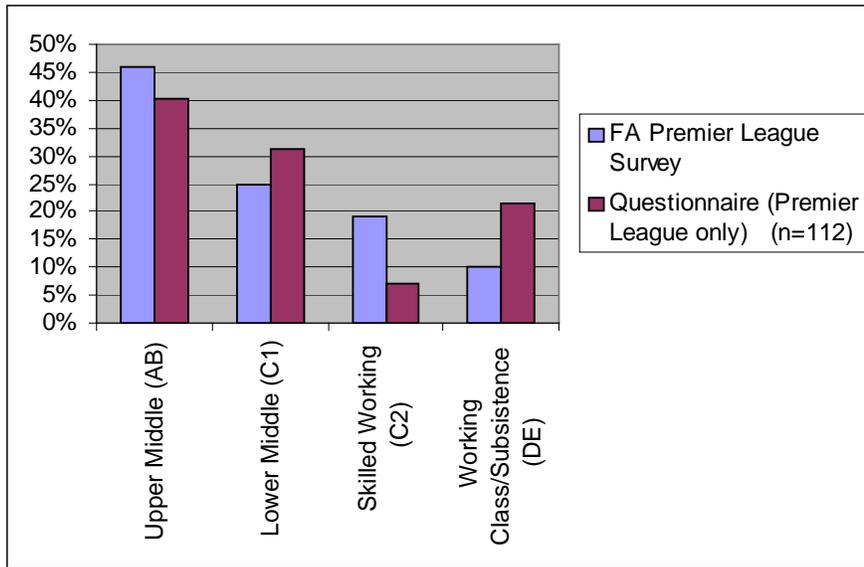


Figure 4.1 Premier League demographics - socio-economic grouping

4.2.2 Football League Demographics

Here, the Football League respondents in the survey conducted in this thesis is compared to the Football League's (2006) own survey. As is immediately obvious from Table 4.2, there are differences between the presentation of the Football League survey results and the Premier League survey results (different age groupings are used). As the age groupings for the football league were more "compressed" (i.e. the Football League 2006 survey questionnaire has just three groupings instead of the 6 the Premier League has used), the figures from this thesis match reasonably well. Again, the respondents to the questionnaires are quite young, with 61.4% in the 16 to 34 age group. On this occasion however, this is not that different from the corresponding Football League figure of 50%. The 35-54 age groups match even more closely (33.7% to 35%), with a wider gap noticeable in the 55+ category. Perhaps the most probable explanation for this gap comes from the fact that the respondents to the online questionnaire came from Internet forums, where it is quite likely that the average age of users is relatively low (i.e. there are few users above 55).

Age	Football League Supporter survey (n=43,590)	Questionnaire - Football League only (n=308)
16-34	50%	61.4%
35-54	35%	33.7%
55+	15%	4.9%
Gender		
Male	80%	91.5%
Female	20%	8.5%
Socio-Economic Grouping		
Upper Middle (AB)	36%	45.2%
Lower Middle (C1)	44%	37.4%
Skilled Working (C2)	10%	8%
Working Class/Subsistence (DE)	6%	9.4%
Income		
Income	£38k	£27k

Table 4.2 Football League Supporter survey (2006) and questionnaire (Football League only).

The socio-economic categories again match reasonably well (see Figure 4.2), with small variations in each category. Once again, the reason for the lower income of our respondents is most probably due to the overall younger age of our sample.

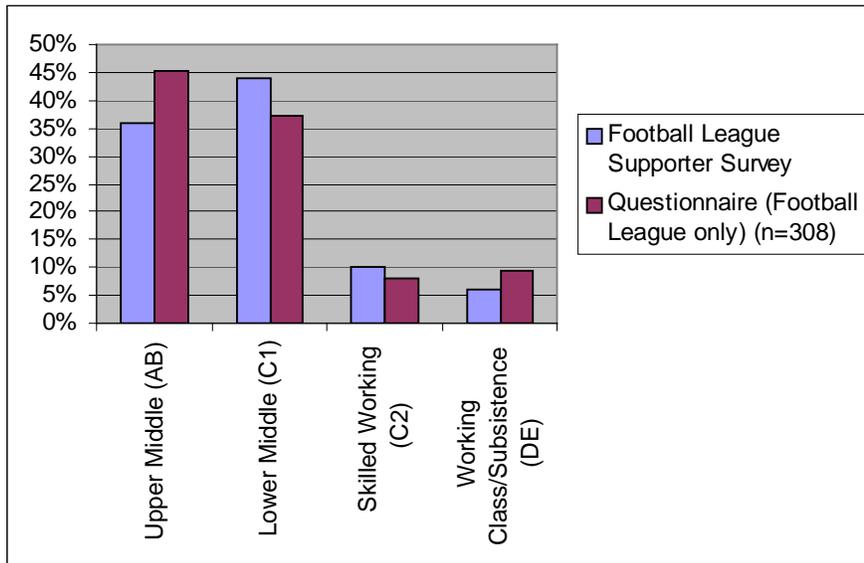


Figure 4.2 Football League demographics

4.2.3 Comparisons Within This Study

Having shown how the demographics of this study's surveys match with other studies, one more table has been included, this time to illustrate the demographics of fans within this thesis, comparing the online questionnaires, offline questionnaires and interviews. The number of interviewees in the Stage One Web-Study study (30) is clearly not large enough to produce the expected demographic figures, but the column has been included as a general guide. To help illustrate how the figures match, a graph has also been included (see Figure 4.3).

Age	Online Questionnaire (n=369)	Offline Questionnaire (n=51)	Interviews ⁶¹ (n = 30)
16-24	38.5%	13.7%	16.7%
25-34	21.4%	41.2%	40.0%
35-44	21.1%	17.6%	26.7%
45-54	13.3%	9.8%	10.0%
55-64	3.8%	11.8%	3.3%
65+	1.4%	5.9%	3.3%
Blank	0.5%	0.0%	0.0%
Gender			
Male	91.3%	78.4%	80.0%
Female	8.1%	21.6%	20.0%
Blank	0.6%	0.0%	0.0%
Socio-Economic Groups			
Upper Middle (AB)	43.6%	37.3%	56.7%
Lower Middle (C1)	35.8%	29.4%	36.7%
Skilled Working (C2)	8.1%	3.9%	3.3%
Working Class / Subsistence (DE)	9.2%	29.4%	3.3%
Blank	3.3%	0.0%	0.0%
Income			
Income	£28.2k	£26.7k	N/A

Table 4.3 Demographics of the two sets of questionnaires and the interviewees

All in all, the demographic figures from the three data collection methods match reasonably well. Although the figures have some discrepancies, they do not appear to be significantly biased one way or the other. The youngest age group is clearly prominent on the online questionnaires, however, that is the only major difference.

⁶¹ Figures of the interviewee responses have still been included in tables where the interviewees were asked questions relevant to the table in question.

Other more minor differences are also present (e.g. the DE group of the Socio-economic categories), however, the trends overall match in all the demographic categories as can be seen in Figure 4.3. The interview demographics also match reasonably well, even though (as already mentioned) the number of interviewees was significantly lower than the questionnaires. As an example, a graph has been created to show the match in terms of socio-economic groupings (Figure 4.3).

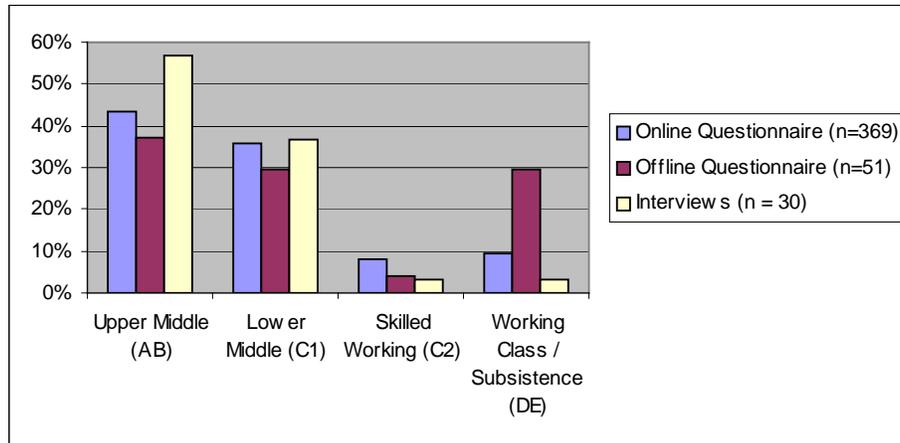


Figure 4.3 Questionnaires and interviews - socio-economic groupings

Overall, as the demographics of the respondents in all categories for this study do not appear to be heavily or inexplicably biased in any single direction, it is accurate to say that a foundation on which the objectives (discussed in the following section) can be met has been provided.

4.3 The Objectives

4.3.1 Tools and Methods Used for Acquiring Football Related Information

As stated, the purpose of this first objective was to find out what tools and methods football fans used to find information about their team. After initial studies, a list of possible tools was compiled. This list was used firstly in the interviews, and then in the questionnaires. By the end of the data collection, two more methods used by fans

were discovered. Table 4.4 shows the percentage of respondents that use the various tools and methods available according to each data collection method.

Tool/Method Name	Online (%) (369 respondents)	Offline (%) (51 respondents)	Interviews (%) (30 interviewees)	Total (%) (450 respondents)
Websites	86.7	72.5	93.3	85.5
Friends/Colleagues	39.0	56.9	70.0	43.1
Reference books	28.5	23.5	6.7	26.5
TV	13.8	21.6	70.0	18.4
Local newspapers	13.6	13.7	46.7	15.8
Teletext	11.9	23.5	13.3	13.3
National newspapers	11.9	15.7	46.7	14.7
Radio	8.7	9.8	46.7	11.4
Fanzines	8.4	5.9	20.0	8.9
DVDs/videos	5.7	3.9	26.7	6.9
Magazines	4.6	3.9	26.7	6.0
Text alerts	2.2	2.0	0.0	2.0
Mobile Internet	2.2	2.0	36.7	4.5
Shareholder reports*	0.0	0.0	6.6	0.2
Computer games*	0.0	0.0	3.3	0.2

*Discovered in the later stages of the interviews.

Table 4.4 Percentage of respondents using tools or methods for information seeking

From the very beginning, it was immediately clear, that in fact, the web plays a very central role in providing information to football fans. All the data collection methods used in this study, indicated that football fans of English football are generally heavy Internet users. In this instance, in all three data collection methods, it is strongly apparent that websites are at the top of the lists. The online questionnaire has 86.7% of respondents saying they use websites, with the offline questionnaire reporting 72.5%, and the interviews at 93.3%. In all the figures, websites are clearly the first choice for the respondents, with the second most popular method being at least 15.6% away (offline questionnaires) and at most a staggering 47.7% away (online questionnaires). The predominance of the websites as a source of information will be discussed in more detail later in this chapter (see 4.3.3 The Role Played by the Web, p. 142).

Other noteworthy aspects of the table are some rather conspicuous differences in responses between the three data collection methods (e.g. the figures for TV and Local Newspapers – seen clearly in Figure 4.4). One explanation might be the levels

of concentration or enthusiasm in filling out questionnaires and taking part in interviews. It is possible that the interviewees were eager to be as helpful as possible, whilst the questionnaire filling, as a more commonly undertaken task, is carried out with somewhat less enthusiasm.⁶²

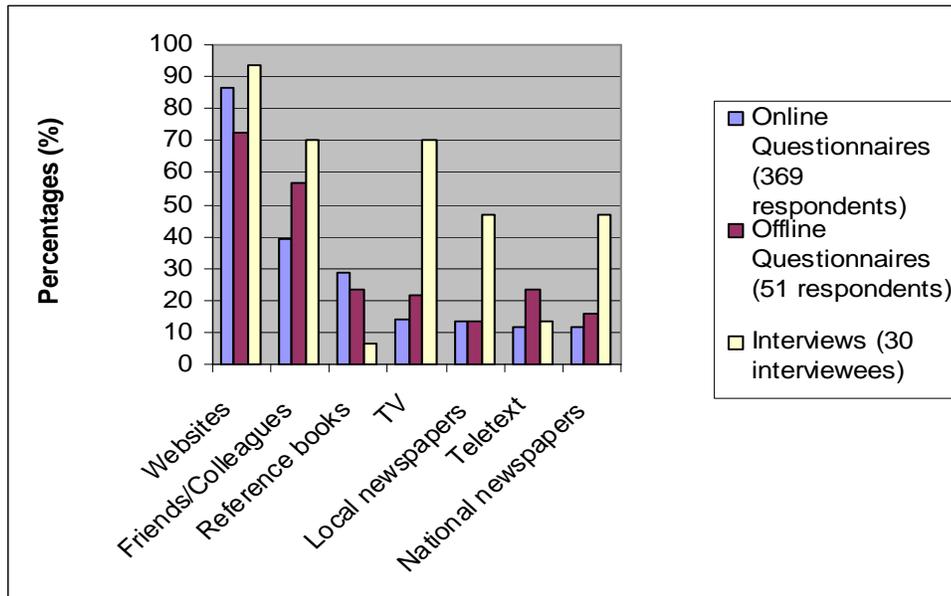


Figure 4.4 Notable tools or methods used by football fans for information seeking

The second most popular source of information for football fans seemed to be friends and colleagues. This is very clear in both sets of questionnaires (39% online and 56.9% offline), but somewhat less clear in the interviews, as many of the methods seemed popular (friends/colleagues was at 70%, but TV was also at 70%). These figures make even more sense when viewed in conjunction with Table 4.5, which shows the amount of time people spend talking about their team to their friends.

⁶² The interviewee would often try to provide information which was as complete as possible, possibly for fear of a follow up question or request for clarification. This would explain the general reduction in figures going from interviewee, to offline questionnaires, to online questionnaires, with the reduction in likelihood of a request for an explanation (note that the researcher was present in the filling of the offline questionnaires).

Time spent per day	Online (%)	Offline (%)	Interview (%)
Under 10 mins	5.4	25.5	20
10 mins - 1 hr	41.2	41.2	60
Over 1hr - 2 hrs	30.1	21.6	16.7
Over 2hrs to 4 hrs	15.4	5.9	0
Over 4 hrs	5.7	2.0	3.3
Don't Know	2.2	3.8	0
Total	100.0	100.0	100.0

Table 4.5 Time spent per day by respondents talking about football (percentages)

The vast majorities in each column (92.4% online, 70.6% offline and 80% interviews) spend at least 10 minutes a day talking about their team to their friends and colleagues, and the number of people spending more than an hour (or longer) then trails off, something that is illustrated clearly in Figure 4.5. This further strengthens the case that fans acquire information from friends and colleagues.

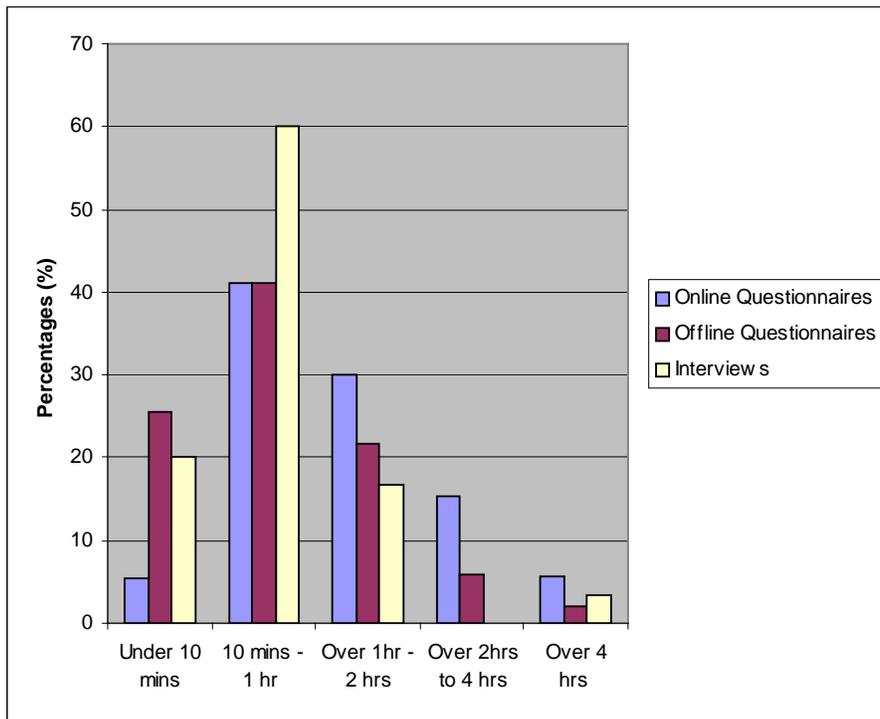


Figure 4.5 Time spent talking about football (per day)

Additional insight into this phenomenon was provided through comments by the interviewees:

"I need to know about the younger players in case my friends or people at work ask me." Jami-ARS

"I think it's a social event as much as anything else so I didn't want to buy a season ticket on my own, what's the point? So it was only when my daughter was old enough that we could go together that it, becoming a season ticket holder, became something that I would consider." Gran-NUFC

"I spend quite a long time talking to my daughter about it and to my father-in-law." Gran-NUFC

"I get more information by talking to the people at the match, and also from overhearing people talk at the match." Macl-NUFC

Even further reinforcement of this view is supplied by numerous comments about the social aspect of football. Various terms are used to describe this aspect of football fandom, all with similar implications. Terms such as community, tribe, religion, culture, football family and male bonding were used. Clearly, talking is a fairly predominant activity in most forms of social activity, including the ones mentioned in the following quotes:

"[football is] communal for men." Ze-LIV

"...feels like a family, which is as important as the actual football aspect."
Emap-BREN

"Chelsea is my religion." Mile-CHE

"Football is a big part of English culture. It was a big part of my childhood."
Makr-ARS

"Football is a social activity." Beto-CHE

"Football unites communities." Sylv-CHAR

Next in popularity on the list were reference books. These are available for all football teams, and several of the interviewees claimed to possess them, whether purchasing one themselves or receiving them as gifts. Clearly, the information found in all the books is not identical, but broadly speaking, these books will have plentiful statistics on past and present players, as well as some featured stories, players and so on. Some of the interviewees mentioned that their use would be for very obscure or old facts.

"I use the reference book for obscure stuff or old stuff." Frei-FUL

"...if I wanted to know who was the third highest scorer in the league in 1975." Mile-CHE

Numerous other sources of information in paper form were also on the list including local newspapers, national newspapers, fanzines and magazines. The popularity of local newspapers vis-à-vis national newspapers depends on the overall popularity of the team in question, and the area (geographically speaking) that the fan is living in. Local papers are generally available only to locals, so if the fans do not support a local team, they would probably have trouble accessing their team's local paper. At the same time, if the team the fan supports is a historically popular one (e.g. Manchester United), this is not such an issue. As the national papers have such extensive coverage of major teams, fans can get similar information about their teams without the need for the local papers. Such a theory might also apply to a certain degree to fanzines and magazines. Supporters of teams in the lower divisions are more likely to want to read fanzines about their teams, as the amount of coverage on their team elsewhere (e.g. in generic magazines) is less, whereas for those who support historically popular teams, this is not such an issue because of the abundance of information available. Overall, for fans of historically popular teams, fanzines and local newspapers are less important, whereas for fans of less popular teams, magazines and national newspapers are less important.

Meanwhile, common information disseminating tools such as TV (including teletext), radio, DVDs and videos were always going to have a certain degree of popularity, as well established appliances in every household. With so many other sources of information, and their role of late being primarily for entertainment rather than information dissemination (excluding teletext) purposes, their levels of popularity seem commonsensical.

Perhaps the surprises in the list are the text alerts (2.2% online, 2% offline and 0% interviews) and mobile phone Internet (2.2% online, 2% offline and 36.7% interviews), both of which were remarkably unpopular. With the mobile phone industry boom of the previous years, one would have expected the mobile phone to play a more active role in the football industry.

Finally, at the bottom of the table, two new sources of information were revealed, both of which were acquired through the interviews. The first was the annual shareholder report (Cook-LEYT, Brad-LIV) that club shareholders receive and the second was a computer game (Dion-LIV).

A football club's annual report will contain various pieces of information. Traditionally, the information found in a report will indicate how the club has done in the year that has just passed and how it expects to do in the coming year. The report will also mention what the club has tried to achieve and what the club will try to achieve. Additional information includes the salaries paid to senior management, along with justifications, accounts and operational summaries for the club as a whole. To give a clearer picture, the table of contents from a football club's annual report has been included in Table 4.6.

In terms of accessibility, if the club is a public limited company (PLC), then the annual report will be publicly available (as was the case with Newcastle United FC in 2006). However, the ownership of the clubs in the top four professional leagues in England is varied (certain clubs are owned publicly, others privately). As an example, the Annual Report of Fulham FC is not publicly available. In this instance, fans who own shares in a privately owned company (which can still have widely distributed shares) have access to information that those outside the company might not.

As this source of information was discovered in the later stages of the interviews (after the questionnaires had been designed and distributed), it was not possible to further gauge its popularity. However, the shareholder's report is primarily a business document. This coupled with the few mentions in the interviews or other research, indicates that for ordinary football fans at least, it is unlikely to be a popular source of information.

Contents	Page number
Financial highlights	1
Operational highlights	1
Chairman's statement	2
Financial review	8
Directors and senior management	10
Directors' report	11
Directors' remuneration report	14
Statement of corporate governance	18
Independent auditors report.	21
Group profit and loss account	22
Group balance sheet	23
Company balance sheet	24
Group cash flow statement	25
Reconciliations of movements in shareholders' funds	25
Notes to the accounts	26
Notice of annual general meeting	45
Corporate information	47
NUFC contacts	48

Table 4.6. Contents of the Newcastle United FC Annual Report (2005/06)

The second tool revealed as an information resource was a computer game called Pro Evolution Soccer.

Pro Evolution Soccer (Pro Evo) is a game created by Konami, the version of which was discussed during the interviews, is commonly referred to as Pro Evo 5. The game is a football simulation game, where the player has the ability manage a team and play the games in a league of his choice. Certain players' and teams' real names are used (those sponsored by Adidas), and in other cases a similar name of the player or the team is used. As an example, Liverpool FC is referred to as Merseyside Red, and Dutch striker Van Nistelrooy is known as Von Mistelroum. Additional software

updates though are released by unofficial parties which can override these fictional names with real names. Information and statistics are available on each player.

Below is an exact quotation from the interviewee:

"In terms of gathering information, something a bit outside the interview, I'd like to say that I get a lot of information about the players from Pro Evo, the computer game. Football games have a remarkable amount of information that is not available on the web. I mean, some things, like the finishing⁶³ of the player is 78%, well, that's not measurable, so we can't say how correct it is. Other things though, what position he can play, whether he's left or right footed, languages he can speak etc. Even the 78% which cannot be accurate scientifically speaking, does give an insight into the perception of people on the abilities of the player. The accuracy of the information is not guaranteed, but I imagine the makers do a lot of research, as they want the game to be as realistic as possible, and with the millions of pounds spent making the game, it's quite probable that the information is very close to correct, even if not 100%." Dion-LIV

The information in this quote was verified using Pro Evo 5 on a Sony PlayStation 2. All the information was found to be accurate, with the exception of the languages the player can speak, though that information is available on other football computer games (e.g. Sega Football Manager 2006), but not Pro Evo 5.

The gathering of information using a computer game is certainly unusual. There was no mention of this in any of the other data collection methods utilised in this thesis. However, at the same time, it is unlikely that the interviewee is the only individual who uses computer games in such a manner. As such, the interviewee's comments have been taken into account, and a new tool for the acquisition of information has been recorded, though like the shareholder's report, it is quite unlikely that it is a primary source of information for many fans.

⁶³ Finishing refers to a player's ability to score a goal.

Overall, it appears that the web is by far the most popular tool or method used to acquire footballing information. This is perhaps unsurprising as other studies have also indicated high Internet use amongst supporters (FA Premier League 2005, Football League Survey 2006), however, this is the first time that the use of websites has been looked at within a broader view, looking at all possible tools and methods for the acquisition of footballing information. As a source on information, websites are clearly some way ahead of any other tool or method. The next most popular method for acquiring footballing information was through friends and colleagues. This finding is also not greatly surprising, as football is said to have played an important role in English society (Fishwick 1989) and various studies have been carried out which examined various of its social and community aspects (FA Premier League Survey 2005, Football League Survey 2006, Jaquiss 2003, Tacon 2005). One additional notable finding which came about as a result of meeting objective 1 was the discovery of individuals who attained footballing information in methods unmentioned before, notably computer games.

Though not entirely unexpected, these results were positive from the viewpoint of this study. Certainly, the high instance of activity concerning websites facilitates the study of PHPs in this area. This, combined with the findings concerning the communal aspects of information behaviour, provided additional motivation to continue along the chosen path for this thesis, looking into PHPs and how their benefits can be maximised through a communal effort.

4.3.2 Reasons for the Use of Particular Tools and Methods

The questionnaires included a question which asked how heavily respondents used certain tools and methods, and in conjunction asked the reason for this. Table 4.7 shows the reasons why heavy users (those that gave the rating for the highest use on a five point scale) use the methods they use.

	Auth/ rity	Compl/ ness	Conv/ ence	Most up- to-date	Only Source	Other	Don't Know	Total
Friends	2	5	20	4	2	5	1	39
Websites	7	19	53	134	12	1	28	254
Radio	1	0	13	8	3	0	6	31
Teletext	5	1	13	12	0	3	7	41
TV	8	2	13	20	0	1	8	52
Loc News	11	11	11	14	6	2	12	67
Nat News	4	8	7	4	1	1	6	31
Fanzines	0	4	6	6	4	1	7	28
Magazines	0	1	0	1	1	1	1	5
Ref Books	3	2	1	0	0	0	0	6
DVD/Video	2	1	1	0	0	0	1	5
Text Alert	2	1	3	7	0	0	2	15
Mob Inet	0	0	9	3	2	1	3	18
Total	45	55	150	213	31	16	82	592

Table 4.7 Reasons given by heavy users of any method for the use of that method (online and offline combined).

The prominent columns in Table 4.7 are the most up-to-date (213) and convenience (150) columns. It appears that these are the two most important factors in determining the tool to use in the search for footballing information. These are then followed by completeness (55), authority (45) followed by only source (31). When looking at the tools or methods individually, there are some minor variations, but as expected, up-to-date and convenient information seem to also feature strongly.

Looking at websites is done primarily because it is the most up-to-date information source (52.8%) followed by convenience (20.9%) and completeness (7.5%). More information on the domination of the web as a source of information is available later in the chapter (see 4.3.3 The Role Played by the Web, p. 142). Other tools and methods which are used primarily for up-to-date purposes include text alerts (50%), local newspapers (20.9%) and TV (38.5%).

As for the remaining individual tools and methods, talking to friends or colleagues is predominantly done for convenience (51.3%) as is using mobile Internet (50%) and listening to the radio (41.9%). With Teletext, again convenience is more popular as a

reason with up-to-date a close second (31.7% and 29.3% respectively). Finally, fanzines are used equally for convenience and up-to-date purposes (both at 21.4%).

The three remaining tools and methods were used heavily by less than 10 respondents, making any conclusions difficult to come to.

Having inspected these figures, there can be no doubt that convenience and the need for up-to-date information are the strongest motivators for the use of any tool or method for the purpose of acquiring footballing information. This is something that fits with findings of past studies on different user groups, which have also shown that convenience and timeliness are important factors in preferring one type of information resource over the others (Ramos Linscheld and Schafer 2003, Whitmire 2004, Vondracek 2007, Tenopir King and Bush 2004).

4.3.3 The Role Played by the Web

Initial "feeler" talks with football fans were the first to indicate how involved they were with the Internet. These early indications were that football fans who were not web active were actually missing out on a critical resource and that they were one step behind web active fans (Narsesian 2004). Further evidence was not far away, as interviewee after interviewee talked about how web active they were. In fact, two interviewees mentioned that they had been reprimanded at work concerning the amount of time they spent looking at football websites.

Clearly, the main tool for acquiring footballing information is the web. This ranks as the most popular tool on both the online (86.7%⁶⁴) and offline questionnaires (72.5%), as well as the interviews (93.3%). These figures have already been mentioned in this chapter (see 4.3.1 Tools and Methods Used for Acquiring Football Related Information, p. 131). However, there is no shortage of additional data to reinforce this idea.

⁶⁴ It is a curious statistic that 86.7% of people who use football forums say that they use websites to acquire footballing information. One explanation might be that people consider their use of football forums to be for purely entertainment purposes, not involving the acquisition of particular information. Another less likely, but also possible explanation is that some of the respondents heard about the questionnaire from friends who use the forums. It was not possible to determine the exact cause for this peculiarity.

Figure 4.6 shows the tools or methods that football fans use heavily (i.e. respondents who selected the strongest option on a five point scale). Once again, the websites are the predominant force, something that is particularly clear in the figure. In this instance, respondents were allowed to make more than one choice, so the percentages do not add up to 100%. According to the online questionnaire, 63.4% of fans use the web heavily. The second most popular tool is the local newspaper at 16.0% (47.4% less than the websites). The offline questionnaire has web usage lower, at 39.2%, followed more closely by friends/colleagues at 29.4%. The differences are significant, and for an argument that the offline questionnaire sample is small, there is the counterargument of the biased sampling of the online questionnaire, which was, needless to say, found on a website. However, all in all, web usage is very strong as has been shown in other surveys, which indicate heavy fan Internet use, for example 87% (FA Premier League 2005).

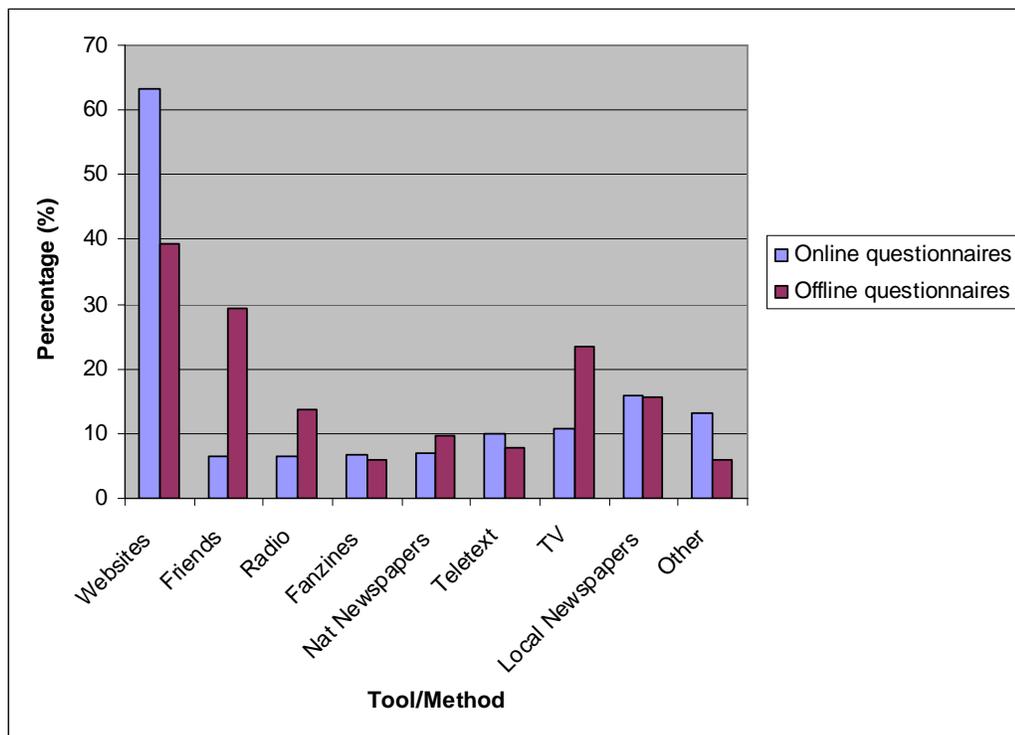


Figure 4.6 Heavy users of tools and methods

Another table which shows yet more evidence of the prominence of the web as an information resource is Table 4.8. This table shows the first tool or method that fans

would use to carry out information searches. Once again, the web is by far the most popular destination (75.9% online, 64.7% offline and 70% interview).

Tool/Method	Online (%)	Offline (%)	Interview (%) ⁶⁵
Websites	75.9	64.7	70.0
Blank ⁶⁶	9.8	5.9	6.7
Friends/Colleagues	8.4	19.6	0.0
Reference Books	4.2	3.9	0.0
Radio	0.5	3.9	0.0
TV	0.3	0.0	0.0
Text Alerts	0.3	0.0	0.0
Fanzines	0.3	0.0	0.0
Teletext	0.3	2.0	3.3
Local Papers	0.0	0.0	10.0
National Papers	0.0	0.0	6.7
Mobile Internet	0.0	0.0	3.3
Total	100.0	100.0	100.0

Table 4.8 The first tool or method respondents would use when looking for information.

The reasons why the web is as much a force as Tables 4.8 suggests are numerous, and some light can be shed on this, as the questionnaires asked for the reasons people use the web. The primary reason for this heavy web use seems to be the same for both online and offline respondents and this was in order to get the latest information (34.7% online, 29.4% offline). According to the online respondents, the second most important reason was to communicate with other fans (28.5%), however, this was not so important for the offline respondents (5.9%). The next notable aspect was the convenience offered by the web, for which the response rates were very similar in both versions of the questionnaire (25.2% online, 23.5% offline). Overall, with the exception of fan communication, the figures from the two questionnaires match reasonably closely (see Figure 4.7). Concerning the fan communication, the difference is only natural when one considers that the online questionnaires were advertised on messageboards (online).

⁶⁵ The question asked at the interview was slightly different, but with very similar implications. The question asked: Out of all the tools you use to look for information, which one do you use primarily?

⁶⁶ Blank entries are from respondents that in an earlier question said they never carry out football information searches.

Reason	Online (%)	Offline (%)
Latest Information	34.7	29.4
Fan Communication	28.5	5.9
Convenience	25.2	23.5
Completeness	6.8	15.7
Authority	1.6	5.9
Other	3.2	19.6
Total	100.0	100.0

Table 4.9 Main reason why fans use the Internet in relation to their football needs

Overall the convenience and benefits of the web are overpowering. This is as a result of the combination of easy Internet access (i.e. many fans have Internet access at work as well as home), the high speed with which the information is available (i.e. as opposed to Teletext or mobile Internet), the completeness of the information as numerous interested parties (news services, the clubs themselves, players, fans, other commercial and non-commercial organisations) feed this information onto the web, and finally the fact that for most practical purposes, it is free (in monetary terms).⁶⁷

⁶⁷ The Internet is generally accessible via a PC and an Internet connection, both of which need to be paid for. However, with the majority of football fans already having access to the Internet for other reasons, it is considered free.

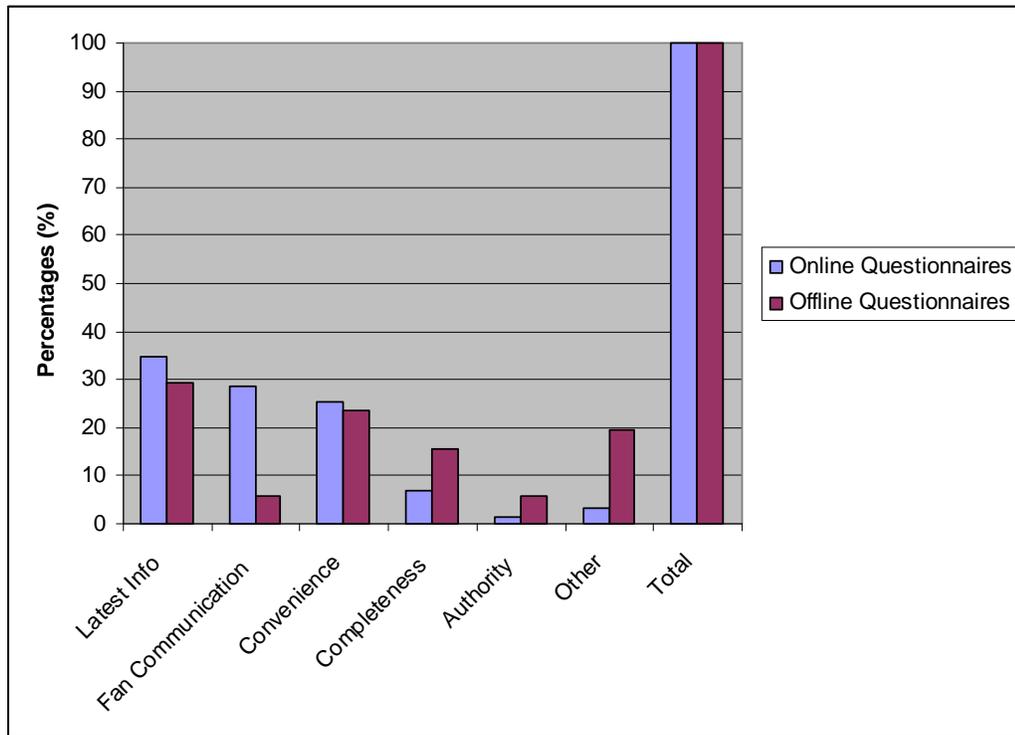


Figure 4.7 Reasons fans use the Internet (for football purposes)

No other information medium can boast this vast array of benefits. Newspapers, magazines, fanzines and reference books need to be paid for separately, as do text alerts, DVDs and mobile Internet. Teletext is free and available around the clock, but the information available is limited, the service is slower and not interactive. TV and radio are free⁶⁸ and available practically 24 hours a day, however, they are not interactive (for the most part), and often do not contain the specific team information as and when it is needed. A friend or colleague might (in theory) be able to provide all services provided by the web as an information resource, but then it is unlikely that this individual will be available for more than a handful of people.

All these benefits seem to be the reason for the amount of time spent by respondents reading about their team on the web, which was a question asked on the questionnaires (but not in the interviews). According to the questionnaires, the vast majority of respondents spend at least 10 minutes a day online (97.8% online, 80.4%

⁶⁸ As mentioned about the Internet, a price has to be paid for a TV or a radio as well. However again, with the majority of the population owning a TV or a radio regardless of their affinity for football, the TV and the radio are considered free.

offline). Meanwhile, a very significant amount of respondents spend over an hour online daily (45.3% online, 23.5% offline). Here again, there is a reasonably close match between the two versions of the questionnaires (Table 4.10).

Time Spent	Online (%)	Offline (%)
Under 10 mins	2.2	19.6
10 mins - 1 hr day	51.7	43.1
1 - 2 hrs day	33.3	19.6
2 to 4 hrs day	9.8	3.9
Over 4 hrs day	2.2	0.0
Don't Know	0.8	13.8
Total	100.0	100.0

Table 4.10 Time spent on the web reading about their team

Additional insight can be found in comments made by the interviewees. Football messageboards and forums were particularly praised by interviewees. An online forum is a site where users can register and communicate with each other concerning any subject they wish to discuss. Forums exist dedicated to almost every professional football club in the top four English leagues. Forums differ from standard sites in terms of information provision, and the layout of the forum can be significantly different (i.e. in most cases a forum website will not look like a news football site at all). On a forum, if information on team statistics is required, instead of clicking on a section (e.g. which might have been called “team stats”), a question is posted asking for the exact information required (e.g. “Who has scored the most goals this season?”), and an answer is provided by another user (normally a few hours later). Forums differ in terms of information provision in that discussion threads are “born” and “killed”, meaning search engines often will not index large parts of any forum.

Regardless, those interviewees who used forums seemed to be very heavy users (including the two fans who said they were reprimanded at work for spending too much time on football websites) who would check for fresh news several times a day, and even more regularly during times of heightened tension or activity (e.g. if a great player was about to be signed). There were several mentions of how often forums get updated.

"Normally, I check it five or six times a day. During the off season, where there's nothing going on, I'll check it all the time because the main sites will have nothing at all, and I need something..." Frei-FUL

"Messageboards are constantly being updated. There is a tremendous amount of information on them." Look-BREN

"Lots of times, I post something, and I want to see the replies, so I'll check it every couple of hours." Tonx-BREN

"I check the GPG⁶⁹ every couple of hours. If something important is going on, I'll check it more. I'll have it up, and I'll keep looking." Tonx-BREN

Numerous other positive aspects of forums were highlighted. One of the most notable was the use of the forum in order to initiate money raising activity. Brighton and Hove Albion used a forum called North Stand Chat to coordinate the carrying out of tasks that would help build a new stadium, including money raising.

"The forum was the focal point of a campaign to get a stadium. We organised all sorts of activities through the forum, who was going to go where, and do what. It was all coordinated using North Stand Chat, and new people who wanted to come in and help could also come in, again through the forum."
Coot-BRI

Other positives about the forums were the ability to get "inside information".

"People on the forums often have links to the club, and will give information you will not get otherwise." Moz-LIV

⁶⁹ GPG is an acronym for Griffin Park Grapevine, the most popular of the Brentford FC forums.

"TFI [The Fulham Independent forum] is better than the BBC, Sky and the official site because of the people who post there. Some players' sisters and wives post there, a guy who works at the clubs posts there. If the team wants to leak something, it'll turn up there." Frei-FUL

Users also commented that any piece of information they needed was available through the forums. One went so far as to say newspapers (i.e. journalists) often acquire information they print by reading the forum sites.

"On the forum, you will always get a reply." Sylv-CHAR

"Some people on the forum have been Fulham fans for 40 years. There's nothing they don't know." Frei-FUL

"If the newspapers have nothing to print. If it's a quiet period, you know in the summer. They scour the messageboards, and print stuff from there." Frei-FUL

Overall, the role played by the web in providing information to fans has to be considered significant. Clearly, fans are using the web heavily, as it has the latest information, is convenient and allows fan interaction (with forums playing a notable role).

4.3.4 The Role of the PHP Within the Web

In this section, Table 4.11 shows the percentage of football fans who use the various types of websites. Of the various types of websites which are included, it is clear that the PHP is one of the less popular (in all except the interviews). Once again, it is possible that the interviewees mentioned their use after some thought, and perhaps even if they use them extremely rarely. It is possible that the figures for the actual use of the PHPs by those who filled out the questionnaires is higher, but that their use is so rare, or their understanding of what constitutes a PHP so poor, that they chose to answer no. Regardless, with the online questionnaires (7.3%) and offline questionnaires both showing very low figures (15.7%), in terms of percentages and

also in order of popularity (8th out of 9 online, 6th out of 9 offline), it is hard to say that the PHP is anything but a marginal tool (as far as football fans are concerned). The evidence from Table 4.12 is only slightly better, with 28.8% of online respondents saying they use PHPs once a month or more, and 33.3% of offline respondents. The low use of PHPs is illustrated even more clearly in Figure 4.8.

Type of Website	Online (%)	Offline (%)	Interviews (%)
Official	85.4	76.5	76.7
Forums/Messageboards	84.6	35.3	53.3
Specific professional	62.9	17.6	36.7
Generic News	54.2	58.8	80
Generic Sports	43.4	37.3	73.3
Webzine	28.2	15.7	10
Fan sites	18.7	5.9	26.7
PHP	7.3	15.7	63.3
Player sites	4.6	2.0	3.3

Table 4.11 Number of people that use each type of website⁷⁰

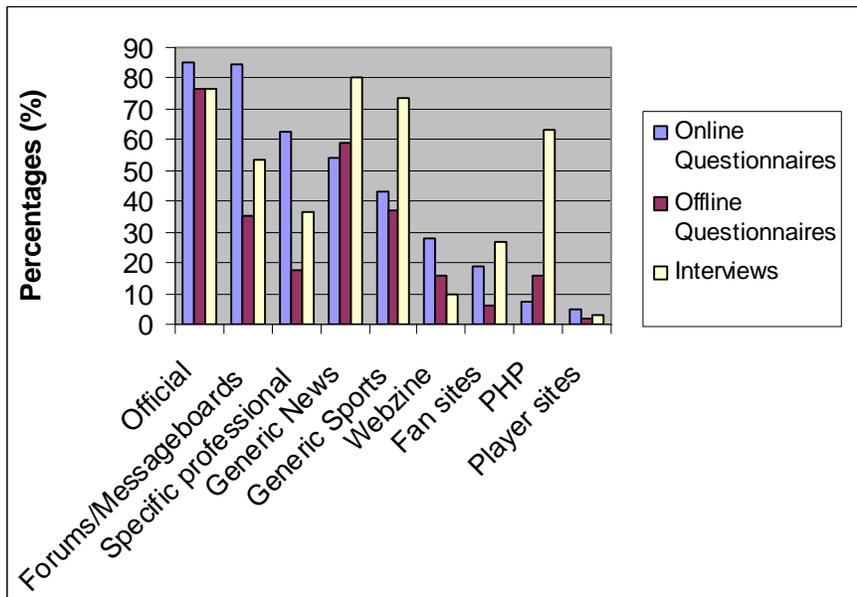


Figure 4.8 Types of sites used by fans

⁷⁰ The totals have not been included in Table 4.11, as they do not add up to 100% on account of respondents being allowed to choose more than one of the options.

Frequency	Online (%)	Offline (%)
Every day	6.8	5.9
A few times a week	8.7	3.9
Once a week	6.8	5.9
Once a month	6.5	17.6
Rarely	35.4	37.3
Never	16.5	2.0
Don't Know	19.3	27.4
Total	100.0	100.0

Table 4.12 Frequency of PHP use

Meanwhile additional evidence supporting the low use of PHPs came from the interviewees, with many claiming not to know enough about them to comment, and several interviewees dismissing their value out of hand.

"PHPs don't really play a role at all." Look-BREN

"PHPs could not possibly match the quality of information found on non-PHPs because a [single] person could never be anywhere near resourceful enough."
Moz-LIV

"PHPs do not play a major role." Dion-LIV

"PHPs have never been important to me." Thom-WAT

One user commented that the big media companies are responsible for the demise of the PHP.

"PHPs were more popular in the beginning, before the big companies picked up every Internet morsel they could find. However, as the companies picked these up, PHPs started dying." Hein-TOT

In order to shed more light on PHP usage, a cross-table was created (using the online questionnaire results), to look at whether the amount of time spent by users on the web had an effect on how often they looked at PHPs. The table shows no clear

relation between the two, something which was confirmed by carrying out a chi-squared test ($p = 0.3411$).⁷¹

	Everyday	Few times /week	Once /week	Once /month	Rarely	Never	Don't Know	Total
Under 10 mins	1	0	0	1	1	5	0	8
10 mins - 1 hr day	13	16	11	15	69	35	32	191
1 - 2 hrs day	8	9	12	6	47	16	25	123
2 to 4 hrs day	3	7	1	2	10	4	9	36
Over 4 hrs day	0	0	1	0	4	1	2	8
Don't Know	0	0	0	0	0	0	3	3
Total	25	32	25	24	131	61	59	369

Table 4.13 Frequency of PHP use against time spent on the web (online)

According to these figures, the PHP is not a popular source of information. The majority of respondents use it once a month or less, and it is also clearly not popular when compared to other types of websites.

4.3.5 Football Fan Satisfaction with the Availability of Football Related Information

Several questions asked in the questionnaires and interviews help shed light on this issue (objective 5). The first table comes from a question where respondents were asked (on a five point scale) how closely websites meet their footballing information needs. On both questionnaires the results show that respondents are to date, quite pleased with the information found on sites. The online questionnaire showed that 89.2% of respondents chose four or higher. The offline questionnaire responses, though not as positive, were also significantly positive, with 72.6% choosing four or higher. Meanwhile, during the interviews, interviewees were asked if they were happy with the overall quality of the sites they used, 80% answered yes. Figure 4.9 has been included as a visual illustration of both sets of results, but also the closeness of the match between the online and offline questionnaires.

⁷¹ Only the online questionnaire results were used. Some of the categories had to be amalgamated before carrying out the test.

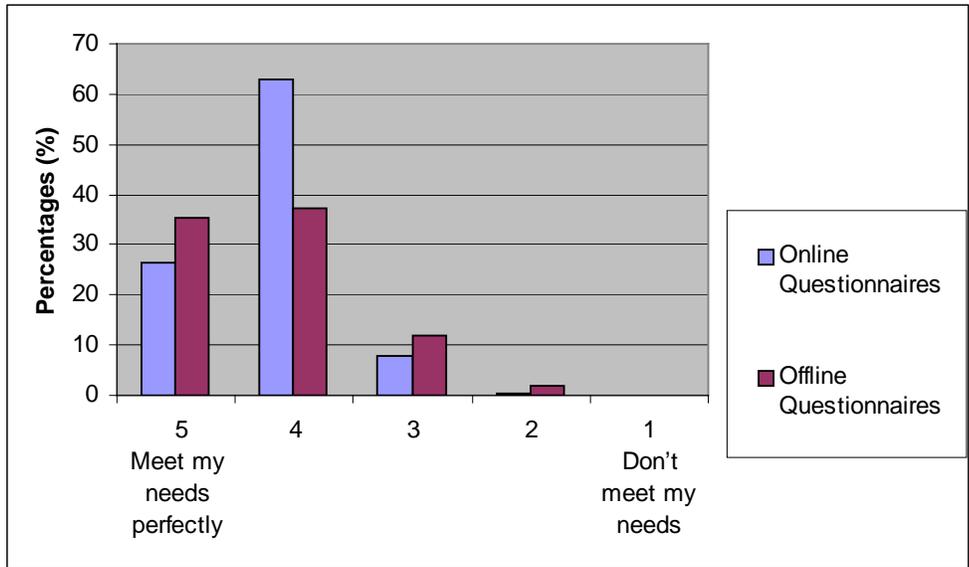


Figure 4.9 Sites meeting the needs of fans

	Interviewees (%)
Yes	80.0
No	13.3
Don't Know	6.7
Total	100.0

Table 4.14 Sites meeting the needs of fans (interviewees)

Table 4.15 was created by putting together the amount of time respondents spent on the web against how closely their information needs were met (using the online questionnaires). Perhaps those that spent more time on the web felt their information needs were met because they went through all the information available. Perhaps, the more time one spends on the web, the hungrier one gets for information. The results though, do not match these theories. According to the Table 4.15, how closely one's information needs are met is completely independent from how much time one spends on the web.

	1	2	3	4	5	Total
Under 10 mins	0	0	0	6	2	8
10 mins - 1 hr day	0	1	22	123	42	188
1 - 2 hrs day	0	1	5	75	40	121
2 to 4 hrs day	0	0	2	21	11	34
Over 4 hrs day	0	0	0	6	2	8
Don't Know	0	0	0	1	0	1
Total	0	2	29	232	97	360

Table 4.15 Sites meeting the needs of fans (5 point scale) against time spent on the web (online).

For further evidence on the availability of footballing information on the web, fans were asked how often they found the specific piece of information they were looking for. The online questionnaire results show that perhaps the availability of information on the web is not as high as indicated by the earlier tables. Looking at the table it is clear that whereas fans are clearly more likely to find what they are looking for than not, only 7% always find what they are looking for. According to the results in Figure 4.9 and Table 4.15, it might have been expected that Table 4.16 would have results in further positive territory. However, only 43.3% could claim that they found what they were looking for always or almost always, showing that there clearly is room for improvement when it comes to the availability of certain information. Those who filled the offline questionnaire were surprisingly more confident at their ability to locate information on the web, with 25.5% claiming they would always find what they were looking for. This confidence is again somewhat surprisingly present in the interviews, where (without the scale provided on the questionnaires) 80% of interviewees said that they would always or almost always find what they were looking for. Meanwhile, though there were some discrepancies, the general trends on both questionnaires matched.

	Online (%)	Offline (%)
Always	7.0	25.5
Almost Always	36.3	35.3
More often than not	36.0	13.7
Sometimes	10.0	11.8
Rarely	0.5	0.0
Never	0.0	0.0
Don't Know	10.2	13.7
Total	100.0	100.0

Table 4.16 How often fans find what they are looking for (online and offline)

	Interview ⁷² (%)
Always/almost always	80
No	13.3
Don't Know	6.7
Total	100.0

Table 4.17 How often fans find what they are looking for (interviewees)

Another table was created to see whether the possibility of finding the information being searched for increased with the time spent on the web. Having combined the results of both questionnaires (Table 4.18), there seems to be little evidence showing that with an increase in time spent, there is an increase in the probability of finding that information. According to the proportions of respondents, those that spend over four hours a day online are most likely to always or almost always find what they are looking for (62.5%). However, as this is not a proportion that increases gradually, this figure should be treated with caution.

⁷² The question asked whether the interviewees always found the information they were looking for.

	Always	Almost Always	More often than not	Sometimes	Rarely	Don't Know	Total*
Under 10 mins	2	4	6	0	0	0	12
10 mins - 1 hr day	17	73	73	26	2	0	191
1 - 2 hrs day	12	56	44	11	0	0	123
2 to 4 hrs day	5	13	13	4	0	1	36
Over 4 hrs day	0	5	3	0	0	0	8
Don't Know	3	1	1	2	0	0	7
Total	39	152	140	43	2	1	377

*The fewer number of respondents are as a result of people who do not generally carry out searches

Table 4.18 How often fans find what they are looking for against time spent on the web (online and offline).

One last table looks at how often respondents carry out searches. Fewer searches might indicate that the information is already readily available. The results do not indicate anything immediately obvious. Both questionnaires and interviews indicate that a notable percentage of fans only carry out searches as frequently as once a month or less. Looking at the figures in order in Table 4.19 (60.4% online, 50.9% offline, 43.3% interviews) hints that those who use the web more, carry out fewer searches.

	Online (%)	Offline (%)	Interview (%)
Once a day	5.4	3.9	6.7
Few times a week	15.7	19.6	13.3
Once a week	16.8	25.5	16.7
Once a month	18.7	23.5	20.0
Rarely	32.5	17.6	23.3
Never	9.2	9.7	0.0
Don't Know	1.7	0.2	20.0
Total	100.0	100.0	100.0

Table 4.19 How often fans look for a specific piece of information

Further insight can be found though, in the comments made by interviewees, particularly on what information fans think is available on the web. In almost all the interviews, either directly or by implication, fans gave the impression that they felt the footballing information on the web was abundant.

"The majority of information you need, 99% of the information is on the web."

Ze-LIV

"Websites are very comprehensive these days. I've never needed anything which I couldn't find on the web." Smar-NUFC

"The web is quite comprehensive." Mile-CHE

"Websites are pretty good these days." Abot-GILL

"[on the web] the truth always comes out eventually." Sylv-CHAR

Certainly, the picture painted by these figures and statements suggest that football fans are pleased with the availability of footballing information on the web. However, fans do not always find what they are looking for, something which indicates that there is still room for improvement.

4.4 Chapter Conclusion

Overall, the findings of the study to this point were encouraging. The objective 1 results (see 4.3.1 Tools and Methods Used for Acquiring Football Related Information, p. 131) show that there is heavy use of websites by football fans, which is done because the information found is up-to-date and the process involved is convenient (objective 2, see 4.3.2 Reasons for the Use of Particular Tools and Methods, p. 140), so much so that websites are by far the most popular tool or method used for the acquisition of footballing information (objective 3, see 4.3.3 The Role Played by the Web, p.142). These are all positive findings, as it means there is a great deal of activity worth investigating. In addition, there is an implication that with the websites being so much more popular than any other tool or method of football information seeking, it will take some time for this state-of-affairs to change, meaning that the results of such a study could potentially be used for some time. The popularity of PHPs is quite low (objective 4, see 4.3.4 The Role of the PHP Within the Web, p. 149), and this is something that suggests that there is room for improvement, even if

in terms of overall football information availability, football fans generally feel that their information needs are being met (objective 5, see 4.3.5 Football Fan Satisfaction with the Availability of Football Related Information, p. 152). All in all, the findings resulting from the research so far encouraged further work to be carried out. The next chapter deals with objectives 6 to 11.

Chapter V

Results II *- Perceptions and Facts about PHPs*

5.1 Introduction

Earlier in this work, a list of aims and objectives was given (see 1.4 Aims and Objectives, p. 17). The previous chapter dealt with a number of objectives concerning the informational behaviour of football fans. This chapter deals with a group of objectives concerned with various properties of PHPs and the way they are perceived by football fans. More specifically, the objectives are:

6. To determine the number of sections that have instances of unique information on football PHPs.
7. To determine the number of sections that have instances of archival storage on football PHPs.
8. To determine the degree of accuracy of the information available on football PHPs.
9. To determine the factors that have an effect on the availability and quality of PHPs.
10. To determine the perceptions that football fans have concerning PHPs
11. To determine the causes behind the perceptions that football fans have of PHPs.

The results are compiled from the investigations of the PHPs (Appendix H), responses to the questionnaires (Appendix F and Appendix G) and interviews (Appendix A and Appendix B).

5.2 The Football PHPs

In the process of meeting the set aims and objectives, a number of PHPs had to be studied. The intention was to examine three PHPs of four teams from each of the four divisions, which would have made a total of 48, however ultimately 42 PHPs were examined for reasons⁷³ discussed in Chapter III (see 3.4 Stage One Web-Study, p. 86).

Once the clubs were chosen and the PHPs were found, the first step was to categorise the PHPs. These can be a standard PHP, a blog/SN site or a photo-site, whilst at the same time classified either as expressive or instrumental (see 3.4.3 PHP Classification, p. 90, for more details). Under various circumstances, PHPs are excluded from the study. Such cases include a site with a certain number of broken links, a site in a foreign language or a site with access restrictions (see 3.4.2 PHP Inclusion, p. 89).

Table 5.1 shows the number of PHPs in the various categories per club. It also splits the number of PHPs into their respective divisions (and subtotals). The table shows that the majority of PHPs (whether expressive or instrumental) are standard PHPs, with the study containing 262 of them (out of 282, or 92.9%). Out of those, 219 are instrumental (83.6%). These figures are reasonable as expressive PHPs are likely to only mention football in passing as part of the author's likes or dislikes. It is almost certain that the vast majority of expressive PHPs contain no reference to football at all. Photo-sites meanwhile, are very few (again whether expressive or instrumental), only 6 out of 282 (2%). There are two likely causes for this small number. Firstly, most PHP authors incorporate photographs as part of their standard PHP⁷⁴ and secondly, most football photographs are included as part of the photographer's personal collection, mixed with pictures of friends and family, making it difficult for other fans (or search engines) to find. The most peculiar result in the table is the low

⁷³ For three clubs, three PHPs were not available. The three clubs which were represented by just one PHP were Brentford (League 1), Rochdale and Chester City (both League 2), all of which are in the two lowest professional divisions.

⁷⁴ For a site to be classified as a photo-site, it must contain nothing but photographs.

number of blogs/SN sites that were come across, 14 in total (5%). Here, the most likely explanation is that most blog/SN site authors do not comment exclusively on football matters, thereby reducing the frequency with which they are linked to by other sites. As explained in more detail earlier (see 3.4.1 Finding PHPs, p. 88), the searches for all PHPs were carried out by exhausting all links on the sites initially discovered, as well as carrying out searches using Google (and exhausting all links found on those sites too) making this outcome all the more surprising.

Team	Instrumental (e.g. Manchester United unofficial)			Expressive (e.g. Jack's page)			Total
	Photo-site	Blog/SNS	s-PHP	Photo-site	Blog/SNS	s-PHP	
Manchester United	1	1	88	1	0	17	108
Newcastle United	0	0	34	1	1	8	44
Middlesbrough	0	0	9	1	0	2	12
Charlton Athletic	1	4	4	0	4	1	14
Premiership sub total	2	5	135	3	5	28	178
Derby County	0	0	8	0	0	3	11
Wolverhampton Wanderers	0	0	22	0	0	3	25
Leeds United	0	1	20	0	1	3	25
Brighton Hove Albion	0	0	7	0	1	3	11
Championship subtotals	0	1	57	0	2	12	72
Hull City	0	0	4	0	0	0	4
Walsall	0	0	8	0	0	0	8
Brentford	0	0	1	0	0	0	1
Huddersfield Town	0	0	5	0	0	0	5
League 1 subtotals	0	0	18	0	0	0	18
Southend United	0	0	2	1	1	3	7
Cheltenham Town	0	0	5	0	0	0	5
Rochdale	0	0	1	0	0	0	1
Chester City	0	0	1	0	0	0	1
League 2 subtotals	0	0	9	1	1	3	14
Totals	2	6	219	4	8	43	282

Table 5.1 PHP classification

5.3 Unique Information on Football PHPs

This first objective for this chapter (objective 6) attempted to determine whether PHPs as a whole made any "original contribution" to the web as an information resource. To

that end, every section in all the PHPs was looked at to try and determine how many of them had instances of unique information.

Having examined the 42 PHPs, ratings were given to each section of each PHP, showing a number of properties. As described earlier in this work, in the first step, the sections in the PHPs were compared to the sections in the three non-PHPs and ratings were assigned (see 3.4.4 Comparing PHPs, p. 96, and 3.4.5 Relative Information Content Ratings, p. 97).

Table 5.2 shows the PHPs, and alongside them the number of Web-Wide Unique (XX),⁷⁵ Archival (A), Locally Exclusive (X), Equinferior (\leq) and Not Applicable (N/A) rated sections they contain. It has been included to give an overall picture of the PHP investigations.

Looking at the table as a whole, some noteworthy observations can be made. Out of 368 sections in total, 150 (40.8%) contained Locally Exclusive information, with 70 (19%) containing Web-Wide Unique information, 112 (30.4%) containing Equinferior information and 106 (28.8%) sections where the comparisons were Not Applicable. Looking at an overall picture where the unrelated sections of the sites are disregarded (i.e. without taking N/A ratings into account), there are more sections with Locally Exclusive ratings than there are with Equinferior ratings (40.8% against 30.4%), meaning that there are more sections on PHPs that have Locally Exclusive information than there are that do not. Once again, it must be stated that the size of the sample is not large enough for the figures to be conclusive, but the number of Locally Exclusive and Web-Wide Unique ratings are certainly surprisingly high.

⁷⁵ When looking at the X and XX ratings, it is easy to conclude that the X rating should not be given much importance, as the information in question is available elsewhere on the web, probably on a non-PHP. Though this might be true, there are two important points to bear in mind. The first is that the information that is provided on the PHP is more often than not relevant to a specific team. The same information which is available elsewhere, is more than likely available on a generic football/information site, a place many fans might never come across. Secondly, on many occasions, the information found elsewhere on the web, was only available on other PHPs.

Whilst examining the first five objectives (objective 4 in particular, see 4.3.4 The Role of the PHP Within the Web, p. 149), it was clear that PHPs were not very popular with football fans. With that data, it would not be unreasonable to assume that there is little information of value on PHPs. However, according to Table 5.2, this is not the case.

Examining the figures division by division in terms of percentages (see Table 5.3 and Figure 5.1), what is immediately notable is the significant increase in the number of Locally Exclusive and Web-Wide Unique ratings in the League 2 subsection. Here a very substantial 32.4% of the sections have a Web-Wide Unique rating, with a mammoth 51.4% having Locally Exclusive ratings. However, looking more closely at the individual PHPs in Table 5.2, it can be seen that these unusual figures are all the result of one site, the Southend United Database. Without this PHP, the percentages would be in line with the other divisions.

Name	XX	A	X	≤	N/A	Total
Newcastle Football Club	1	0	2	2	0	4
Toon Down South	3	0	6	2	1	9
ToonInfo	1	0	5	2	6	13
Geraldine's Viva Manchester United	1	8	13	9	7	29
United World	2	2	6	4	0	10
Manunited FC Homepage	0	1	2	0	6	8
Steve Bridge	1	0	1	0	0	1
The King of Sweden	0	0	0	2	0	2
Ball:Sports	0	0	1	0	0	1
Ayresome Park	4	1	4	2	6	12
2004 Carling Cup Final Shrine	3	0	0	3	0	3
Paul Thompson's website	0	0	4	3	3	10
Premiership subtotals	16	9	44	29	29	102
Derby County (Steve Eyre)	4	0	5	1	6	12
www.the-rams.derbycounty.co.uk	0	0	0	4	6	10
Eyal Glozer's Derby County Website	2	2	4	7	4	15
jfEidem's Wolverhampton Wanderers page	1	3	5	2	1	8
Black Country Derby	2	0	4	0	1	5
WolvesWeb	1	0	1	4	3	8
Little Yellow Daffodil	2	0	6	0	1	7
We'll Be Back	0	1	3	8	1	12
Super Leeds Utd	2	0	7	2	3	12
Albion Album	3	0	3	5	6	14
Goodbye Goldstone	0	0	0	3	4	7
BHAFC24	0	3	3	0	0	3
Championship subtotals	17	9	41	36	36	113
Tigerland	0	0	0	1	2	3
Mclean	2	0	3	1	4	8
Hull City Pics	3	3	3	2	6	11
Par Sterner's Unofficial Walsall Homepage	1	2	2	5	2	9
Walsall F.C Online	0	0	4	9	3	16
Steve's Walsall Football Club Pages	3	0	7	8	3	18
Winston Bee	0	0	1	0	0	1
Those were the days my friends	0	1	3	0	4	7
HTFC-world	3	0	3	0	2	5
HTAFC Links	1	0	1	0	0	1
League 1 subtotals	13	6	27	26	26	79
Lewis Hunt	1	0	1	4	4	9
Southend United Database	14	5	19	4	4	27
One Team in Essex	1	1	2	0	0	2
Cheltenham Town FC Rock	0	0	0	3	2	5
The Robin's Nest	2	1	5	4	0	9
Brayson For England	1	0	1	0	0	1
True Blue	2	2	6	4	2	12
Rochdale FC Picture Gallery	3	1	4	2	3	9
League 2 subtotals	24	10	38	21	15	74
Totals	70	34	150	112	106	368

Table 5.2 Ratings per PHP - These are ratings (RICRs) acquired from PHP section comparisons (i.e. the numbers represent PHP sections).⁷⁶

⁷⁶ XX and A values are not included in the far right column totals (see 3.4.4 Comparing PHPs).

Of course, these figures alone cannot mean that PHPs contain more information than, or are more popular than non-PHPs. Even with these figures, it must be noted that a section on a non-PHP might have several times the amount of unique information that is found on the PHP. Sections on the non-PHPs examined are also in general much larger and less likely to contain dead links or out of date information. Furthermore, sections of PHPs that were given N/A ratings could have acquired the rating on account of dead or external links as well as for containing football-irrelevant information.

Name	XX ⁷⁷	X	≤	N/A	Total Sections
Premiership	15.7%	43.1%	28.4%	28.4%	100%
Championship	15.0%	36.3%	31.9%	31.9%	100%
League 1	16.5%	34.2%	32.9%	32.9%	100%
League 2	32.4%	51.4%	28.4%	20.3%	100%

Table 5.3 Ratings per division

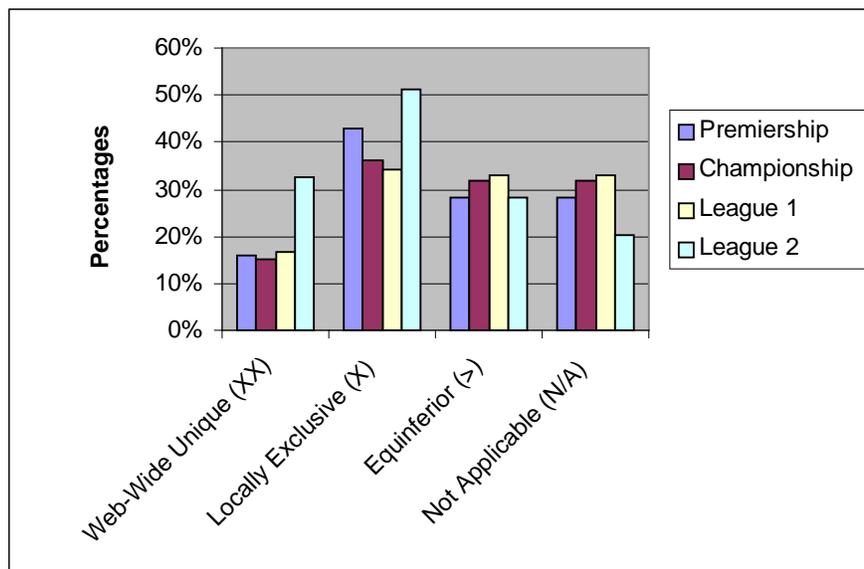


Figure 5.1 Types of information according to division

Regardless though, these figures as a whole show clearly enough, that there are definitely reasons to look for ways to harness the PHPs as an information resource.

⁷⁷ The Web-Wide Unique figures are not included in the totals column.

5.4 Archival Storage on Football PHPs

This seventh objective essentially looked at whether there was a possibility of deriving any "older" information from PHPs. It was an attempt to determine whether this "old" information which is available on PHPs is not available elsewhere on the web.

Hence, in addition to assigning Locally Exclusive, Equinferior and Not Applicable ratings, the first stage of the comparisons also looked for instances where an Archival (A) rating could be given (see Table 5.2). Locally Exclusive ratings can be assigned if historical or old information is available on the PHP (see 3.4.5 Relative Information Content Ratings, p. 97). It should be noted that sections are assigned an Archival rating only when they already have a Locally Exclusive rating (so actually, the same section is given two ratings).⁷⁸

In the PHP investigations, 34 out of 368 (i.e. 9.2%) sections were given an Archival rating, meaning that a good deal of historical facts and figures are available on PHPs. When looked at as a percentage of the total number of Locally Exclusive ratings given, the figure is 22.7%; a significant portion. Figure 5.2 shows a clear illustration of these percentages according to division.

⁷⁸ Meaning that the number of X ratings does not go down as A ratings increase. A section can have both an X rating and an A rating.

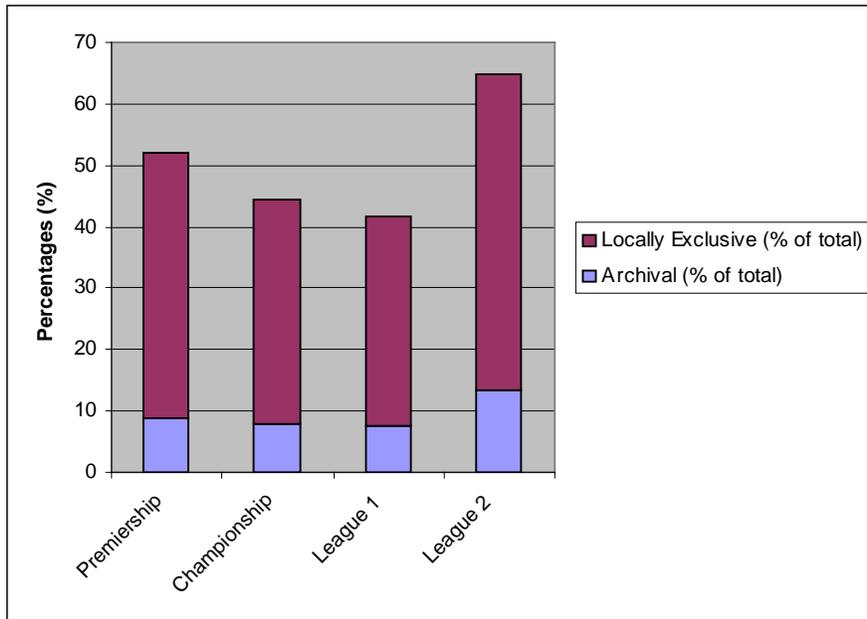


Figure 5.2 Archival (A)/Locally Exclusive (X) Ratings as percentage of total

Looking at the divisional subtotals in more detail (Table 5.4), again, League 2 is the odd one out, though this time the differences are more subtle. The 51.4% of Locally Exclusive ratings is a figure already mentioned earlier. The higher figures of Archival ratings, both as a percentage of the Locally Exclusive ratings (26.3%) and total sections (13.5%) are not too different. Of the two, the 13.5% figure seems surprisingly high, but once again, this is a result of the Southend United Database.

Name	Archival (% of X)	Archival (% of total)	Locally Exclusive (% of total)
Premiership subtotals	20.5	8.8	43.1
Championship subtotals	22.0	8.0	36.3
League 1 subtotals	22.2	7.6	34.2
League 2 subtotals	26.3	13.5	51.4
Totals	22.7	9.2	40.8

Table 5.4 Instances of archival storage according to division

All in all, according to this sample, 9.2% (approximately one tenth) of the sections of PHPs contain older information which is not available on the non-PHPs. The

percentages are not insignificant as they indicate that the existence of PHPs could potentially provide an insightful snapshot of football history on the web.

5.5 The Accuracy of the Information on Football PHPs

The purpose of this eighth objective is to determine broadly to what degree the information on PHPs is factually accurate. As mentioned earlier in the study (see 3.4.8 Accuracy of Information, p. 109), 7 pieces of information are looked at in any one section and depending on how many of them are inaccurate, a rating of accuracy is given. To get a High rating, there must be no more than one error out of the 7 instances of information. To get a Low rating, there must be four or more (i.e. more than half the information must be inaccurate).

	Number of Sections	Percentage
High	26	70.3%
Medium	0	0.0%
Low	11	29.7%
Total	37	100%

Table 5.5 Degree of accuracy of PHPs

All in all, just over 10%⁷⁹ of all the sections of the PHPs were examined (37 out of 368). Out of these, 70.3% had a High rating, and 29.7% had a Low rating. There was not a single instance where a section had a Medium rating. A striking illustration can be seen in Figure 5.3. These would come across as very peculiar figures if it were not for the fact that almost every instance of inaccurate information discovered was accurate at one point in time.⁸⁰ All of the instances of inaccurate information were discovered on pages of sites that were clearly out of date. The implication here is that the inaccuracy of this information is not due to web authors making errors (either on purpose or by accident), but rather that the information is generally accurate, but ceases to be so as the sites become outdated.

⁷⁹ The exact figure is 10.05% (37 sections out of a total of 368 sections).

⁸⁰ Though the out-of-date pieces of information were not always verified to be accurate, not a single piece of contemporary information was found to be inaccurate.

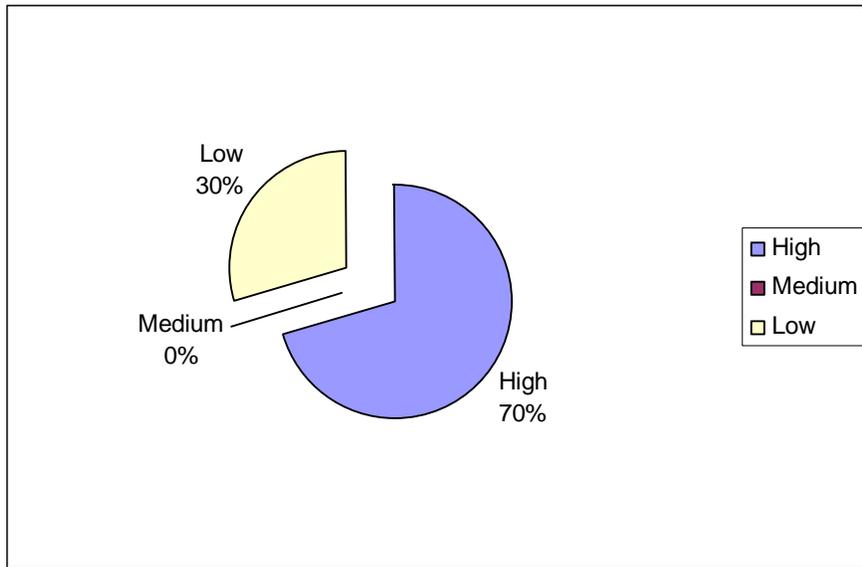


Figure 5.3 Degree of accuracy

On the other hand, information which is out of date cannot be considered accurate. If a user did not realise that the information was out of date, they would certainly be misled. Furthermore, it is extremely unlikely that (even if out-of-date information is treated as accurate) PHP authors never make a mistake, either in gathering or entering information on their web pages. This thesis looks at a small sample where any figures are indicative rather than absolute.

However, it appears that the evidence strongly supports the idea that the information gathered and put up by PHP authors on their websites is, in the vast majority of the cases accurate, at least in the field of football, and this was not an expected outcome.

5.6 Factors that Affect the Availability and Quality of PHPs

The purpose of this objective was to investigate from several angles the various factors that might have an effect on the quantity and quality of PHPs.

5.6.1 Quantity of PHPs

Looking at the number of PHPs found by division (Figure 5.4)⁸¹ it is clear to see that there are great differences between the Premiership and Championship (178 and 72 PHPs respectively), and then between the Championship and the remaining two leagues (72, 18 and 14 PHPs respectively). More specifically, looking at the teams in the Premiership (Figure 5.5), 178 PHPs were discovered for four clubs, whereas the same number of teams in the Championship produced less than half that number of PHPs, at only 72. The difference to the lower leagues is even greater. League 1 teams were represented by just 18 PHPs, with Brentford being represented by a single blog. League 2 had an even weaker representation, only 14 PHPs, with two teams being represented by single PHPs (Chester City and Rochdale).

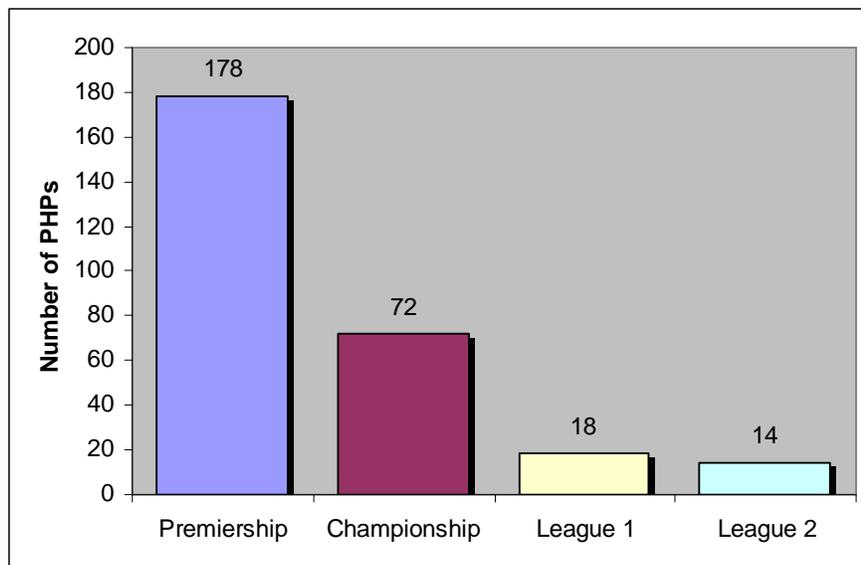


Figure 5.4 Number of PHPs per division

Of course, the results represent a sample of the divisions (just the four teams randomly chosen), and not the entire divisions. However, the sizes of the differences between the samples are still notable and can be seen even more clearly in Figure 5.4 above.

⁸¹ The subtotals for every division are just the number of PHPs found for the four teams in that division added up. It is not the number of PHPs of all the teams in that division.

Looking further into Figure 5.5, at a team by team basis, there is evidence that helps explain the figures. At the upper end of the football leagues, it is immediately obvious that the number of PHPs in existence seems to be strongly related to the performance of the team on the pitch. This can also be seen both by looking at the PHPs per team (Figure 5.5) or by division (Figure 5.4). It must be noted at this point that since the completion of the PHP investigations (which used the 2004/05 league tables), many of the teams have changed division.

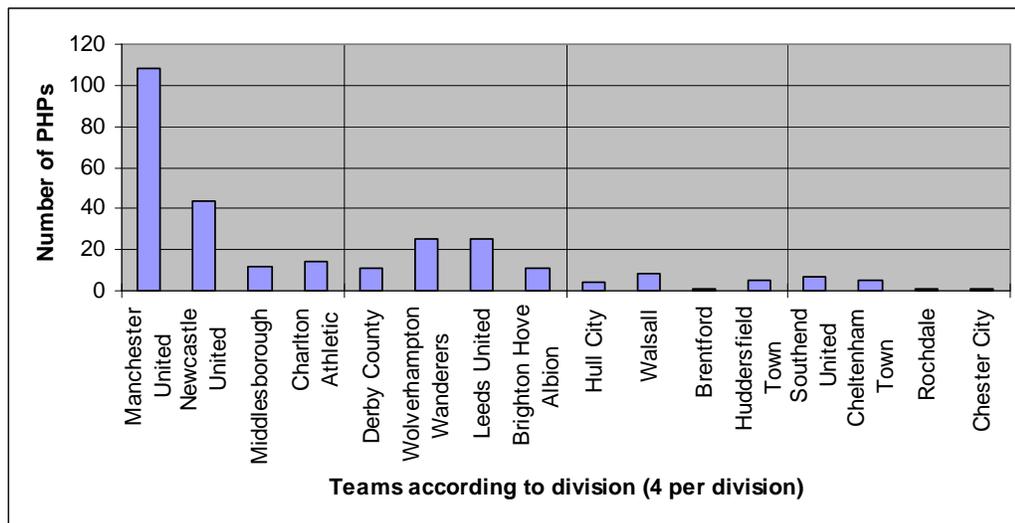


Figure 5.5 Number of PHPs per team

Looking at the top of the table, it is immediately clear that Manchester United and Newcastle United both enjoy a sort of PHP superiority, at least when it comes to the number of PHPs that their fans create. Manchester United in particular has more PHPs (108) created to represent it than all 12 teams of the lower three divisions put together (104). Of course, Manchester United is one of the greatest teams in the world, as already mentioned. It was the football club with the highest turnover in the world for eight years in a row, before being overtaken by Real Madrid in 2005/06 (Deloitte and Touche 2006). They are the team with the most Premiership titles since its inception in the 1992/93 season (Manchester United Official Site 2006a), and are one of two English teams to have won the European Cup in the last 10 years with their

victory in May 1999.⁸² They also have the stadium with the highest capacity with Old Trafford holding 76,212 all seated fans (Manchester United Official Site 2006b), and the highest average attendance in the Premiership of 75,826 (Soccerstats.com Site 2007). Clearly, Manchester United have history enough to support their representation in PHPs, as the most successful English football club of recent years.

Newcastle United, whilst not having enjoyed this sort of success on the pitch, also have comparative figures to support them. Having had until recently the stadium with the second highest capacity⁸³ in the Premiership at 52,387 (Newcastle United Official Site 2006), they had the second highest average attendance too, at 51,800 (Deloitte and Touche 2006) and can claim to have near-enough the most fanatical support, having had average crowds of over 50,000 since St James Park expanded (Deloitte 2007), whilst having won little in that period.⁸⁴ In their defence, the club have qualified for European football every year consecutively from the 2001/02 season to the 2007/08 season with one exception (the 2005/06 season). Finally, Newcastle United had an era of strong performance in the mid-nineties with two consecutive second place finishes.

Of the remaining two Premiership sides, none have had this kind of success. Middlesbrough have had some success, winning the League Cup in the 2003/04 season, and subsequently qualifying for the UEFA cup twice in a row. However, in terms of the league, Middlesbrough have managed no better than a single 7th placed finish in the previous 10 seasons. In terms of stadium capacity and attendances, again their performance is not noteworthy, with a stadium capacity of 35,100 and average attendances of 32,012 (10th out of 20 in 2004/05, Soccer-stats.com Site 2006).

Charlton Athletic are surely the "smallest" team in this group of four, having spent the majority of the past 50 years outside the top division. Though consolidating their Premiership status with five mid-table finishes, including a highest placing since 1953

⁸² Manchester United again won the Champions League after the completion of this research with their victory in May 2008.

⁸³ In the 2006/07 season, Arsenal moved to their new Emirates Stadium which now has the second highest capacity in the Premiership (Arsenal Official Site 2006).

⁸⁴ Newcastle United's last professional trophy of note was the Fair's Cup (today's UEFA cup) in 1969. Since then, they have won the Intertoto Cup (in the 2006/07 season), considered unimportant by many fans.

(7th), they have not managed to win any trophies of note, with no forays into the European competitions either. Once again, in terms of stadium capacity and average attendances, there are no figures of note, with capacity at 27,111 (Internet Football Ground Guide Site 2006), and average attendances at 26,403, placing them 13th out of 20 in 2004/05 (Soccer-stats.com Site 2006).

Moving further down the divisions, it is immediately noticeable that two of the clubs in the Championship are represented by more PHPs than both Middlesbrough and Charlton. Wolverhampton Wanderers and Leeds United are both represented by 25 PHPs each, compared to 14 and 12 respectively for Middlesbrough and Charlton. Brighton Hove Albion and Derby County meanwhile, are both represented by 11 PHPs, not significantly less than the two "smaller" Premiership sides.

These figures can be explained without much difficulty for Leeds United. In the late 1990s, they managed to get to both the Champions League and UEFA cup semi-finals, whilst at the same time managing a string of top five Premiership finishes. As devoted football fans are not known to abandon their sides, it makes sense that those who created the sites in the late nineties kept their interests alive. The 11 PHPs for Derby County are also not too surprising, as the club quite recently spent six seasons in the Premiership (1996/97 to 2001/02). Perhaps the relatively low figure is due to the fact that though the team have a significant fan base, the fans feel a little demoralised by the fact that they spent several consecutive years outside the top division, affecting the creation of PHPs. A similar idea can be used to explain Brighton Hove Albion's 11 PHPs. Though Brighton had a brief adventure in the old Division 1 (from 1979/80 to 1983/84), they are currently at the Withdean Stadium in the Championship, having the lowest capacity in the division. However, it is possible that the fans finally consider themselves as a club "on the up" (with the club getting promoted in recent years and a new stadium finally in sight) and have given the club a stronger online presence.

The intriguing figures are for Wolverhampton Wanderers (Wolves), who have only had one season of note in the previous 10 years (in 2003/04 when they were promoted to the Premiership), yet have 25 PHPs. However, that single successful year was immediately followed by failure, when the team were relegated back to the

Championship after only one season in the top flight. In fact, Wolves have been outside the top division in England since 1984 with the exception of the aforementioned single season. In that time, they dropped as far down as the 4th professional division, and did not win any trophies of note. Wolves are historically significant, being founding members of the football league in the 19th century, and having won several league titles, FA cups and league cups (Wolverhampton Wanderers Official Site 2005). However, their most recent major accomplishment was a foray into the UEFA cup in 1980/81 season, some time before the birth of the World Wide Web. Furthermore, Wolves have a number of local rivals who have been more successful in recent years, most notably Aston Villa, but even Birmingham City and West Bromwich Albion (both of which have spent more time in the Premiership in recent years).

Team	Stadium Name	PHPs	Average Attendance	Stadium Capacity	Total Annual Attendance
Man Utd	Old Trafford	108	67,871	67,700	1,289,541
Newcastle	St James' Park	44	51,844	52,150	985,040
Leeds	Elland Road	25	29,207	40,204	671,769
Wolves	Molineux	25	26,620	29,400	612,259
Middlesbrough	Riverside Stadium	14	32,012	35,100	608,236
Charlton	The Valley	12	26,403	26,400	501,656
Derby	Pride Park	11	25,219	33,597	580,039
Brighton	Withdean Stadium	11	6,426	6,973	147,809
Hull City	Kingston Comm Stadium	8	18,027	25,000	414,629
Walsall	Bescot Stadium	8	6,081	11,300	139,872
Southend	Roots Hall	7	6,077	12,343	139,777
Huddersfield	McAlpine Stadium	5	11,909	24,500	273,914
Cheltenham	Whaddon Road	5	3,648	7,400	83,911
Brentford	Griffin Park	1	6,082	12,763	139,877
Chester City	Deva Stadium	1	2,812	6,012	64,680
Rochdale	Spotland	1	2,690	10,218	61,878

Table 5.6 Number of PHPs, average attendance, stadium capacity and total annual attendance per team (2005/06 season, Soccer-stats.com Site 2006)

Going into the lower two divisions, it becomes difficult to come to any conclusions with the very small samples. Table 5.6 shows each club, with the number of PHPs it is represented by, then the average attendance, stadium capacity and total attendance

over the year. Rochdale and Chester City are relatively straightforward to explain by looking at the table. They have the lowest average attendances of all the teams and have not been above the lowest professional division since 1974/75 (Rochdale) or 1994/95 (Chester City).

One peculiarity is Brentford FC. They have been in League 1 for most of their recent history and had an average attendance of over 6,000 (approximately as many as Brighton, Walsall and Southend) in 2004/05, and yet have only one PHP to represent them. A reason for this could be the financial instability of the club, having changed hands three times in the past ten years (Brentford Official Site 2006, BBC Website 2006b), though perhaps it is more likely that the close proximity to a number of traditionally more popular teams (e.g. Chelsea, Fulham and Queens Park Rangers) has played a more significant role in discouraging fans from creating PHPs about this club.

The other peculiarity is Huddersfield Town. They have been averaging an attendance of almost 12,000 (almost double the figures for Southend, Walsall and Brighton, all of whom have more PHPs) and yet only manage five PHPs to represent them. Furthermore, they used to be further up the leagues (in the Championship), as recently as 2001. Again, a possible explanation for this might be the financial troubles of 2003, as Huddersfield went into administration for four months (Huddersfield Town Supporters Trust 2005).

All in all, it seems that success on the pitch does indeed increase the number of PHPs created about a club. Perhaps the connection is indirect, and the increase in the number of fans a football club has increases the number of PHPs created as a result. It is unclear as to whether the percentage of fans that create PHPs about their football team increases or decreases in relation to the team's success on the pitch (this is something that this study has not attempted to verify for feasibility reasons). Whatever the mechanism, there appears to be a link between the success of a team on the pitch, and the number of PHPs created about that team. To illustrate this point, Figure 5.6 was created. Using Manchester United as a benchmark, the graph shows the proportion of PHPs to various football club attributes such as stadium capacity, average attendance and total annual attendance.

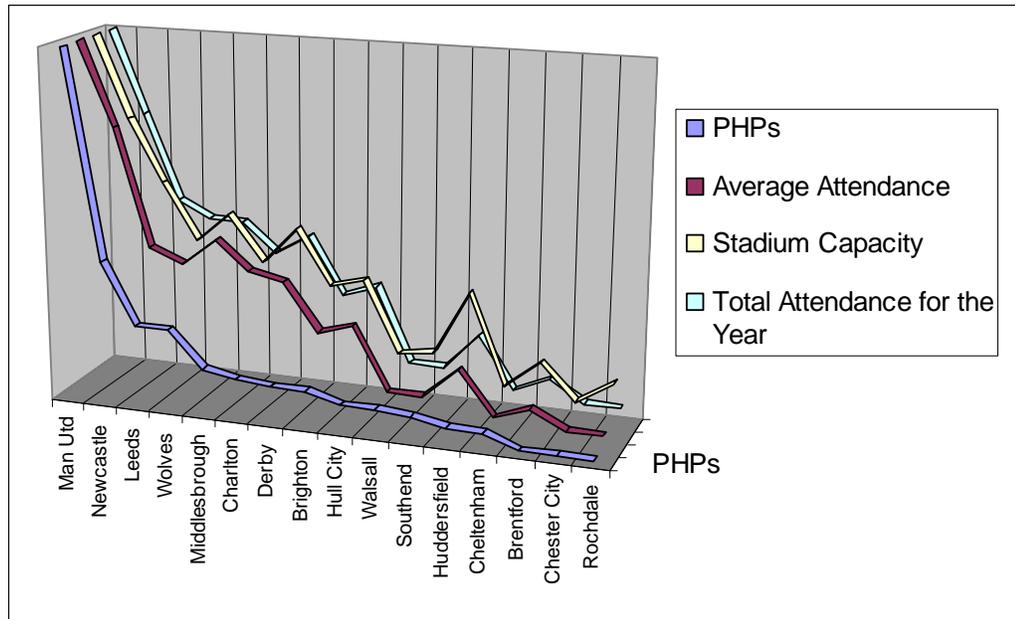


Figure 5.6 PHP proportions (to stadium capacity, average attendance and total annual attendance)

However, this idea of the number of PHPs being linked to the performance of the team on the pitch, goes against one of the preliminary ideas of this work, which was that PHPs might cover gaps in the informational environments left by the principal information providers. Whereas this was an area where the PHPs have larger gaps to fill (with the less coverage that the lower divisions get in general), this table suggests that the fans are doing little to help. Of course, Table 5.6 alone is not conclusive, both because of the size of the sample, but also because the quality and comprehensiveness of PHPs that do exist could still play this role (something the table does not represent). Regardless, the table does not support the idea that fans (and by extension PHP authors) are coming in to fill the gap left by the larger information providers.

5.6.2 Quality of PHPs

In terms of the quality of PHPs, a number of sites are notable for various reasons, either in terms of originality, amount of work required or overall quality.

Perhaps the most unusual PHP encountered was the HTFC-world site (2005). This PHP contains humorous match reports (in cartoon form) of the vast majority of matches that Huddersfield Town have played since 1999. Each match report reads like a comic book with the characters of players, fans and so on making humorous statements throughout the key incidents in a match. In addition the site contains some Huddersfield wallpapers and screensavers, and some information about the habits of the author. The sheer number and size of the match reports alone suggest that a great deal of effort was put into creating the site and the impressive result is clear.

Another noteworthy site is the Albion Album (2005) which contains hundreds of photographs, the vast majority of which were taken by the author himself. The author has been taking club related photographs regularly since at least 1998. Meanwhile, the site has other features too, such as additional information about matches (e.g. date, score and attendance), links to external sites, and even access to information about mobile phone "ringtone" downloads.

Moving higher up the divisions, Geraldine's Viva Manchester United site (2005) is a notable PHP for this study, as it is the one that has the most sections (see Table 5.2), and this reflects the generally high standard of the site. However, it must be noted that this PHP is definitely not the most remarkable come across (see Figure 5.7). The site has many sections, with plenty of information including details about the players, the stadium, the club's history and the club's contact details. The site also has a good deal of statistics, including the league and cup history (detailing final league and cup positions), player statistics and club records (e.g. victory with the highest margin). It also has links to other sites, and a forum for discussion (though this is an external link). However, there are also a few very noticeable problems. Firstly and most notably, the site is no longer being updated, and all of the statistics end at the 2002/03 season. There are also a fair number of dead links (most probably as a result of the

changes on other sites which have not been reflected on this PHP). Overall, it is quite likely that fans have stopped regular visits to the site. In terms of unique information however, the PHP has plenty of facts and figures not available on the official, rivals and BBC sites.

Black Country Derby (2005) is an unusual site about the rivalry between Wolverhampton Wanderers and West Bromwich Albion. The site has statistics about every encounter between the two teams, including the date, score, venue, attendance, scorers (with times) and the tournament (e.g. FA cup). A statistics site is not that unusual and even though the site has statistics going back to 1886, this is not the aspect of the site which makes it unusual. This site was notable for the fact that it was about a rivalry between two clubs rather than the usual site dedicated to a single club.

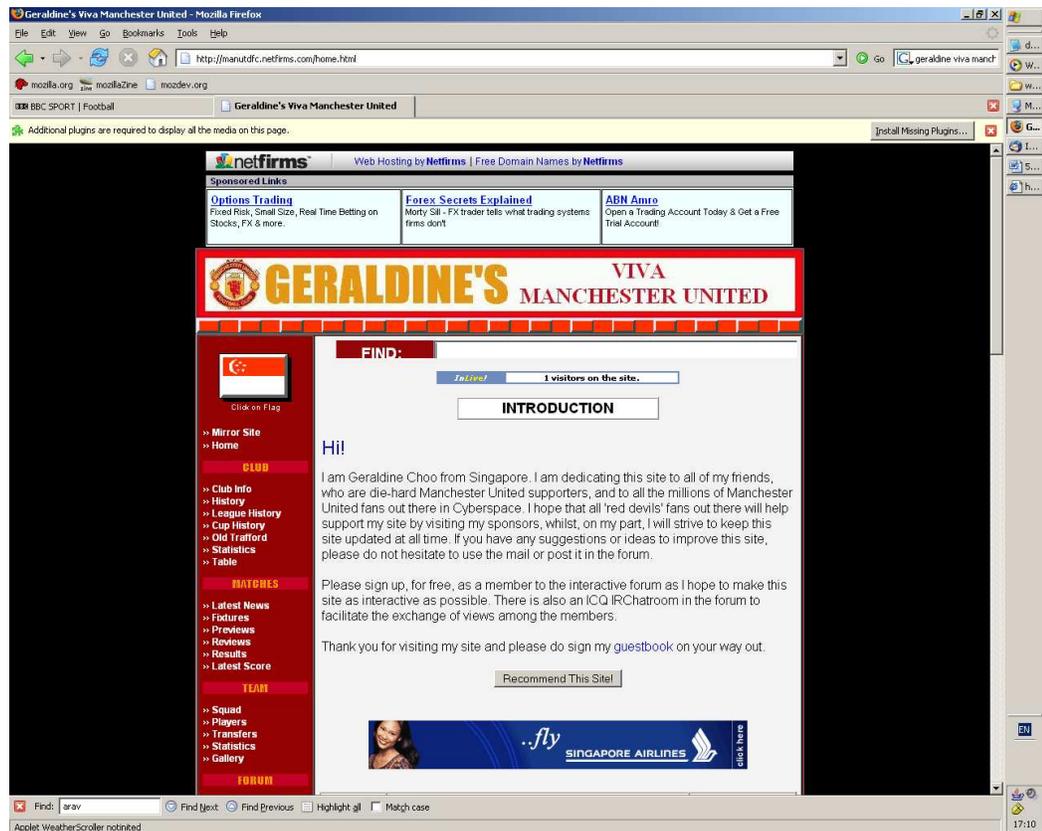


Figure 5.7 Geraldine's Viva Manchester United

Meanwhile, in addition to Geraldine's Manchester United site (2005) and The Black Country Derby site (2005), both of which contain a fair amount of statistical information, there are a number of PHPs containing a great deal more statistical information, which must have required a significant effort from their creators. These include the Robin's Nest (2005), Par Sterner's Unofficial Walsall FC Homepage (2005) and the Southend United Database (2005).

The Robin's Nest (2005) contains the results, goal scorers and final tables for every match Cheltenham Town have played since 1950. In addition, it also contains a history of the club including descriptions of historically significant matches, additional club information, relevant links and some digital Cheltenham artefacts, such as wallpapers and "winamp skins".⁸⁵

Par Sterner's Unofficial Walsall FC Homepage (2005) is again impressive because of the statistics it contains, but also because of its all round design and quality of implementation. It is visually impressive and functionally efficient, with various statistics dating back to 1988. As a football site, it covers most aspects of the team, including club information, details about the squad, the history and then several years' worth of statistics on fixtures, transfers and individual players.

Both of the "statistically heavy" sites above contain a great deal of facts and figures about their clubs. However, the complete statistical guide to a club must be the Southend United Database site (2005). This is by far the most complete of all the PHPs come across in this study (see Figure 5.8). The sheer volume of statistics is unparalleled, and the organisation and technical sophistication of the site is very rare for a PHP. The date, opponents, attendance, goals scored and scorers for every match ever played are available, even the league position and goal difference after the conclusion of the match are included. The technical sophistication of the site is such that this information can be re-ordered and filtered so as to provide even more

⁸⁵ Winamp is a media player for the Microsoft Windows operating system. A "skin" is a decorative design that can be added to the player to change its appearance.

information such as top scorers, results against various teams, performance graphs, attendances and even matches that took place and player birthdays on this day.⁸⁶

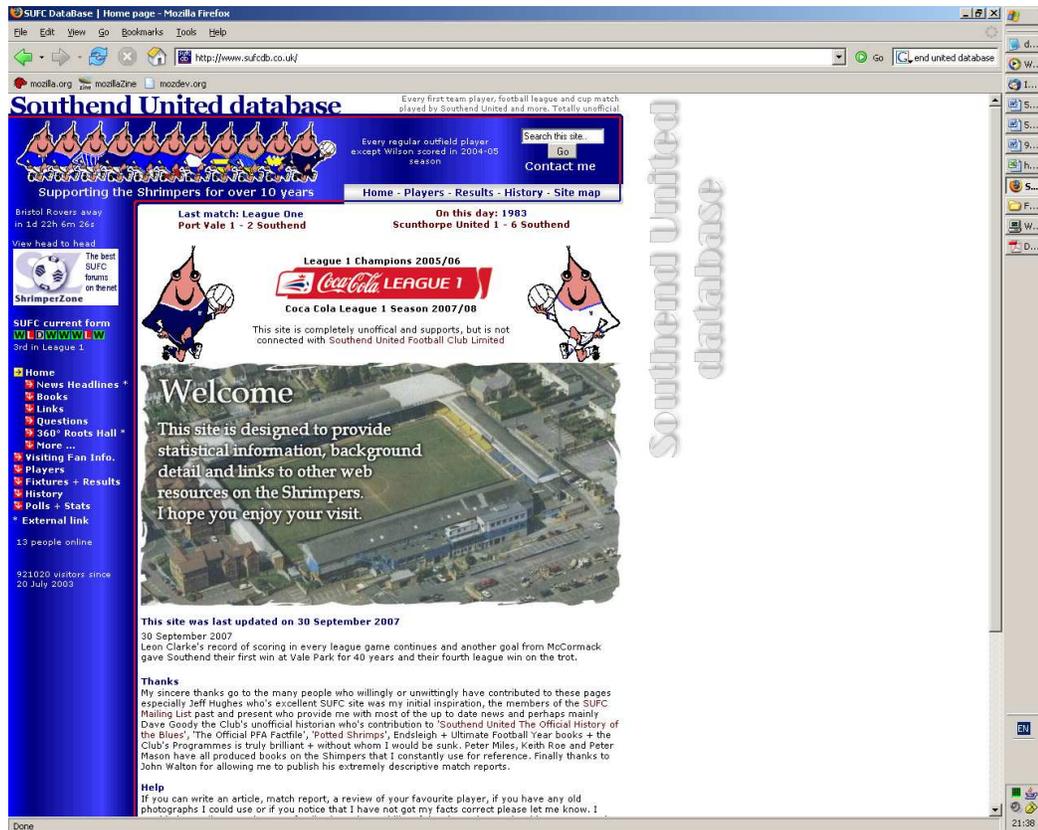


Figure 5.8 Southend United Database screenshot

In addition, the PHP has the basic information required from any such site, such as the squad details, the club history and so on. Furthermore, the PHP has a section detailing all the strips the club has ever had, as well as information on all previous managers. There are also links to other sites so as to provide match reports and latest news, and even information on books. Finally, the PHP has a logbook detailing work that has been done to the site, and even work scheduled to be done on the site. Without a doubt, the Southend United Database (2005) is the most noteworthy of all the PHPs this study has come across.

⁸⁶ According to the date in question (e.g. September 18th), the site provides a player birthday or the result of a match that took place in the past.

Table 5.7 has been created to present a summary of the features of each of the sites in a convenient form. Overall, these 7 sites demonstrate that PHP authors are able and willing to put significant effort into their sites. They appear to have the ability not only to create sites, but also to collect significant amounts of data. They are technically skilled and seem to be willing to work continuously over long periods of time.

They have shown that they can create standard sites of high quality (e.g. Geraldine's Viva Manchester United), create technically difficult sites which are the equivalent of online reference books (e.g. Southend United Database), create very professional looking sites (e.g. Par Sterner's Unofficial Walsall FC Homepage), incorporate media (e.g. the photographs on the Albion Album), create sites looking at facts from unusual standpoints (e.g. Black Country Derby) and they can also be quite imaginative in their creations (e.g. HTFC World). All in all, the PHP authors have shown that they are willing to commit a notable amount of effort, creating not just standard but unusual sites, where the added value is clearly present.

However, all this strengthens the argument that more order is required in the creation of such sites. These are examples of noteworthy sites, yet even here, there is an example of a site that has not been updated in years (i.e. Geraldine's Viva Manchester United site). The reason for the abandonment of the site is unknown, however, even though the site is still online and accessible (at the time of writing), it is quite likely that at some point, this site will be taken down, and the effort put in by the author will be lost. If the author had participated in a group effort, where the contributors could take part for as long or as short a period of time as they wanted, then perhaps this effort would not have been wasted (in the long term).

Name	Features	Comments
Geraldine's Viva Manchester United (Man Utd)	Player details, stadium information, club general history, club contact details, league and cup history, club records, forum and links.	No major updates since 2003 – Notable instances of unique information
Black Country Derby (Wolves/West Brom)	Statistics for every Wolves v West Brom encounter, including date, score, venue, attendance, goal scorers (with time of goal) and tournament.	Dedicated to derby matches rather than a one club
Albion Album (Brighton)	Match photos (sizeable collection), match information, links and mobile phone ringtones.	Photographs used on other sites (with permission)
Par Sterner's Unofficial Walsall FC Homepage (Walsall)	Squad details, player details, transfers, club information, club history and fixtures.	Extensive statistics available – High quality of implementation
HTFC-world (Huddersfield)	Comic graphical match reports, wallpapers, screenshots and author information.	Sizeable collection of comic graphical match reports
Southend United Database (Southend)	Detailed match information for every game, including date, opponents, attendance, goals scored, scorers, league position, top scorers, results against specific teams, performance graphs and player birthdays.	Very extensive statistics available – Technically sophisticated
The Robin's Nest (Cheltenham)	Statistics for every match Cheltenham Town have played since 1950 including results, goal scorers and end of year league tables. Also club history, descriptions of significant matches, links, wallpapers and winamp skins.	Extensive statistics available

Table 5.7 Features of notable PHPs

Finally, Table 5.8 lists the notable PHPs this study has come across along with their RICRs. However, having looked at the PHPs with the highest quality, it is not clear as to why they are connected to the teams that they are. Six of the 7 sites belong to teams outside the Premiership, adding weight to the earlier argument that the fans from lower divisions try to cover the gaps left by the media. Clearly, the sample size here is not significant enough to carry this argument on its own; regardless however, there does appear to be evidence to support further study.

Name	XX ⁸⁷	X	≤	N/A	Total Sections
Geraldine's Viva Manchester United (Man Utd)	1	13	9	7	29
Black Country Derby (Wolves)	2	4	0	1	5
Albion Album (Brighton)	3	3	5	6	14
Par Sterner's Unofficial Walsall FC Homepage (Walsall)	1	2	5	2	9
HTFC-world (Huddersfield)	3	3	0	2	5
Southend United Database (Southend)	14	19	4	4	27
The Robin's Nest (Cheltenham)	2	5	4	0	9

Table 5.8 Notable PHPs and their RICRs

5.7 Football Fan Perceptions Concerning PHPs

This objective looked at the figures that the questionnaires and interviews produced from the questions concerning the football fans' perceptions of PHPs. The responses from the 10 questions for the questionnaires are examined as well as a number of statements from the interviewees. A number of charts are shown in an attempt to illustrate these results more clearly. For this objective, the results of the questionnaires have been merged, excluding respondents (on the offline questionnaires) that do not use the web.

Perhaps a peculiar starting point has been chosen, but it appears that the predominant aspect of the relationship between PHPs and football fans, is the lack of knowledge that football fans have concerning PHPs. Respondents were asked questions concerning various aspects of PHPs (e.g. ease of use, comprehensiveness) to which they often replied "Don't Know". Table 5.9 shows the number of Don't Know replies in all categories.

⁸⁷ Figures not used in total column on the right.

	Frequency (n=413)	Percent (%)
Unique information	146	35.4
Innovative qualities	153	37.0
Accuracy	168	40.7
Comprehensiveness	173	41.9
How up to date	165	40.0
Ease of use	158	38.3
Speed of update	163	39.5
How far back does the information go	170	41.2
Interactivity	164	39.7
Independence	159	38.5

Table 5.9 Don't Know replies on questionnaires (online and offline)

The lowest number of Don't Know responses in any category is 35.4% (unique information), with the highest at 41.9% (comprehensiveness). The ten questions asked about the content of PHPs are identical on both questionnaires, with the Don't Know option always being the most common response. Clearly, football fans do not have a good understanding of PHPs.

Looking at impressions of fans towards PHPs as a whole, according to the questionnaire responses, as already implied, the modal class is always the Don't Know category. The second largest category is always the mid-point three (on the scale of one to five). The majority of the remaining responses lean towards the negative (i.e. one or two). In terms of percentages, on average 39.2% of respondents could not respond to questions asking about the content of PHPs. These figures make sense when read in conjunction with figures of PHP use amongst football fans. As mentioned in the previous chapter (see 4.3.4 The Role of the PHP Within the Web, p. 149), according to the questionnaires, many fans rarely or never use PHPs (51.3%).

Looking further into the data, a number of graphs representing the responses to the questions (stripped of the Don't Know categories) are available below.

The first graph (Figure 5.9) shows the up-to-date, innovative, comprehensive and ease-of-use categories. Of the four, three have negative perceptions as can be seen by the height of the graph on the left side. The category with the most negative

perception is up-to-date, whilst the only category with a positive perception is ease-of-use, where the right side of the graph is clearly taller than the left side.

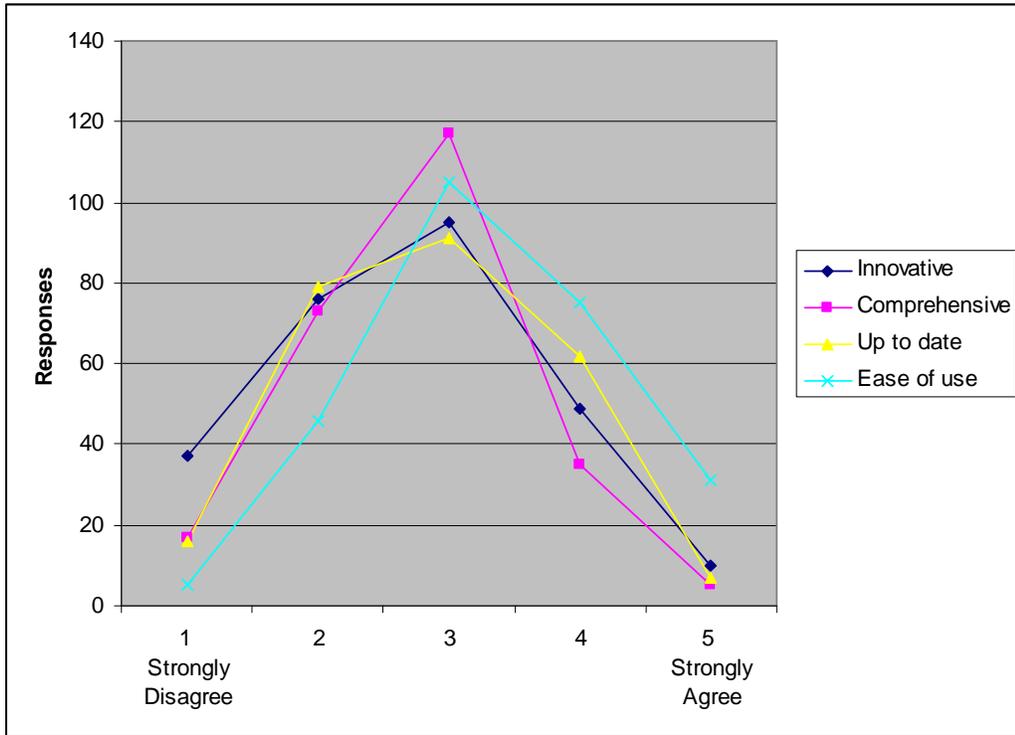


Figure 5.9 Fan Perceptions

This chart shows the responses of fans when they were asked if they thought PHPs were innovative, comprehensive, up to date and easy to use (including online and offline questionnaires).

Figure 5.10 is that which contains the PHP element which is seen in the most positive light of all the elements investigated by this study. Football fans clearly believe that the information found on PHPs is independent. These positive perceptions are even more apparent in light of the other three similarly negative categories (authority, interactivity and speed of update). All three have notably tall left sides, indicating that the perception of the remaining three categories by football fans is not favourable.

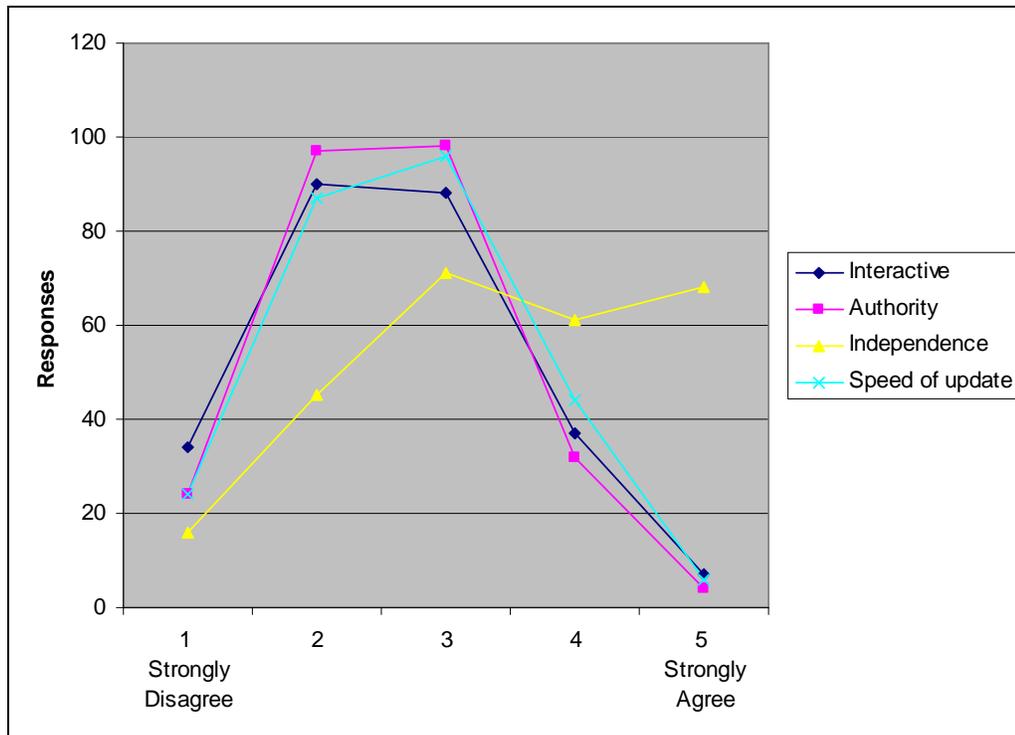


Figure 5.10 Fan Perceptions

This chart shows the responses of fans when they were asked if they thought PHPs were interactive, authoritative, independent and quick to update (including online and offline questionnaires).

The remaining three aspects of PHPs about which respondents were asked, were three aspects that the Stage One Web-Study also looked at. Here, the frequencies are shown with the Don't Know categories also represented (Table 5.10).

Scale	Unique Information		How Far Back		Accuracy	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
1	49	11.9	17	4.1	14	3.4
2	80	19.4	85	20.6	58	14.0
3	95	23.0	99	24.0	136	32.9
4	42	10.2	35	8.5	40	9.7
5	8	1.9	14	3.4	4	1.0
Don't Know	139	33.7	163	39.5	161	39.0
Total	413	100	413	100	413	100

Table 5.10 Replies from respondents when asked how Unique, Far Back and Accurate the information on PHPs was (including online and offline questionnaires).⁸⁸

The figures from Table 5.10 are shown on Figure 5.11. According to the responses (Figure 5.11), the uniqueness of information, how far back the information goes and the degree of accuracy, all have negative perceptions.

⁸⁸ On the scale shown, five (5) represents Strongly Agree and one (1) represents Strongly Disagree.

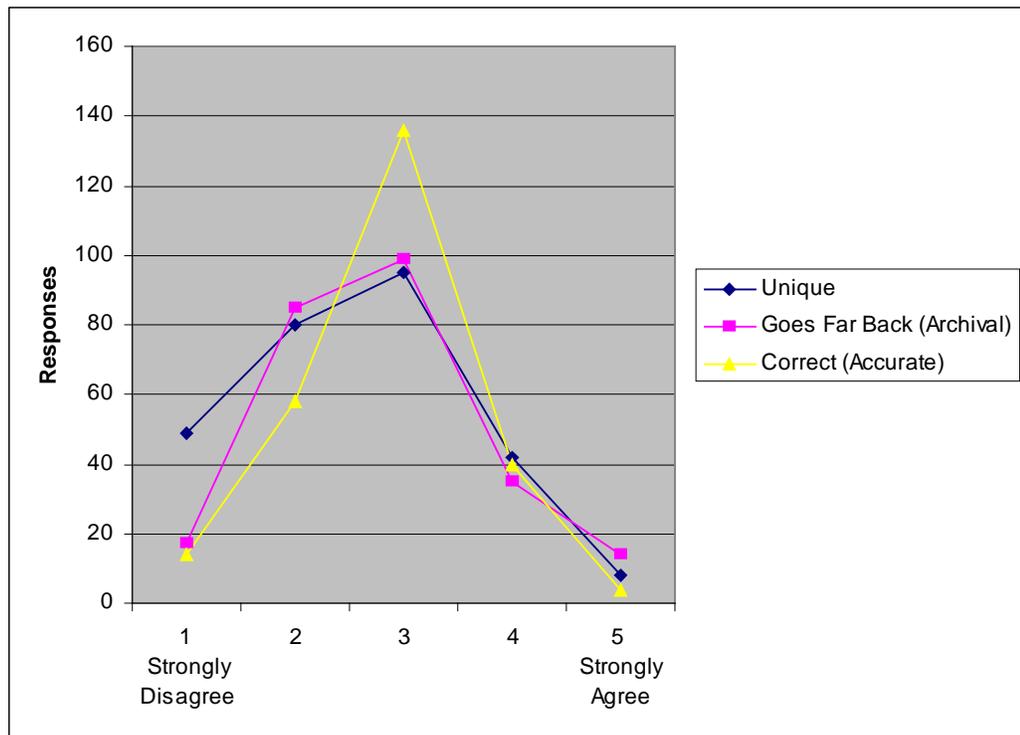


Figure 5.11 Fan Perceptions

This chart shows the responses of fans when they were asked if they thought PHPs were unique, archival (meaning the information found on them goes far back) and accurate (including online and offline questionnaires).

All in all, according to the questionnaires, it must be said that the general perception of the PHP by the football fan is either absent or clearly negative. The high proportion of Don't Know responses paints a picture suggesting that football fans are not overly concerned about PHPs as an entity in their own right. Meanwhile, the remaining fans, who feel they have a better understanding of PHPs, tend to think poorly of the PHP.

The interviews in general reinforced this unconcerned and negative perception of the PHP. Ten out of the thirty interviewees mentioned when probed that they use PHPs, however, none of the thirty listed a PHP as a site they primarily use.

Evidence of the lack of concern of PHPs from the interviewees exists in two forms: firstly, in the lack of comments and secondly, in the negative comments. At the end of

each interview, all the interviewees were told that part of the study was concerned with the role played by PHPs in providing information to football fans. However, only 8 of 30 (26.7%) had any comment to make, with the remaining 22 interviewees (73.3%) having nothing to say on the matter. Meanwhile those that had comments to make were mostly negative (five out of 8), suggesting once again, that fans are not overly impressed with PHPs as a source of information:

"PHPs don't really play a role at all." Look-BREN

"PHPs do not play a major role." Dion-LIV

"PHPs have never been important to me at all." Rict-WATF

However, there was one aspect of the interviews which revealed that the interviewees did indeed make some use of PHPs. The same interviewee (who unknowingly uses PHPs) made both the following statements, which to a certain extent, seem to contradict each other.

"PHPs could not possibly match the quality of information found on non-PHPs because a person could never be anywhere near resourceful enough." Moz-LIV

"Actually, there's a guy who makes Liverpool clips, you know, with music in the background, he puts them on his website. I check those out sometimes." Moz-LIV

Another interviewee also used PHPs unknowingly, in this instance talking about a PHP without realising it.

"I use Toonarama occasionally." Smar-NUFC

However, one statement seemed to sum up the overall impression given by the interviewees.

"Whether a site is a PHP or not, is completely irrelevant." Brad-LIV

This statement is natural when looked at from the point of view of the football fan who is concerned only with the actual information required. This is a fan (apparently like most fans) who wants to feed his information need and is not concerned about how that need is fed.

Overall, the interviews and the questionnaires seem to support the same idea. Fans for the greater part do not know much about PHPs, without this being a concern. Those who feel they understand PHPs a little more, have negative perceptions.

5.8 The Causes of Football Fan Perceptions Concerning PHPs

This eleventh objective lists the observations made about PHPs during the investigations that could be considered as causes for the generally negative perceptions that football fans have about PHPs.

Both the interviews and the questionnaires revealed that football fans feel that one of the most important aspects about footballing information is that it be up-to-date (54.7% of online questionnaire respondents, 30% of offline questionnaire respondents, 30% of interviewees). One interviewee specifically said that the single most important thing about a website must be that it is up-to-date, whilst another said that PHPs could be good if they were updated more often.

"Up-to-date [sic] most important thing about websites." Brok-MANU

"PHPs could be good if they were updated everyday, a couple of times a day."
Harl-NUFC

"PHPs though are not so good if they're not updated often enough." Harl-NUFC

"People who check sites all the time are looking for current news, not old stuff" Smar-NUFC.

Perhaps this is the primary cause for the low use of PHPs among football fans with web access. Having observed hundreds of PHPs, either in passing or in detail, it is very clear that this is a common failing of PHPs. Table 5.5 shows that 29.7% of the sections of PHPs checked contained a low degree of accuracy. However, as already mentioned, all of those sections have been deemed inaccurate because they have become outdated.

Furthermore, even this figure does not represent what football fans see, as it does not include all the PHPs that are too incomplete to be considered. The methodology used in this study has measures that prevent the entry of sites where 50% of the internal links are broken (dead) or otherwise unusable. Though the number of these sites was not recorded, it is significant. An example has already been mentioned earlier with Geraldine's Viva Manchester United site.

Meanwhile, a problem related to out of date PHPs, at which this study did not look, could very well be the time it takes to update a PHP. Major information providers update their sites as events take place, with live commentaries and score updates. It would not be unreasonable to assume that a single individual might be unable to provide such a service regularly. At times of major events, people with access to just the web, will be forced to use non-PHPs.

There are also numerous other problematic characteristics that PHPs have. One problem is the issue of PHPs that simply disappear. A number of sites that this study found as possible PHPs to investigate disappeared before they could be examined.

Furthermore, in addition to PHPs containing dead links and disappearing, another issue is that of the PHPs containing external links where internal links should be. An external link in this instance is not a link to an affiliated or relevant site in a "links to

other sites" section, rather a link that changes the central frame on the site to an external site, whilst keeping the content frames of the original site. Essentially, the site provides information taken directly from another site, in a manner which is obvious to the user. In these instances, users of the site might decide to go directly to the site being linked to. This study recorded the number of external links throughout the 42 PHPs, and found 17 external links in 368 sections (approximately 5%). Clearly, this is not a significantly common occurrence, however, it is still an aspect of PHPs that might be considered negative, and in that sense, makes a contribution to the poor image PHPs have with certain web users.

Another issue is the fact that PHPs tend to be notably smaller than non-PHPs. Sites belonging to large organisations such as the BBC Sport site, the SkySports site or any of the newspaper sites are vast repositories of footballing information, containing statistics, features and other details on all the clubs of all the professional divisions. These sites are actively seeking to be the number one port of call for the football fan. With these vast resources working at the hands of direct competitors, PHPs can at best hope to be the number two port of call (if not number 6 or 7). Meanwhile, there are non-PHPs focused solely on any specific club as well, not least the official sites of the clubs themselves, but also "web-franchises", which have individual sites for every club (e.g. rivals.net, Footy-Mad and so on). Any creator of a football PHP will soon realise that they have a good deal of competition, and this could well be a reason for the disappearance of many of the PHPs, and even a reason that more PHPs are not born in the first place. The idea of the PHP being unable to compete against the bigger sites was also brought up by an interviewee, whilst another commented that PHPs started reducing in numbers as "big" companies started creating footballing websites.

"PHPs could not possibly match the quality of information found on non-PHPs because a person could never be anywhere near resourceful enough." Moz-LIV

"PHPs were more popular in the beginning before the big companies picked up every Internet morsel they could find. However, as the companies picked these up, PHPs started dying." Hein-TOT

It is clearly a logical thought that suggests the unlikelihood of an individual alone to have the resources to create a site that could compete with sites of organisations with tens of staff members working on accomplishing a similar task.

Other causes include viewpoint and representation issues. PHPs can be created by anyone, including fans who are "unreasonably biased" or offensive (e.g. hooligans), fans who do not have an in depth understanding of English (e.g. fans from overseas), or for other reasons are unable to communicate in a suitable manner (e.g. lack of writing skills). Finally, there are aesthetic and technical issues too which can afflict any PHP. Sites can be created which do not function properly on various browsers, or those that do not meet other technological requirements (e.g. smaller than necessary bandwidth) due to a lack of technical understanding. They can also be created in a manner deemed ugly by the general public, due to a lack of "artistic" understanding.

It is quite probable that the three main causes for the negative perceptions of PHPs are the first three mentioned. These are not keeping up-to-date, not updating quickly enough and containing too many dead links. In addition, numerous more minor issues such as sites linking to other sites, relatively small sites, sites with unpopular viewpoints, sites written in poor English and sites with poor presentation or functionality, reinforce these negative perceptions originally created by the main causes.

5.9 Chapter Conclusion

In the process of meeting objectives 6, 7 and 8, it became apparent that PHPs can be deemed useful as an information resource for more than one reason. Indeed, the results of the research carried out were surprisingly positive. With many instances of unique information, both current and archival, and with the accuracy of the information also better than expected, there is added weight to any argument which supports the organisation of the efforts of PHP authors. Certainly, there is now empirical evidence (which was previously unavailable) that the efforts of these authors are notable, and every effort must be made to encourage a raise in the standard of work, in terms of regularity, so as to improve the longevity of the sites and

to avoid duplication. Clearly, there is a will on the part of the authors to provide a resource, and this will should be utilised in the best ways possible. For objective 9 (see 5.6 Factors that Affect the Availability and Quality of PHPs, p. 169), there is clearly an increase in the number of PHPs created for a club with improved performance on the pitch. However, concerning the factors which affect the quality of PHPs, there was no conclusive evidence, and this is something that should be looked into further. The results of this study lean towards a suggestion that PHP authors are more likely to create PHPs containing plentiful information if that information is not already available online; however, for a definitive conclusion further study is necessary.

Finally, the results concerning objectives 10 and 11 (see 5.7 Football Fan Perceptions Concerning PHPs, p. 183, and 5.8 The Causes of Football Fan Perceptions Concerning PHPs, p. 190) further strengthen arguments supporting the organisation of the efforts of PHP authors. Though there appear to be notable instances of unique information on PHPs, the two most noteworthy observations concerning ordinary fans and their relationship with PHPs are both negative. Football fans appear to either not know anything about PHPs or very seldom use them. However, as already mentioned, knowing that there is a significant amount of effort going into the creation of PHPs and the significant amount of unique information contained in them, is further cause to encourage the organisation of web author efforts. The following chapter looks at a blueprint for a site that would aid this organisation.

Chapter VI

Results III *- The Club Community Composite Page* *(CCCP)*

6.1 Introduction

Having carried out the Stage One research, and with the results of that research forming, it became apparent that the PHP as a source of information was notably less utilised than non-PHPs. As a result of this, having gained an understanding of both football fan information seeking behaviour and football PHPs, it was decided that an attempt should be made at producing a proposal in order to better utilise the efforts of web authors. Specifically, the objective set was:

12. To craft an outline for a communal site that can encourage web authors to work together in the creation of web pages for the online community.

In order to meet the objective, web author interviews and website collaboration investigations were carried out (see 3.6 Stage Two Web-Study, p. 112). The results of these interviews and investigations, helped craft an initial blueprint for a site called the Club Community Composite Page (CCCP), where web authors can coordinate their efforts to produce and maintain a football site. These results are described here with additional material available in Appendix K.

6.2 Web Author Interviews

The interviews with the web authors concentrated primarily on the motivations which encourage site creation and the difficulties faced by web authors. The purpose was to

gain a better understanding of individuals involved in web authoring so as to be able to determine more accurately whether an idea such as the CCCP could be feasible. The interview questions can be seen in Appendix A.

6.2.1 Demographics

For this portion of the research, eleven web authors were interviewed; the majority of whom (five out of 11) were aged 25-34. The youngest respondent was 17 years old with the oldest respondent at 52 years of age. All respondents were male, and 9 of them supported teams in the top tier of professional English football, with the remaining two supporting League 2 clubs.

Overall, it must be said that this number of respondents for this portion of the research was not ideal. However, this aspect of the research was carried out specifically for the purpose of ensuring that the CCCP was feasible, and this was accomplished.

A comparison of this sample, however, is not possible as another similar study interviewing web authors was not found. In addition, due to the size of this sample, a meaningful comparison is near impossible to make.

Age	Web Author Interviews (n=11)	Web Author Interviews (%)
16-24	1	9.1%
25-34	5	45.5%
35-44	3	27.3%
45-54	2	18.2%
55-64	0	0%
65+	0	0%
Gender		
Male	11	100%
Female	0	0%
Teams		
Premier League	9	81.8%
League 2	2	18.2%

Table 6.1 Web Author Demographics

6.2.2 Web Author Motivations

During these interviews, six types of motivations were declared.⁸⁹ The predominant motivation for the creation of websites seemed to be driven by the need to express oneself. This matches well with findings in related literature about PHP author motivations, with self-expression noted on several occasions (e.g. Buten 1996, Morris and Ogan 1996, Wynna and Katz 1997, Maruyama 1999, Papacharissi 2002b) not just in older PHPs but blogs as well (Nardi et al 2004b, Li 2005). Love of the team, along with interaction and the desire for feedback all came second (three out of 11). Two out of these three motivations (love of the team and a desire for feedback) are responses which cannot be matched to other findings in other literature, as these are questions that have not been asked in such a context previously, unlike ideas of interaction, which have commonly been mentioned (Morris and Ogan 1996, Maruyama 1999, Dominick 1999, Papacharissi 2002b). Finally, other less popular responses included being involved in team related work or an interest in computing.

Motivations	Number of Responses	Percentages (%)
Express myself	5	45.5
Love of the team	3	27.3
Want feedback	3	27.3
Want interaction	3	27.3
Already involved with team related work	2	18.2
Learning computing	1	9.1

Table 6.2 Motivations of Web Authors

Concerning the opposite related aspect of web authoring (“demotivation”), a high proportion of web authors were surprisingly immune to getting “demotivated”. Though the highest number of responses suggested that a lack of feedback (e.g. comments on their site) reduces their motivation to work (54.5%), a surprisingly high percentage (45.5%) maintained that nothing would stop them from working on the site. These results paint a picture of football web authors as a very dedicated group of people who have quite a strong desire to express themselves. These questions were

⁸⁹ A web author could provide more than one answer for this question.

asked in order to ascertain whether any major “demotivations” might act as obstacles towards the CCCP concept. However, nothing to this end was revealed from asking these questions.

With regards to the CCCP, these motivations (and demotivations) indicate that such a communal site should involve aspects from all the responses, including the possibility of learning computing, being involved in interaction and receiving feedback. The last of these is particularly important as the lack of it has also been put down as a demotivation.

Demotivation	Number of Responses	Percentage (%)
Lack of feedback	6	54.5
Nothing	5	45.5
Football related events	2	18.2

Table 6.3 Demotivations of Web Authors

6.2.3 Website Difficulties

When it came to obstacles in the paths of web creators, the majority claimed that the primary stumbling block was a lack of time. Ten out of the 11 respondents mentioned time as the primary issue. In addition, responses included other commitments as an obstacle (27.3%), the lack of technical ability (18.2%) and finally, one single respondent claimed that with the help of other people he might get more done. Meanwhile, three of the respondents who mentioned time, did so whilst claiming it was not a serious problem.

"Time, a little bit, but it's just a side project and I'm not too busy." HAVI-Barnet

"Of course, more time would be better, but I'm happy" WIDA-Everton

"Nothing really, more time would be nicer, but actually, I'm ok there too." PAUS- Everton

One issue which was not a great concern for most of the web authors was that of competition. Seven out of the 11 interviewees knew their competitors to a certain extent, but none were overly concerned about them. The most interesting response came from a respondent who thought of his competition also as a promoter of his own site.

"I know which sites I like and I link too, and I'm linked back. There's a small network of us. So I know that people probably come to me from them and vice versa... " HAMM- West Ham

Meanwhile, an issue where web authors had little agreement amongst themselves concerned the reasons behind the failure or abandonment of websites. The most popular reasons (though marginally) were a lack of time (27.3%) or not understanding the effort necessary to maintain such a site in the first place (27.3%). Meanwhile, though most respondents spoke with caution,⁹⁰ their subsequent statements were occasionally quite forceful.

"Laziness!" TIMD-Newcastle

"The authors lack passion" IMIT-Peterborough

"People just don't love enough. If you love it enough, you wouldn't stop"
WIDA-Everton

It appears that web authors themselves can only speculate on why other web authors might stop updating their sites, though time and dedication are both necessary in order one needs to maintain a site over a prolonged period. Once again, the literature on this subject in this context is quite scarce. As a result it is difficult to tie this in with the work of others.

From the point of view of the CCCP these results are far from discouraging. The basic ideas behind the CCCP are to increase the benefit currently gained from PHP authors.

⁹⁰ Many responses to this question started with a statement such as "I am not sure, but...".

As such, the basic aims include spreading the burden of having to constantly update a site amongst numerous authors, thereby also reducing the time required by any single author (yet not limiting the time one can spend), the lack of which seems to be the primary obstacle for web authors.

6.2.4 Site Popularity and User Participation

When it comes to keeping a site popular, the web authors overwhelmingly agreed that it must first and foremost be up-to-date. Nine out of 11 interviewees mentioned the importance of keeping the site updated. Other responses included having good material (18.2%) and making sure the site is well connected, with links to other sites (27.3%).

When it came to increasing participation on the site, the general sentiment went towards interesting subject matter. Eight out of the 11 interviewees gave a response which included the content of the site, with three saying they did not know. Interesting comments include:

"..the subject matter is important. Subjects with no 'answer' are very popular. E.g. should we get a new owner or not?" BOLD-Arsenal

"[site participation will increase] if you say something outrageous. People can't help but comment on something outrageous." GUS-Blackburn

"Detailed information which people will identify with. Essentially, what you write must be good. You can't just write anything." HAVI-Barnet

According to web authors, relevant and up-to-date content makes a site popular and encourages user participation.

On any site, the content depends on the individuals involved who produce this content. As such, it is difficult for a site blueprint to increase the quality of the authors' contributions. However, by reducing the commitment necessary from any

single individual, and by increasing the number of authors, it becomes possible to increase the probability of having good material (by giving authors more time to work on any single contribution). This same benefit of having to contribute fewer hours to the site should also increase the probability of being able to keep it up-to-date, something that will also increase the popularity of the site.

6.2.5 Site Collaboration

Interviewees were also asked what would further improve their sites, and whether collaboration would help. When it came to thinking of things that could be done to facilitate the improvement of their sites, authors were surprisingly unsure. Three mentioned possible technical improvements, two mentioned more people or contributors with 6 (54.5%) saying they either could not think of anything, or did not need anything.

When it was suggested that perhaps further collaboration might improve their sites, 6 interviewees (54.5%) agreed, with three thoughtful of whether it would actually make their lives easier or facilitate the improvement of their sites. Two interviewees were convinced collaboration with others would not help them in any way.

From a CCCP viewpoint, in this instance, though it is positive that the majority of authors think increased collaboration is a step forward, it is surprising to discover that as many as 45.5% of respondents were less than certain of this. Regardless, with over half the respondents positive (albeit from a sample of just 11), and collaborative sites already available online, there is reason enough to pursue the idea further.

6.2.6 Web Author Opinions And The CCCP Concept

All in all, the opinions of web authors appear to encourage a concept such as the CCCP. Web authors are clearly very dedicated to their sites, and are not easily dissuaded from publishing their content with the primary obstacle to site maintenance being time. A CCCP site, having a greater number of contributors, should reduce the amount of time any single author would spend on a site, and by doing so, facilitate an

improvement in the quality or “depth” of the content contributors produce. During interviews, it was revealed that sites must be up-to-date with relevant information to maintain popularity and encourage user participation. Again with less required of each individual contributor, this should be achievable. Another interesting revelation was that web authors are not overly concerned about their competition, and even consider their activities to be beneficial to their own sites. In terms of collaboration, overall, web authors seem to believe that working with more individuals would benefit their sites, though this is not something they actively sought to accomplish. However, it is the hope of the author of this work, that with a blueprint readily available (after the publication of this work), that perhaps this will change and more web authors will seek to work collectively.

6.3 Website Collaboration Investigations

Having interviewed web authors, the next step was to examine actual collaborative efforts online. Throughout the investigations, numerous examples were found where groups of people were working in collaboration. These include groups of people working on football website franchises (e.g. rivals.net), official supporters’ trusts, other supporter groups (e.g. Portsmouth fans in America), fanzines (including both online only fanzines and fanzines that started in paper form), forums and certain other unusual sites (see Appendix L).⁹¹

6.3.1 Football Website Franchises

One common variety of online collaboration, is the aforementioned, football website “franchise”. Many such sites (i.e. rivals.net sites) are created by a group of people who get together for this purpose. This often includes an editor, some technical support (often from the franchise company) and other contributors, such as match reporters, columnists and individuals that deal with statistics. Figure 6.3 shows a screenshot of such a site (the Manchester United rivals.net site).

⁹¹ A site might fall into more than one of the categories (e.g. fanzines whose online presence is a rivals.net site are quite popular) whilst other sites have multiple sections which could fall in other categories (e.g. it is common for sites to have forums as part of the site). However, the classification of the sites is beyond the scope of this project and serves no purpose for what is being studied at this time.

Such football site "franchises" include rivals.net, footy-mad and football network. In fact, this same concept is used for the vast majority of official club sites, with Premium TV providing the technological base for the clubs' online presence. However, in such cases, each club site is a separate entity from other sites (including having a separate URL) and so this scenario should not be confused with that of news sites (e.g. Sky Sports, BBC, newsnow) which have sections for each club. In addition to this, there are also sites which provide specific information about each club, for example the history of a club (<http://www.fchd.info>) or historical football kits (<http://www.historicalkits.co.uk>), where again information is provided on a team by team basis. However, in all cases, as the sites share the same design for the clubs, all such sites (or site subsections) were not examined in this process.⁹²

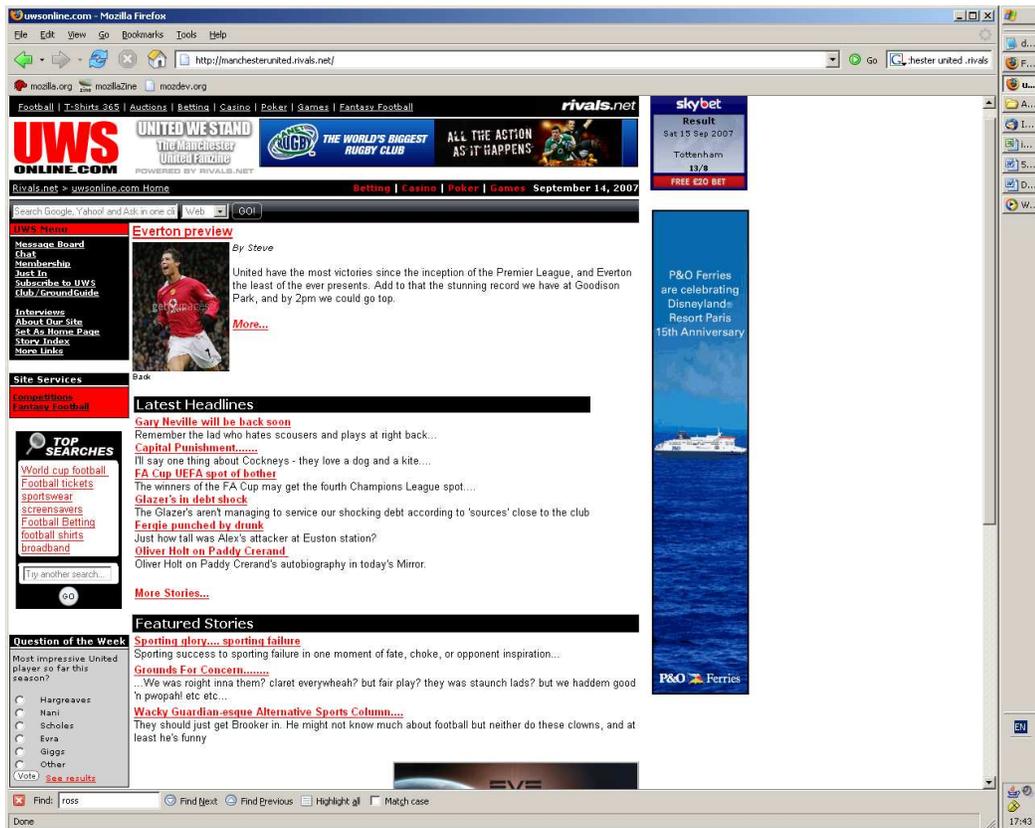


Figure 6.3 Screenshot of United We Stand (the Manchester United rivals.net site)

⁹² It must also be noted that in the search for sites, many sites were come across that appeared to be large scale collaborations using the phrase "for the fans, by the fans". This was come across on numerous occasions, however, it was discovered though that technically, it might be true, more often than not, it just means that the editor and webmaster are fans of the team in question.

6.3.2 Supporter Groups

A second common online collaboration in football is that of supporter groups. Concerning professional football in England, there is a strong presence of general supporter groups (often geographically distant from the club's hometown) and more specifically, official supporter trusts.

The long established supporter groups are simply groups of fans that try to come together for the purpose of following their team.⁹³ With the advent of the web, many of these groups have created an online presence in order to facilitate various tasks that need to be carried out as part of the everyday existence of such a supporter group. For example, whereas in the past, the organisation of meetings might have required numerous phone calls, now the time and place is simply advertised on the site. The sites created for such organisations are often the effort of several of the members of the organisation and vary from technically (and content-wise) very basic, with few features and little content, (e.g. astonvilla-supporters-eastcounties.com) to sophisticated, with many features and plenty of content (e.g. arsenalamerica.com).

Official supporter trusts, meanwhile, technically speaking, are supporter organisations that are registered as an Industrial and Provident Society (IPS) with the Financial Services Authority (FSA). The starting point of these trusts though, are supporters who pull together in order to create an entity in an attempt to exert more influence over the club and ensure that the opinions of the fans are heard (and taken into account). Currently, there are over 100 supporters' trusts across England, Wales and Scotland (Supporters Direct Site 2007), the vast majority of which have an online presence.

6.3.3 Fanzines

Fanzines in paper form have a history of being created from fan collaborations (Haynes 1995). Even before football, but certainly concerning football too, fanzines

⁹³ Some of these groups are then granted official status by their clubs if they meet certain criteria, though many are informal groups.

often started as the effort of a group of friends. From the very beginning they accepted fan contributions, something that they still do to this day. The online versions of fanzines meanwhile actively advertise for contributors (e.g. bcfcfanzine.com, onlinegooner.com). The online content of fanzines is generally similar to the offline content, with opinion articles about current affairs being the most popular piece of writing. Over the course of the Web-Studies, fanzine sites were come across for almost every professional club in the English leagues, showing a strong online presence.

6.3.4 Forums

Internet forums are sites that facilitate discussions on the web by providing functions to post and read statements by users. The general service offered is comparable to earlier newsgroups or mailing lists that predate the forum. Much like the newsgroups before them, the forums generally have a group of users who post and read regularly. Currently, such forums exist in abundance on the web, dedicated to almost any subject.

Football forums, like all forums, can be created and maintained by individuals or groups of individuals. However, the content of the forums is always supplied by the members. In the case of football, there are forums dedicated to every professional team, some extremely popular ones with upwards of 10,000 members (e.g. redcafe.net).

6.3.5 Notable Collaborative Efforts

Beyond the common types of collaborative sites mentioned above, the process of looking into sites of clubs in the top and bottom tier of professional football in England also revealed several notable sites as examples of collaborations. These were carefully examined as described earlier in this thesis (see 3.6 Stage Two Web-Study, p. 112), and descriptions of these sites along with screenshots can be found in Appendix L. On several of these sites, it was clear that the content placed on the site came from numerous sources.

Collaborative Website	Team
Addickted.net	Charlton Athletic
Beespedia	Brentford
Forever Man Utd (FMU)	Manchester United
Swindon-Town-FC	Swindon Town
ToffeeWeb	Everton
West Ham Online (WHO)	West Ham

Table 6.4 List of Notable Collaborative Websites

For a site to be a collaborative effort, there must be ways in which interested individuals can contribute to the content of the site. However, for this to occur, the site must advertise that it is looking for contributors, and that such a mechanism (which allows individuals to contribute) exists. This advertising can be done in a number of ways, including openly advertising on the front page (Forever Man Utd, Addickted.net), by strongly incorporating forum aspects (West Ham Online) or by creating a wiki site (Beespedia). Collaboration can then be incorporated in a site in various ways. Such a site can be “completely open” for collaboration, with contributors not needing to register and being allowed to create new sections (e.g. as many wiki sites are). The most “open” collaborative site come across in this research did require registration (i.e. Beespedia), though another site did allow the publication of certain pages without registration (Swindon-Town-FC). Most sites, including perhaps the most successful ones come across, exercised some form of editorial control before allowing publication, choosing to screen articles before publishing and then giving contributors increased permissions as they demonstrated over time that they are capable and trustworthy (e.g. Forever Man Utd, West Ham Online). Furthermore, the idea of creating a community is also present, with one site (addickted.net), describing itself as the result of an attempt to offer a "centralised place for all Charlton fans to use, enjoy and most importantly contribute". The site creators want the fans to have their Charlton blogs on the site, columnists to write columns and fans to provide match photos.

In terms of creating a site as a centre for a community, according to the investigations carried out, Addickted.net is alone in so openly and thoroughly adopting the concept. Having said this, ToffeeWeb is also notable for the number of ways it emphasizes fans and communities. The site has two separate sections on the front-page menu

called "The Fans" and "Community". Within The Fans there are 9 subsections and within Community, 6 subsections. Though not all of the subsections are directly linked to collaborative efforts, amongst them there are numerous fan articles commenting on current affairs, specific matches, what it is to be an Everton fan and letters to the editor. In addition, fan articles often make it to the front page of the site, giving the site a more communal feel.

Of these unusual collaborative efforts examined, most appeared to be quite popular, and up to date, though there were those with problems too. One site had organisational problems such as the same content appearing in more than one place (West Ham Online), whilst another had a noticeable lack of contributors (Beespedia), giving the site an incomplete feel, with a third site also suffering from a lack of contributors, though this had a less significant effect (Swindon-Town-FC has match reports for most games if not all).

6.3.6 PHPs to Collaboration

In addition to the notable collaborative efforts mentioned above, there is also the issue of a natural movement from PHPs to group efforts. With a few exceptions, PHPs in general are the work of single individuals, and as it stands, they appear to be providing unique information fairly consistently to the general public. However, as shown in this section overall (i.e. 6.3 Website Collaboration Investigations), there are numerous examples of online collaboration too, particularly in relation to football.

Furthermore, concerning non-football sites, there are a few high profile examples. A few notable ones include:

- The Open Directory Project (dmoz.org) which is a human-edited directory of the web (offers services similar to a search engine), constructed and maintained by a global community of volunteer editors.
- The OpenOffice.org Organisation which, through a gathering of volunteer developers and end-users, has created an Office Suite comparable to Microsoft's Office software.

- Wikipedia, which is a multilingual, web-based free content encyclopaedia project, where articles can be written and edited by almost anyone with access to the website.

However, other instances were also come across in several of the interviews carried out, and the parts of the Web-Studies (both in Stage One and Two). Throughout the interviews there were two mentions of PHP authors moving away from their own site. In one instance, "Ormondroyd", (author of Ormondroyd's Virtual Match Reports) moved from his PHP to the Guardian Newspaper. In another instance, a PHP author was given the position of webmaster at the official club site⁹⁴ (Leyton Orient). In other instances, sites that were originally PHPs started advertising for contributors, an idea which certainly existed before the advent of the web, originally in paper fanzines.

Meanwhile, on numerous occasions online, instances were come across where the author of a PHP had abandoned his or her own creation to work on a larger web project (e.g. one of the "franchise" sites such as rivals.net) and these are shown in Table 6.5. Meanwhile, during the time it took to complete this thesis, there was even a site that went from a PHP to a non-PHP (SteveRoy.com).

⁹⁴ This was mentioned by an interviewee, and attempts were made to verify this information by contacting the club site. Unfortunately, no response was ever received and hence this information has not been verified.

Team	Original Site	Original URL	Current State
Rochdale	The Alternative Website	http://members.aol.com/upthedale/index.html	All three sites became the Rochdale rivals.net site (Rochdaleafc.com)
Rochdale	DIAZ - The Rochdale Audio Zone	http://website.lineone.net/%7Ediazone	
Rochdale	Halliwel's site	http://hometown.aol.com/halliwel69/index.html	
Brighton and Hove Albion	Bob Gear's Albion Links page	http://www.bhalinks.co.uk/	Became the Brighton-Mad site
Walsall	SteveRoy.com	http://www.steveroy.com	Became UpTheSaddlers.com -a popular non-PHP

Table 6.5 Instances of PHPs that became non-PHPs

As already mentioned, these discoveries were peripheral to the research, and were not specifically looked out for. However, having been noticed they have been included here, as relevant to the research as a whole. Specifically, the movement of the authors in these sites from PHPs to group efforts bodes well for the CCCP as a concept. Certainly, web authors in this instance have shown that they agree with such a concept, and by extension, such a blueprint is beneficial to this field.

6.3.7 Summary of Online Collaborative Efforts

Overall, it is clear that there are currently numerous examples of online collaboration. There are a variety of ideas being utilised in a number of ways, ranging from almost complete freedom to update a site (e.g. Beespedia) to a collection of contributions centrally (e.g. ToffeeWeb). Various controls are placed over who can and cannot edit the site directly, again to varying degrees (e.g. Swindon-Town-FC allows direct uncontrolled publication, whereas FMU gradually grants more rights to contributors). However, there is also clearly room for improvement, with more being done to

encourage community (i.e. PHP author) collaboration, for example by combining aspects of the sites examined. The design for the CCCP does so (see 6.4.2 The Club Community Composite Page (CCCP), p. 213), by taking ideas from these sites, and combining them to create a blueprint that should be applicable to any football club.

One final interesting observation concerning online collaborations is that more communal efforts appear to be found online for the Premiership teams rather than the League 2 teams. Again, this does not appear to support the idea that web authors are fully grasping the opportunity to fill in the gaps in the information supply about their teams (again with the assumption that these gaps are more prevalent in the lower divisions). Regardless, this should not have a major effect on the plausibility of a concept such as the CCCP.

6.4 Crafting A Communal Site Outline

Throughout the course of this research, as a result of examining, closely or in passing (see 3.4 Stage One Web-Study, p. 86, and 3.6 Stage Two Web-Study, p. 112), hundreds of football websites, it became possible for the author to put together a blueprint for a recommended collaborative effort by combining attributes of several collaborative sites already in existence.

6.4.1 Collaborative Sites - An Ideal Tool

Throughout the course of this study, several points have come up concerning football fans' efforts and needs. According to the results, fans of English football clearly enjoy discussing football, whether this be online or in person. They enjoy reading about their teams, but an important aspect is that the discussion topics be recent. In addition to this, they are happy to use the web to meet their football information needs. Those who are web authors like to create sites concerning the clubs they support, expressing their ideas and getting feedback. They are dedicated to this process, and are supportive of collaborative projects. As the popularity of the web spreads, and as it becomes more commonplace, it is likely that these trends will continue. Internet communities exist supporting most (if not all) professional football teams in England,

and with the help of these communities, it should be possible to maximise the benefits of the efforts put in by PHP and other web authors, as well as others active in the football Internet community.

Based on the results of this thesis up to this point, a suitable tool or method for the dissemination of football information should have certain attributes for it to become popular. These include the following:

Be available online.

As already mentioned, a very high proportion of football fans currently use the Internet (e.g. 95% according to the FA Premier League 2008) and with the trend of web use in the UK steadily rising (National Statistics Omnibus Survey 2005b)⁹⁵ it would make sense for this information tool to be available online. Perhaps the strongest case for this argument is that currently, 95% of 16 to 24 year-olds have used the Internet (National Statistics Omnibus Survey 2005a) showing a low likelihood of reduced web-use in the future.

Have up to the minute news (including match reports and transfer information).

The most important reason for using any tool or method according to the questionnaires was that the tool or method had the most up-to-date information, whilst Green (1999) has produced results which also support this idea (the most common searches are for information which has been recently updated). To further reinforce this notion, most of the web authors interviewed, as well as several of the football fan interviewees, mentioned that the most important aspect of the sites they used was that they be up-to-date.

Be convenient.

The same part of this study which revealed that up-to-date information was the most important, revealed that convenience is the second most important aspect of information. Both questionnaires and interviews strongly highlighted the importance of convenience in the search for information.

⁹⁵ From 2002 to 2006, there has been a 26% increase in the number of households that have Internet access.

Be free.

This "free" attribute of any information disseminating tool or method has not been explicitly identified. However, the majority of tools and methods mentioned by fans did not have additional costs associated with them. Concerning websites in particular, there was very sparse mention of any financial aspect from the fan's point of view.

Offer a sense of community.

Major clubs have long been the focal point of the local community (Fishwick 1989) and in return, the prosperity of football has benefited the local community (e.g. BBC Website 2004). Such ideas were supported by numerous interviewees suggesting that the importance of football is closely linked to the community. As such, it makes sense to incorporate communal ideas into the creation of an information dissemination tool.

Have interactive aspects (e.g. a forum).

The figures from this thesis suggest that a high proportion of football fans use football forums and messageboards. In addition to this, web author interviewees mentioned the benefit of having a forum for the purpose of increasing traffic to a site, but also as a feedback mechanism. Furthermore, the fact that such interactive aspects would also help reinforce the sense of community mentioned earlier, makes it sensible to include this as an attribute that a communal information disseminating site should have.

Have plentiful additional and peripheral information.

According to this thesis, completeness of information was not an essential factor in determining whether a fan would use a specific site or not. However, for the purpose of catering to as large an audience as possible, this has been included as a required quality.

6.4.2 The Club Community Composite Page (CCCP)

The Club Community Composite Page (CCCP) is the name given to a blueprint of a site which will help maximise the end product of the effort of people who have the will to produce material about their own football club (clearly including PHP authors). Sites already exist that meet the seven broad guidelines mentioned above (e.g. those in Table 6.4), and these were examined whilst preparing a design for sites such as these, that could be implemented to represent any professional football club in England.

Conceptually, the idea of the CCCP is that the site has the approximate layout of a traditional football website (e.g. a footy-mad site), yet be immediately editable by the community. The reason for this is so as to encourage potential contributors as much as possible. As for the appearance of the CCCP, from the point of view of the users, it should be very similar to that of ordinary football sites. Traditional sections of such sites such as fixtures, results, match reports and transfers should be in place to make the site attractive to potential users (Green 1999) as should player information, statistics, history and additional features which should help appeal to as wide an area of user as possible. As a starting point, some of these sections would have to be filled by the initial individual or group involved with the launch of the site. To ensure initial interest, the latest news section should be as up-to-date as possible. A forum should also be included as part of the site (something encouraged by web author interviewees). Basic necessary instructions, such as how to register and post or edit pages should also be readily available, either in the help section or any suitable area of high visibility. This is necessary to fit in with the idea that the site should be editable by the community. Finally, the way the CCCP is intended to function (i.e. with the community editing the site) should be clearly described, perhaps on the front page, or in a location with equally high visibility.

Possible Implementation

The CCCP is a model that should theoretically work for most, if not all, professional football clubs in England. Some form of community site currently exists for both the

top professional league in England and the bottom professional league. Meanwhile, for any club in the lower divisions, it is not uncommon to have a handful of PHPs for a club that has a 7,000 or 8,000 fan average attendance. For example, in League 1, according to the samples used in this study,⁹⁶ the average number of PHPs per team is 4.5. The corresponding average attendance (for the same four teams used in each division) is 7,623 fans. Meanwhile, though the quality of the PHPs created for these teams vary, there are certainly numerous instances where fans willing to put the effort in (collectively speaking) would benefit from better opportunities to do so. The interviews of the web authors carried out in this thesis reinforce this notion, with most web authors agreeing that additional collaboration would benefit their sites (see 6.2.5 Site Collaboration, p. 201).

For any implementation, as already mentioned, a notable amount of effort will be necessary to produce and launch the site. Some central organisation is required, an individual or a group who will take the initiative to publish the site and advertise it to begin with. Once the site is launched and active, it will still need some central administration to ensure that a standard is maintained, both in terms of informational context and technology, while the CCCP tries to reach a certain point of popularity. However, once a certain point of activity is reached, the administrative efforts necessary should reduce significantly.

Open to the Public - Ideas Currently in Use

As one would expect, there is no commonly accepted single formula for creating a collaborative web effort. As such, there currently exist a variety of types of collaborations, including online fanzines, blogs, forums, wiki sites, "ordinary" sites (with many contributors) and even a site which calls itself a community site has been come across. Many of these sites are also "cross-breeds", having attributes of more than one type of site. For example, some instances of collaborative sites found are clearly primarily forums with certain features that give them the appearance of a traditional football site. An example of this is West Ham Online (WHO) where a

⁹⁶ These figures are not by any means definitive, but rather just for guidance purposes.

navigation menu and an article can both be found on the front page (not the customary appearance of a forum).

However, exactly how publicly open the collaborative sites are is not immediately clear. In the case of WHO, when it comes to getting an article on the front page, the site administrators want to exercise some editorial control. Two sites which are closer to what a CCCP should be like are Forever Man Utd (FMU) and Addickted.net. However, here too there is not much freedom afforded to potential contributors. Both sites advertise for contributors, but in both cases the sites request that contact is made (by the potential contributor) and discussions undertaken before allowing contributions to be made and materials to be published. At least on WHO, contributors can start publishing immediately, even if only on certain sections of the site. Other collaborative sites discovered also want to exercise a certain degree of editorial control before allowing articles to reach certain parts of the site, with the exception of Beespedia, which is a straightforward wiki site. Beespedia though, has other issues, in that it appears to have not taken at all into consideration the informational needs of football fans (e.g. notable popular sections are missing, such latest news, match reports, forums and so on). In order to meet the needs of football fans, a CCCP must also have all the usual sections of a site with particular importance placed on news, results, fixtures, match reports and transfers (Green 1999). Another collaborative site, Swindon-Town-FC, has many of these sections (with the crucial exception of a Latest News Sections) and provides a mechanism for automatic submission and publication of various materials, specifically Match Reports and Match Photos (a feature that was not found on any other site). Many of the collaborative efforts found had all the necessary sections (e.g. Forever Man Utd, ToffeeWeb), though certain sites prefer to "outsource" them (e.g. West Ham Online has a direct link to parts of the West Ham official site).

However, something that should also be possible is the ability for contributors to add further sections as deemed appropriate, something that is only really possible with Beespedia. The reason that giving contributors such freedom is so important is born from the fact that the effort required to maintain a CCCP (from an administrative point of view) should be minimal. This is an idea that filtered through from the interviews with web authors (the vast majority of whom cited time as their primary

obstacle). The maintenance of a website is clearly a very time consuming activity, and as a result, any process which can reduce the workload of a web author is very welcome by the authors themselves. This is also linked to the need for feedback. In order for web authors to keep their sites active, it is important for them to receive feedback. Activity on the site is one type of feedback (according to two web authors), so in addition to reducing the workload of web authors, the contributions of others should motivate them further. Hence, the idea of the CCCP site is that the community can take over many of the roles that the single PHP authors would have played in the past, whilst also encouraging them. This in turn acts as one additional layer of protection against the possibility of a site being abandoned or left incomplete. In fact, once active, in addition to being constantly updated, the site will also be constantly evolving and expanding.

Potential Obstacles

Naturally, there will also be potential problems with the CCCP which would have to be overcome. The first and most obvious is the need for a "critical mass of participants". Many currently active sites appear to be using a handful of core contributors extensively to maintain the site (something that was noticed in the Stage Two Web-Study). Though this is a formula that appears to be working, it still places significant responsibility on certain people. The idea behind the CCCP is to reduce this as much as possible, by spreading the weight of this responsibility over as many contributors as possible. However, should the critical mass of participants not be attained, the site would never reach the desired state of functionality. The site needs to have a foundation to start, and for this reason, some initial effort will inevitably be necessary to obtain this critical mass.

Meanwhile, an unrelated second potential problem is the competitive nature of football. Clearly, if a site is open to all for editing and updating purposes, it is possible that fans of rival clubs could attempt to unsuitably alter the content of the site, in an act that could be called "electronic vandalism".⁹⁷ Though not mentioned by web authors in the interviews carried out, such precedents exist online, with one of the

⁹⁷ This was not mentioned by the web authors in the interviews. However, those interviewed were not involved in sites open to the public for editing.

high profile incidents involving The Million Dollar Homepage (BBC Website 2006f), though in that instance, the motives of the offenders were purely of a criminal nature. Regardless, safeguards should be put into place to prevent a sudden collapse of the site (see Appendix K).

Finally, a CCCP will often be in an "incomplete" state, as new sections of the site are proposed and created. However, the CCCP should still be thoroughly usable and maintain a high level of interest, particularly if it has an up-to-date and reasonably active latest news section (e.g. including live updates on matches and transfers). Indeed, all the collaborative sites discovered through the Stage Two studies are still active at the time of writing.

6.5 Chapter Conclusion

Naturally, as already mentioned the CCCP will match all the characteristics of the ideal tool as suggested by the findings of this thesis. These include being available online, having the latest information, being free, convenient and offering an enhanced sense of community. One key aspect of the CCCP though, is that it be as open as possible for editing by the community at large. As discussed, this may have negative consequences (e.g. electronic vandalism), and as a result, it may take several attempts for the right balance to be found. However, as much as possible should be done to encourage community participation. Existing sites have gone some way to encourage member participation (though there are still obstacles to instant editing, and they still employ ample editorial control) yet more is felt could be done. Overall, the CCCP should complement the facilities already provided by numerous sites, both PHPs and non-PHPs. In doing so, it should go some way in increasing the options available to both providers and consumers of information. Whereas providers will have the ability to feel part of a larger project, one on which they can work as their lives allow, whilst also enhancing links within the club's online (and possibly offline) community, consumers will have access to an additional, reliable, relatively independent and representative view of the goings on of their clubs. Meanwhile, as football has helped the community in the past (Oughton et al 2003), so such a site can help football's virtual community. Advocates of virtual communities firmly believe that they are a

rich source of information (Wellman and Gulia 1999, Rheingold 1992), and indeed studies have shown that strangers in such communities give other strangers plenty of information (Constant, Sproull and Kiesler 1996).

On the whole, the CCCP should survive and prosper according to the results and indications of the research carried out in this thesis. Collectively, there is a lot of effort put into individual PHPs concerning football, and indeed there are successful collaborative efforts already in place. In addition, the Internet community as a whole has shown that it is possible to collectively create and maintain thoroughly resourceful sites (e.g. wikipedia.org). As such, the majority of fans involved in both providing and consuming football information, should make good use of such additional means of information management. The following chapter discusses the findings of this thesis as a whole and proposes future work to further facilitate the information seeking and finding processes in this field.

Chapter VII

Discussions and Conclusions

7.1 Introduction

In the course of meeting its aims, this study examined various aspects of the information seeking behaviour of football fans, and within that sphere the roles of the web and more specifically, the Personal Home Page. The overall purpose for this was to gain an understanding of the PHP as an information resource, the study of which has implications for numerous sub-fields of information science, such as information seeking behaviour, web genre studies, grey literature and virtual communities to name but a few. The scope of the study included teams from the top four professional football leagues in England.

It was determined early on, that in order to gain as full an understanding as possible of the PHP as an information resource, it would be ideal to use a holistic approach. This approach meant that the information seeking behaviour of football fans would be looked at as well as the informational content of the PHPs, so as to understand the point of view of the football fan as well as the actual content of the PHPs. Hence, of the 12 objectives set, the first five (see Table 7.1) provide the foundation from which the PHP can be examined and are primarily concerned with the football fan's point of view. The examination of the PHPs is carried out in a more in-depth manner through objectives 6 to 11 (see Table 7.2). Having completed the research into the specific goals, an attempt was made at formulating a recommendation to move the field forward (the Club Community Composite Page). However, further research was deemed necessary to strengthen the basis of this recommendation, and hence a 12th objective (Table 7.3) was set.

7.2 Foundations of the Thesis - Football Fan Information Seeking

The first five objectives started by determining the tools and methods fans use to acquire their footballing information (objective 1), and it is clear that football fans have an affection for the web. When asked how they would search for a piece of information, regardless of the data collection method used, the greatest number of responses pointed to the web. The lowest figure was from the offline questionnaires where it was 72.5%, and the highest figure was from the interviewees where it was 93.3%. These figures more or less match figures by the Premier League (FA Premier League 2005) which show that approximately 81% of football fans (of Premier League teams) accessed the Internet for football purposes "in the last month", but also match research in other fields (such as serious leisure) which show that the Internet is a primary source of information for information seeking outside professional environments (Nicholas et al 2003, Rieh 2004, PEW 2005).

Objectives 1 to 5	Conclusion
1 - To determine the tools and methods ordinary football fans use to acquire their "footballing" information.	The vast majority of football fans in all data collection methods cited websites as their preferred method of looking for a specific piece of information. Other popular tools or methods were asking friends or colleagues, and checking reference books.
2 - To determine the reasons for which these tools and methods are used	The main reason for this choice of web appears to be that it is convenient to do so, and that the information is up-to-date.
3 - To determine the significance of the role played by the web in providing this "footballing" information.	More football fans use websites more often than any other tool or method. In addition, websites are the first tool or method that football fans use when looking for a piece of information.
4 - To determine the significance of the role played by the PHP within the role of the web in providing this "footballing" information.	Personal Home Pages are of the least popular types of sites that football fans use. The majority of football fans use PHPs once a month or less.
5 - To determine whether, all in all, football fans are satisfied with the availability of "footballing" information on the web.	Overall, football fans feel that the web meets their information needs. Also, the majority of football fans who use the web find what they are looking for more often than not, with those who spend more time on the web, finding their required information more often than those who spend less time.

Table 7.1 Objectives 1 to 5

With the growth of the number of households with access to the Internet it is apparent that this trend will only grow more in the near future, as the number of households with Internet access nears a stable number. Two fans in the interviews mentioned that they do not use the Internet for footballing information purposes, however, as future generations grow up with the Internet, it is near inconceivable that those who seek information for such purposes do not use the web. Clearly, the Internet is here to stay, and so are football websites.

The second most common response according to all data collection methods was asking friends or colleagues. A number of categories also had response rates of over 10% in all data collection methods. These included TV, local newspapers, national newspapers, and teletext. One other notable category was the reference book. According to both sets of questionnaires, the reference book was the third most popular tool when carrying out an information search. However, only two of the fans interviewed mentioned the reference book at all. A variety of other methods were given a mention in one or more of the data collection methods, though all with response figures at under 10%. These included radio, DVDs/videos, magazines, fanzines, text alerts, mobile Internet, computer games and shareholder's reports. Here, these results cannot be verified by other studies, as other studies have not examined these aspects of football fan information seeking behaviour.

When it comes to the amount of time spent on the web (objective 3), 66.7% of fans on the offline questionnaires said they use the Internet to read about football for at least 10 minutes a day, with almost a quarter (23.5%) saying they read about football for at least an hour. Naturally, on the online questionnaires the figures were significantly higher at 97% reading for at least 10 minutes, and almost half (45.3%) reading for at least an hour. Clearly, there are football fans who spend an even more significant part of any day reading about football on the web. According to the online questionnaires, 12% of fans spend more than two hours a day reading about football on the web. According to the offline questionnaire this figure is at 3.9%. Even the interviews (where this specific question was not asked) suggested a similar figure as two interviewees (6.6%) mentioned being reprimanded at work for spending too much time looking at football websites. Again, the amount of time spent on the web has not been measured by other studies, though one survey indicates a similar mentality in

football fans, stating that over 60% of fans use the Internet for football information at least two or three times a week (FA Premier League Survey 2005).

Both questionnaires also determined which methods and tools fans were heavy users of (objective 1 and objective 4). In this instance, the results were slightly more varied. Once again, the websites came out on top, significantly more so in the online questionnaire at 63.4% but also offline where the websites were still the category with the most responses at 39.2%. Again, on both questionnaires, there were a few groups with more than 10% of the responses. Online, more than 10% of respondents were heavy users of teletext, TV and local newspapers. According to the offline questionnaires, the categories with a higher than 10% response rate were friends and colleagues, TV, local newspapers and radio. Once again, in this instance these results cannot be verified by other studies.

Looking at why these tools and methods are used (objective 2), according to the questionnaires (online and offline combined), the primary reason why heavy Internet users use the web was to get the latest information (52.8%). This can be verified to a certain extent by another study (Green 1999), where the most popular (by far) types of information sought were all very recent (News, Results, Match Reports etc). The second category of significance which received 20.9% of responses was convenience. Two reasons which were relatively unimportant to fans were the authority of the websites (2.8%) and the websites being the only source of information (4.7%).

As for the reasons of use of any other given tool or method, generally speaking up-to-date and convenience are the most popular reasons, with completeness also occasionally a popular reason. Again, the authority of a particular medium was not considered an issue for the vast majority of respondents. Meanwhile, the least popular reason for using a specific tool or method was because it was the only source of information, suggesting that the same information is finding itself available on various media.

Concerning football fans' overall opinion of the availability of information with regards to football, the results of this thesis suggest that there is information available in abundance. Looking specifically at the web, the interviews and questionnaires both

indicated that football fans are happy with the amount of information available. In interviews, 80% of the participants said they were happy with the overall quality of the sites they used, whereas in the questionnaires the vast majority of respondents responded positively (selecting four or higher on a five point scale) when asked if the web meets their footballing information needs (89.2% online, 72.6% offline). These results also match results of other studies (Green 1999, FA Premier League Survey 2005).

All in all, websites are clearly the most popular destination for fans of professional football in England who are looking for information. Looking at these results together, it appears that the web as a medium has had an effect, whereby football fans spend more time reading about footballing information than ever before. Though we cannot say this with full certainty, the indicators certainly point in that direction. The reasons for using the Internet are convenience and to get the latest news. There has never before been such a convenient medium which provides such up to date information. Certainly, television and radio provide up to the minute information, however, one could not control (certainly not in the past) when one could watch or listen to a football program. Now with the Internet available around the clock, people are able to spend hours acquiring footballing information. In the past, this would not have been possible. It is difficult to say whether other football information outlets, such as newspapers, are suffering at the hands of the Internet (the investigation of which is beyond the scope of this study), however it appears that regardless of this, the availability of footballing information has increased.

7.3 Personal Home Page Conclusions

Moving on to the PHP (and objectives 6 to 11), of all the relevant studies, surveys and investigations, never has one been carried out examining the PHPs as an information resource on a specific subject and the relevant potential informational benefits. Studies on PHPs have been carried out in several academic fields (e.g. computing, information science, psychology, journalism etc.) and various aspects have been studied, including the examination of their characteristics and design (de Saint Georges 1997, Bates and Lu 1997, Dominick 1999, Dillon and Gushrowski 2000,

Papacharissi 2002a) and reflections that they have on "real life", such as gender differences on the web (Arnold and Miller 2000, Flanagin and Metzger 2003), identity and self-presentation issues (Wynn and Katz 1997, Chandler 1998, Dominick 1999, Nomura, Ishida and Yokozawa 2001). Without doubt however, this is the first instance in which the content of the PHPs has been compared to the content of non-PHPs (i.e. traditional information providers). In that sense, this thesis provides a first insight into the PHP as a potentially useful resource. This in itself presents certain difficulties. For instance, several aspects of this research are difficult to verify using external sources, simply because there are no external sources to compare to. Regardless, for the majority of the research, there are at least some relevant external studies and in those instances, the results of this thesis are compared and discussed.

Objectives 6 to 11	Conclusion
6 - To determine the number of sections that have instances of unique information on football PHPs.	At least 15% of PHP sections (on PHPs from each division) have information which is not available elsewhere on the web.
7 - To determine the number of sections that have instances of archival storage on football PHPs.	Overall, on PHPs, approximately one in 11 sections has information which is unique just because it is older information (e.g. the information is from 1996).
8 - To determine the degree of accuracy of the information available on football PHPs.	Over 70% of information found on PHPs has a high degree of accuracy. In general, the vast majority of information on PHPs is accurate when published, though this information becomes less accurate over time (as the sites are not updated).
9 - To determine the factors that have an effect on the availability and quality of PHPs.	A team which has more success on the pitch will have more PHPs about them created by fans. There is also an indication that the quality of PHPs is likely to be higher if a team is in the lower divisions (though at this time, there is not enough evidence to make a firm case for such an argument).
10 - To determine the perceptions that football fans have concerning PHPs.	At least 30% of football fans were not conversant with PHPs. The remaining fans had overall negative perceptions of 9 of the 11 categories examined concerning information on PHPs. The 9 categories were accuracy, comprehensiveness, innovativeness, up-to-date, interactivity, authority, speed of update, how far back it goes and how unique it is. The only categories seen in a positive light were independence and ease-of-use.
11 - To determine the causes behind the perceptions that football fans have of PHPs.	Football fans most likely have negative perceptions of PHPs because they are often not kept up-to-date, they are not updated quickly enough (even when they are up-to-date), and they contain too many dead or malfunctioning links. Other possible causes include their relatively small size, their links to external sites, their occasionally inappropriate viewpoints and poor presentation.

Table 7.2 Objectives 6 to 11

7.3.1 Characteristics

Throughout the research carried out for this study, many PHPs were examined, and even more were briefly observed (specifically for the purposes of objectives 6, 7 and 8). In many instances, the PHPs had a predictable arrangement. They provided a welcome page by the author, mentioning the team the site was dedicated to and the site was split up into sections containing information about the players in the team,

some news about the team, numerous match reports, a club history, some statistics about the team and perhaps information about the football ground the team currently occupies. However, it was also noticed that the PHP is a site with numerous possible guises. Whereas the PHP was said to be the first uniquely digital genre (Crowston and Williams 1997), the waters have now been muddied somewhat with the appearance of the blog (and SN sites) and more detailed attempts at defining the PHP. This, in turn, has had an effect on the characteristics that PHPs are said to have. So far, studies of PHPs have always sought to point out the types of elements which commonly make up a PHP (de Saint Georges 1997, Bates and Lu 1997, Dominick 1999, Dillon and Gushrowski 2000, Papacharissi 2002a). There have been mentions of elements such as the name of the author, email address, favourite websites, gender and photograph (Bates and Lu 1997), or title, email address, update date, table of contents and a creation date (Dillon and Gushrowski 2000).

However, for primarily two reasons, the results of those studies need to be revisited. Firstly, the number of guises of PHPs has changed since 2001, when the latest of the empirical studies was carried out (Papacharissi 2002a, 2002b). With the more recent explosions in the popularity of blogs and SN sites, the figures created by studies discussing characteristics and elements found on PHPs need to be revisited. The PHP articles written in the late 1990s (and slightly later as well), are mostly discussing a specific kind of PHP (the only kind that existed then) and not any of the evolved versions of today (e.g. the blog, the SN site or the photosite). Secondly, at the time of those studies, a commonly accepted definition of the PHP was not available (indeed, even now, a fully accepted definition still seems to be elusive, though with Doring's (2002) attempt not yet challenged, perhaps we are closer to a consensus). However, in the present day, being closer to a generally accepted definition of PHPs (for which it is easily argued that blogs should be included- see 2.3.2 Blogs and Social Networking Sites, p. 32), relevant studies need to change their sampling methods. For instance, the practice of simply taking lists of PHPs from service providers (e.g. Bates and Lu 1997, Dominick 1999, Dillon and Gushrowski 2000, Papacharissi 2002a, 2002b) could be rethought. Though many of these studies put into place some safeguards to protect from the inclusion of certain sites that they did not consider to be PHPs (notably small business sites taking advantage of the free web page services), this does little to encourage the inclusion of all types of PHPs. Such procedures illustrate

the one-sidedness of the commonly used methodologies, leaving out more sophisticated PHPs created by authors who purchased their own domain names, sites which clearly fit the definition of the PHP. Any results of any aspects of PHPs from such studies will have some reflection of this single provider bias. Even where the larger studies have used more than one service provider (e.g. Papacharissi 2002a, 2002b used four service providers), there is still an exclusion of sites that will never find their way into the examined PHPs though they clearly fit the definition. Even studies which are not observational (e.g. Maruyama 1999, Ujigawa 1999, Flanagin and Metzger 2003), take into account the results of studies which are, and inevitably reflect PHPs from certain sources. These characteristics were not specifically the focus of this study, however, the methodology devised here does combat this issue. The examination of PHPs by topic as has been carried out in this thesis (once again something that, to date, has not been attempted anywhere else), though certainly more time-consuming than previous PHP selection methods, certainly increased the probably of ensuring the selection of all types of PHPs that exist, and indeed aids the discovery of new types of PHPs. For studies which need to find PHPs according to a specific field or subject, the methodology used here is suitable.

However, all this is not to say that all studies from now on should be carried out using the ideas presented in this thesis, nor is it to say that discarding the earlier studies is the way forward. The argument is simply that studies concerning PHPs should, where possible take a more all rounded approach, and not disregard the definition of the PHP, and by extension, the vast repertoire of PHPs in existence. The earlier studies conducted on PHPs should perhaps be viewed as studies on specific types of PHPs, rather than PHPs as whole.

7.3.2 Accessibility of PHPs

This then, raises another issue concerning the accessibility of PHPs. It is questionable as to whether identifying PHPs as PHPs is useful for the web surfing information consumer (objective 4). In fact, the evidence in the research carried out for this thesis would suggest otherwise, with several interviewees making the point that they are not interested in whether a site is a PHP or not. However, from the point of view of

studying PHPs as the work of individuals that contribute to the information available to the general public, it would certainly be useful to be able to easily and quickly identify a site as a PHP. To this end however, the attempts at creating PHP search engines have failed, perhaps not conceptually, but in that none of the search engines are still in operation. However, both attempts at creating the search engines disregarded the definition of the PHP, choosing to focus on either academic PHPs (Hoff and Mundhenk 2001) or the name of the author (Shakes, Langheinrich and Etzioni 1997). Here, the reasons for the failure of the search engines is hard to speculate on, however, a lasting solution could well be the inclusion of a classification scheme in the more popular search engines. At the time of writing, none of the major search engines provide either an option to search specifically for PHPs, or an indication that a site in the returned results is a PHP.

In terms of finding a PHP on a specific subject though, for the purposes of collecting PHPs for study, this work devised an original method. Previous studies carried out on PHPs, have not looked at specific subjects. The tendency has, as already mentioned, been to use generic PHPs from service providers (e.g. Dominick 1999, Papacharissi 2002a) or PHPs of specific domains, such as academic (e.g. Thelwall and Harries 2004, Nomura, Ishida and Yokozawa 2001). This is the first work that searches for PHPs on a specific subject area. The process devised for finding the sites and then classifying them as PHPs is, as a result, completely unique to this thesis.

7.3.3 The PHP as an Information Resource

This thesis is itself unique, in that it is the only study to ever look at the contribution of PHPs within the greater context of information seeking in a specific field. This has brought about a more in depth understanding of issues surrounding PHP use and perception, specifically in the subject area of professional football in England. These will be discussed presently.

Negative Perceptions

When investigating the perceptions of football fans concerning PHPs (objective 10), it was clear that the majority of football fans, as mentioned in earlier chapters, had negative views, something that has also been noticed elsewhere (Bates and Lu 1997, Dominick 1999). One of the major causes for these negative perceptions must be the fact that PHPs are not updated as quickly as traditionally popular football websites, something which was discovered whilst investigating objective 11 with the following comments indicative of the interviewees' thoughts:

"People who check sites all the time are looking for current news, not old stuff" Smar-NUFC.

Investigations concerning other objectives also provide evidence pointing in the same direction (e.g. objective 7 and objective 9). In the process of measuring the instances of archival information available on PHPs (objective 7), it became apparent that many PHPs are not updated often enough, if at all. This has also been noted by other studies (Papacharissi 2002a, 2002b) and has an effect on the football fans, who start to move away from the PHP to the larger sites in search for the latest information. Understandably, every football fan has a finite amount of time to browse the web, and with every occasion where the football fan is forced to use a non-PHP, the general popularity of PHPs goes down. Meanwhile, whilst looking at the factors that have an effect on the availability and quality of PHPs (objective 9), numerous points were noted about the supply of information in general concerning professional football in England, one of which was that on many of the popular non-PHPs (e.g. the official site, the BBC website, the SkySports website), the information is provided by relatively large groups of people rather than individuals. Clearly a single individual working on a PHP will not be able to provide the same quantity of information, updated at the same rate.

Indicators from other studies include a mention of the popularity of recent news (Green 1999) and at least one survey shows that fans visit the same site several times a week (FA Premier League Survey 2005). As the PHP represents a single person, it

is only logical that in most cases, it cannot be updated often enough to satisfy this insatiable appetite and this is surely another cause of negative perceptions.

Unique Information

The PHP has often been used as a tool for mass communication, and by extension a tool for information dissemination (Dominick 1999). Several works have hinted at the potential benefits of PHPs (Haines 1999, Weaver 2000, Thelwall and Harries 2004) and looked at instances where they might provide benefits. However, in terms of information provided, only now, for the first time, can we say that not only do PHPs have the potential to provide additional information to fields, but that they actually do provide additional information (i.e. unique information) to certain fields (objective 6 and objective 8). Having used the subject area of professional football as a case study, it was clear from the outset that the majority of the information would be provided by non-PHPs. However, in a subject area with such great numbers of information providers, there were still clear instances where information was available only on PHPs. In the past, it would not have been possible to say this with any certainty, taking into account the fact that no study has compared the content of PHPs to non-PHPs. Most PHP studies have been observational solely of the PHPs (e.g. Bates and Lu 1997, Dominick 1999, Papacharissi 2002a, 2002b), even those studying blogs (Hourihan 2002, Nardi et al 2004b, Herring et al 2004) as well as the occasional study on improving the ease with which the sites are created (Maruyama 1999, Ujigawa 1999, Flanagan and Metzger 2003), resulting in a lack of insight into the informational content of PHPs. Even the studies that were relevant to the PHP as a resource (Haines 1999, Weaver 2000, Thelwall and Harries 2004), did not compare the PHPs to non-PHPs, so could not conclusively make such a statement.

Accurate Information

This benefit is then fortified by the conclusions of the investigations dealing with the accuracy of information (objective 8). Having discovered that the vast majority of the information found on PHPs is accurate (especially at the time of writing) clearly strengthens a previously weak aspect of PHPs. It would have been easy to doubt the

value of information found on PHPs, if one could simply say that the accuracy of the information is doubtful. Indeed, the investigations carried out by this research concerning this aspect of PHPs are not extensive, and a larger scale study should be undertaken before this can be said with certainty. However, the results for this study were so high in positive territory that, without any evidence to the contrary (which is not currently present in academic literature), there can be no reason to discard it. Hence, knowing that PHPs offer information which is unique to the web, and also for the large part accurate, makes a strong case for a higher value to be assigned to PHPs.

Archival Material and Static Information

Something that was discovered during this study (objective 7), which provides even further value to the PHP as an information resource, was that often the site will be out of date (with information about players who are no longer at the club and comments about events that took place in previous seasons). As the PHP is abandoned, the information provided on the pages remains static and this, in a strange way, is a cause for the added value, in a way similar to which grey literature is considered to have value.

Concerning football websites as a whole, there is no guarantee that a site will archive or store its past material. Certain sites keep archives of their material which might or might not be available to view on their pages. In addition, there are accessibility and completeness issues. Some archives are kept on databases, which might not be as accessible as more conventional web pages (as search engines might not index them), thereby notably reducing their accessibility to the general public. In addition, keeping information stored on databases might mean that the original information was reduced for feasibility purposes. Under these circumstances, an abandoned PHP with out of date information could end up being the sole provider of this information on the web. This characteristic was already known about PHPs in general, something shown in other research (Doring 2002, Papacharissi 2002b, Narsesian and Nicholas 2005). However, what has never been mentioned before is the usefulness of having a snapshot of this information, both as it could be the only instance of the information on the web but also for historical purposes.

Furthermore, in the same way as arguments are brought forward for the collection and storage of grey literature, so they could be brought forward here. With the information stored on PHPs accurate and significant, yet their nature ephemeral, the similarities between PHPs and grey literature are obvious.

The Benefits of Grey Literature

The combination of the static archival information and unique information, which is accurate, found on PHPs makes an argument for the increased value of the PHP as an information resource perhaps easier to make. Essentially, in addition to all the benefits of PHPs stated elsewhere, the reasons why the PHP is of value can be said to be the same as the reasons for which grey literature is of value. Earlier in this work (see 2.4.6 PHPs and Grey Literature, p. 43), the relationship of the PHP with grey literature was discussed, pointing out the similarities and the differences. Comparisons of various aspects of the web to grey literature have already been made (Pace 2002, Thompson and Guistini 2006), however, the benefit of PHPs specifically as grey literature was, to date unclear. Now perhaps, having shown that unique information is available on PHPs, it is easier to assert that the benefits of grey literature extend to PHPs. The study of PHPs can now be said to be of value for several grey literature related reasons. Firstly, the material present on these sites is quite diverse, so it is likely that it would often never have been published through the normal channels (Auger 1998). Secondly, even when the material is also to be published through the normal channels, it will be present on the PHPs well in advance of any such publication, something which acts as a catalyst for information availability and all related benefits (Auger 1998). Finally, for certain studies, such as social history, the snapshots provided by PHPs can be used to shed light on issues that official bodies might not be as willing to share (e.g. official government documentation on sensitive political issues), something which could also apply to football.

Unique Viewpoints

Another strength of the PHP must surely lie in aspects and viewpoints that cannot be found in the mainstream media, provided by people who have the inclination to study a subject from a specific viewpoint. An excellent demonstration of this is the HTFC-world website, which shows match reports in a humorous cartoon form. Again, the speed of update of such a site cannot compare with the mainstream media sites as it takes time to come up with the "script" as well as the actual "drawings". However, even so, this is something that is clearly not commonly offered by the mainstream media. There have already been mentions in academic literature that the PHP is an opportunity given to the ordinary citizen to publish information in a way that was not available before the advent of the web (Dominick 1999, Papacharissi 2002a). Though it was previously thought that perhaps they are not making the most of this opportunity (Dominick 1999), the results of this thesis suggest that they are adding their unique viewpoints, thus complementing the mainstream media. Meanwhile, it should be considered that, as this applies to football PHPs, it also probably applies to other fields too. The results in this instance should be extendible to PHPs as a whole, and thus the benefit of PHPs should apply to many fields.

Availability of PHPs

One final interesting finding about the PHP (in the field of football at least) was the number of PHPs available in relation to the success of the football team on the pitch. Theoretically at least, the PHP, as a webpage that can be created by almost anyone (in the West), should give fans of "smaller" teams the ability to create an online presence as notable as those of their larger counterparts. This however, does not seem to be the case. Though having an accurate mapping is difficult, the findings of this thesis (see 5.6.1 Quantity of PHPs, p. 170), suggests that teams with a traditionally larger following (e.g. Manchester United) have a greater number of PHPs dedicated to them. The reason for this is not clear (and, as it was not an objective of this thesis, it was not actively sought), though perhaps the likeliest explanation is that the number of PHPs ultimately reflects the number of fans a team has. There is little related literature available on the subject, though one study (Joinson and Banyard 2002) has shown that

success on the pitch does not affect the online behaviour of football fans (though the study looked at fan behaviour on a day-to-day basis rather than a longer period of time).

In terms of the PHP as an information resource, it was thought possible that with the fewer mainstream websites providing less information on the teams with a lesser following, web authors would come in to fill this informational gap. However, at the time of writing, concerning football, this does not appear to be the case. Extending the application of this finding to other fields is difficult, but at a very basic level, it would appear to suggest that where few people have an interest, it would be harder to find a PHP on the subject.

7.3.4 Football Fans, PHPs and Online Collaborations

During investigations concerning objective 6 (see 5.3 Unique Information on Football PHPs, p. 161) and objective 8 (see 5.5 The Accuracy of the Information on Football PHPs, p. 168), instances were discovered where PHPs provided something different enough, accurate enough, comprehensive enough and regular enough for them to earn some notoriety among football fans. As an information resource which will not necessarily update every day, the PHP can still play its own role, as demonstrated by the Albion Album, HTFC-world and so on. Even the out of date PHPs, which appear at first sight to be worthless bytes in the giant reservoir of the web, have a role to play. Having completed all the research for this thesis, there is no question that PHPs have a contribution to make, and indeed, make a contribution even in a field as saturated (in information terms) as the top tier of professional football in England.

However, this was also the first ever study to investigate the use of PHPs in terms relative to other tools and methods. Having looked at the PHP from this distant, broad view, it is clear that their contribution, though existent, is minimal. In the case study used for this study it is clear that the amount of information present is more often than not, notably less than the information found on non-PHPs concerning the same club. In the end, even though they are still a source of information otherwise unavailable on the web, football PHPs in general are far from popular. Overall, the research

conducted for this thesis (objective 4) determined that in fact, PHPs are used very sparsely compared to the other information providing tools and methods available to football fans. The positive aspect of this though, is that many PHP authors seem to have realised this, and have started working together to provide more comprehensive sites.

Out of the work carried out for objective 12 (see 6.4 Crafting A Communal Site Outline, p. 210), it was found that there are currently various ways in which football fans work together to produce football websites online, including both common collaborations (such as fanzines and forums) and more unusual instances of collaboration. In the instances of the notable unusual collaborations discovered, often more "typical" football site layouts are maintained, yet editorial control is relinquished to various degrees (with contributors in some instances going through automated mechanisms to publish material).

Concerning web authors in particular, they tend to be well motivated individuals, with a strong desire to express themselves, whose primary obstacle (though not necessarily a major one) is time. Web authors are generally not overly concerned about their competition, and though they generally feel collaboration would aid the improvement of their sites, they do not seem to actively seek it. Concerning site popularity, the most important aspect is that the site be up-to-date and where high participation levels are required (though the responses were varied), the majority of web authors made a mention of relevant or interesting material.

All in all, the CCCP has a design which is intended to meet all these criteria, and certainly, the results of this research suggest that such a design offers a maximisation of the benefit derived from the efforts of such web authors.

Objective 12	Conclusion
To craft an outline for a communal site that can encourage web authors to work together in the creation of web pages for the online community.	Put together a design for the CCCP, having examined existing collaborative efforts and through interviews with web authors.

Table 7.3 Objective 12

7.4 Original Contributions

The efforts of this study have yielded five main original contributions in the study of Personal Home Pages, information seeking and needs, discussed in order of priority here. The first, and most significant, of these contributions is a broad PHP contribution which is made of four discoveries concerning PHPs. All the contributions are discussed presently.

Contribution 1 – Attaining an improved understanding of PHPs as unique information providers online.

This first contribution is made of four discoveries concerning PHPs. These include:

- 1 - Discovering that certain Internet-wide content is available only on PHPs.
- 2 - Discovering that a significant proportion of this unique content is of an archival and static nature.
- 3 - Discovering the high level of accuracy of information found on PHPs.
- 4 - Determining the overall role played by the PHP as an information resource within the greater sphere of information seeking.

As already mentioned, this is the only study to ever look at the informational contribution of PHPs. In the process of doing so, it was discovered that there is indeed information on a subject (professional football in England) that can be found online only on PHPs (objective 6). No other academic work has, as far as the author can ascertain, at the time of writing, looked at the PHP in such a manner, having neither attempted to discover nor discovered this truth. In addition, through the realisation that many PHP authors cease to update their sites whilst still keeping them online (whether intentionally or unknowingly), it became apparent that valuable historical snapshots could potentially be provided by PHPs. This indeed proved to be the case, as not only do the outdated PHPs provide valuable insight into a subject in the same way that grey literature does, but often the information found on PHPs is unique simply because it is not recent (objective 7). Furthermore, having determined that the degree of accuracy of information found on PHPs in general is quite high (objective 8), it is quite likely that this unique information (archival, static and otherwise) is

accurate. Finally, concerning the fourth point above, in addition to being the only academic work carried out which looks at PHPs in a specific subject area, this is also the only thesis, as far as the author can ascertain, to examine the PHP within a broader context, looking at the role played by the PHP in relation to all tools and methods used (objective 4). Presently, the PHP does not play an active role in the supplying of football related information and the causes for this have been identified.

Looking at this contribution overall in relation to available relevant literature, work has been carried out on PHPs in terms of the purpose of PHPs (e.g. Buten 1996, Dominick 1999, Papacharissi 2002b) and the informational value of PHPs has been implied (e.g. Haines 1999, Hoff and Mundhenk 2001, Blood 2002, Thelwall and Harries 2004). Football information services have also been looked at (e.g. Green 1999, FA Premier League 2005) and the idea of treating the web (as well as PHPs) as grey literature has been touched on (Pace 2002, Thompson and Guistini 2006). However, all in all, as far as the author can ascertain, research which has provided insight into PHPs concerning the four points (as shown above), has never before been conducted. Clearly, these findings are of value to the field of football PHPs, however, almost certainly, these can also be extended to numerous fields outside football.⁹⁸

Contribution 2 - Devising a new system for the collection and classification of PHPs on a specific subject area.

To date, numerous studies on PHPs have been carried out in various fields. All of these studies have collected PHPs by using set specific domains such as web page service providers or universities (e.g. Buten 1996, Dominick 1999, Papacharissi 2002a, Thelwall and Harries 2004). By carrying out an investigation of PHPs in the field of professional football in England, this thesis devised a system for the collection of PHPs on a specific subject area (see 3.4.1 Finding PHPs, p. 88, and 3.4.3 PHP Classification, p. 90). This system for finding PHPs is completely original when viewed from the viewpoint of other systems used in similar studies (e.g. Buten 1996 etc). The system for the classification of PHPs, uses ideas seen in other works in combination, specifically Smith (1997) and Doring (2002), along with original ideas

⁹⁸ Further research needs to be carried out to identify these fields.

not seen or proposed elsewhere, in particular concerning blogs and photosites. However, this system, with minor modifications, could be used in research in other fields which also study the PHP, such as librarianship, psychology, computer science and journalism, to name but a few. At the time of writing, as far as the author can ascertain, no study has been carried out looking at PHPs in a specific field and consequently, never before have studies had a procedure for the acquisition and classification of PHPs in a specific field of study.

Contribution 3 – Identifying the trend of football website authors moving from PHPs to collaborative efforts.

Though this was a trend that the study was not explicitly seeking to detect, at the time of writing, there has been no mention in relevant literature of PHP authors in such recreational subjects moving towards collaborative efforts. At this point, it must be stated that with the popularity of collective sites, such as Wiki sites or the Open Directory Project (and noting that both of these examples have been around for several years), such a statement could be criticised for merely pointing out something already known. However, this thesis has noted specific (albeit few in number) cases where PHP authors have abandoned their own sites (see Table 6.5) to work together on a larger project. Once again, as far as the author can ascertain, academic research which has explicitly identified such a trend, has not previously been conducted, making this an original contribution.

Contribution 4 – Detecting the need to re-evaluate the results of certain older PHP studies.

As discussed above (see 7.3.1 Characteristics, p. 225), due to the uncertainty concerning the definition of a PHP and the relatively recent increased activity in blogs and SN sites, previous results of studies concerning PHPs need to be viewed in a different light. The conclusion is that studies concerning PHPs should, where possible take a more all rounded approach and take into account the definition of the PHP. The earlier studies conducted on PHPs (e.g. Buten 1996, Dominick 1999, Papacharissi 2002a) should perhaps be viewed as studies on specific types of PHPs, rather than PHPs as whole. This contribution came as a result of the investigation into the

definition of the PHP, and was not set as an objective. However, though studies have taken in-depth looks at the PHP (e.g. Doring 2002, Papacharissi 2002a), such a conclusion has not been mentioned in any related literature at the time of writing.

Contribution 5 - Formulating a skeleton outline for communal sites that can encourage football web authors to work together in the creation of web pages for the online community.

Throughout the investigations of PHPs, it became clear that, though some of them maintain their active status, a great number of PHPs are abandoned. As a result, using the results from the investigations, an attempt was made to create a blueprint for a communal site (the CCCP) that would spread the responsibility and effort required to maintain an active site over as many individuals as possible. In relevant literature, proposing improvements to PHPs (Haines 1999), and providing guidelines as to how PHPs should be created (Ujigawa 1999) have both been discussed. Identifying the PHP as a genre has also been mentioned (Crowston and Williams 1997, Dillon and Gushrowski 2000) and PHPs have even been created for gender research purposes (Arnold and Miller 2000). However, the examination of collaborative football sites, and the drawing of guidelines concerning what is needed to create and maintain an active communal football website have not been mentioned. As a result, it is considered a completely original contribution. Furthermore, it should be noted that such communal sites could also be developed for other communities and groups in other fields. One obvious example from relevant literature is the application such sites could have for those interested in creating communities of practice. Clearly, the CCCP design is tailored to football, however, conceptually, any field where a significant number of individuals are interested in a common subject, websites such as these can be created.

7.5 Future Work

As discussed, this study has produced various results concerning various aspects of the information seeking behaviour of football fans, and in that context, the use of the web and PHPs as information resources. In addition to the main recommendation, a

number of ideas as directions for future work are discussed as additional possible means to maximise the benefits of the efforts of football PHP creators.

The exploratory nature of this study makes it somewhat unsurprising that in this instance, the process of finding answers has led to more questions. Broadly speaking, in this section there are two avenues of future work which lead on naturally from this thesis. The first avenue is concerned with PHP studies on the web, furthering information science themes such as information seeking behaviour, grey literature, web genre studies and virtual community studies. The second avenue concerns related fields that this thesis did not look into directly, but just came across, where the sub-fields of information science involved are primarily information seeking behaviour and virtual communities.

7.5.1 Web and PHP Studies

Concerning PHPs, there are two areas of interest to investigate. The first and perhaps most obvious option, is to carry out a PHP study, using the methodology devised for this thesis, in a separate, unrelated field. Here, only the field of professional football in England has been examined. Studies should be carried out looking at other fields with differing informational qualities so as to paint a fuller picture of where PHPs are at their most effective and also at their least effective. It would be useful to have the same methodology applied as closely as possible in various fields, so as to have exact figures to compare. Such a study would contribute to information science in a way which would be very similar to this study, touching on areas such as information seeking, web genres, grey literature and virtual communities whilst following on from works such as Papacharissi (2002a, 2002b), Dominick (1999), Herring et al (2004), Nardi et al (2004a, 2004b), Narsesian and Nicholas (2005) and Rheingold (2000).

Another aspect of PHP studies (in the domain of web genres) which is in need of examination, is the difference between complete PHPs and incomplete PHPs (i.e. those under construction or abandoned). As already mentioned, there are so many possible differences between PHPs, that it has become difficult to treat them as a single entity. These numerous guises of PHPs lead to a multitude of feasibility issues

that must be overcome with every PHP related project. A clear line separating the broken, under construction and incomplete PHPs from the web would go some way towards ordering a murky domain. Any increased order to this field would undoubtedly help clarify how best to utilise the resourceful aspect of PHPs, furthering works such as Crowston and Williams (1997), Haines (1999), Dillon and Gushrowski (2000) and Doring (2002).

7.5.2 Related Fields

Although the initial focus of this study was on the use, importance and attributes of PHPs in the information seeking behaviour of football fans, the study created an important insight into different aspects of the information seeking behaviour of football fans. Based on the findings of this study, further research could be carried out on the information seeking behaviour of football fans, which could in turn help the enhancement of research in two important areas of information seeking, which are everyday life information seeking and the role of the Internet in information seeking.

Hence, in addition to possible future paths already mentioned, there are two related and potentially interesting avenues for research. These are both borne from “peripheral” discoveries made in this thesis which are not central to the themes discussed here (but are relevant in both information science sub-fields and more loosely related fields). They concern the forums and the official sites. Both of these were mentioned often throughout the interviews, and they would make interesting future work for differing reasons.

Forums, in the context of this study, are interesting for two reasons. Firstly, as an information resource, and secondly, for the effect they have on the field of professional football.

As an information resource, the interest in the forum comes from two sources. Firstly, forums are an altogether different kind of website to conventional sites (including the PHP), and it is difficult to flatly compare them with other information resources. The general setup of forums is different, with the content always changing, with threads

being "born" and others being "killed" (see 4.3.3 The Role Played by the Web, p. 142). As such, users must have an understanding of the setup of the forum to utilise it. However, the nature of the forum is such that a thread often will have just part of the information required (as in essence, a thread is simply a recorded conversation, and will have fitting characteristics). At the same time, according to the fans interviewed in this study, users of forums find the vast majority of the information that they seek on them. In information science terms, such a study would provide additional insight into everyday life information seeking and the role of the Internet in information seeking.

Secondly, concerning the effects of forums on professional football in England, it was the football fan interviewees alone that brought this up as an issue. There were three notable phenomena that occurred in the forum. In the first instance, it was used to coordinate a campaign to help raise funds to build a stadium. Secondly, as a result of conversations taking place on the site, a player decided not to sign for a club. And thirdly, completely fictitious information was fabricated on the forum which found its way into the mainstream media.

The first of these phenomena is perhaps the most predictable one. Forums will give fans a chance to get together and achieve certain goals, such as raising funds for a new stadium or more commonly applying pressure on the club to take certain decisions. As a tool that was not around 20 years earlier, it would be interesting to see how the forum has changed the fan's behaviour.

The second and third of these phenomena were probably less predictable. Both in fabricating information (thus potentially changing history), and dissuading players not to join the club, the forum has an effect on people not associated with it. In this instance, it would be interesting to examine in more detail how forums can affect entities which might have previously been considered unrelated.

In addition to looking at the forums, an interesting path would be to look at the official sites of the football clubs. The popularity of the official sites is clearly quite high. At the same time however, there appear to be numerous complaints concerning various aspects of the sites. On the one hand, when any single object is used often and

by many people, there will invariably be complaints made about this object. On the other hand, football fans are using many sites in addition to the official sites to acquire information about their own club. In this case, it is probable that the official sites are lacking in certain aspects, and it would be interesting to discover in more detail the possible causes and possible cures to this phenomenon. Looking into these aspects of forums and official football websites would be of interest to several sub-fields of social studies (both related and unrelated to football), furthering works such as Fishwick (1989), Green (1999), Giulianotti (1999, 2002) and Ruddock (2005).

7.6 Conclusion

In this study, football fans of the top four professional leagues are shown to be avid users of a range of web resources. Among these, the PHPs are not the most popular, as they are rather static, although they do often provide web-wide unique information. As PHPs are common in fields beyond football, these results are likely to be of wide ranging interest.

Currently, of all the tools and methods available to football fans (which include books, newspapers, TV, mobile phones and even friends), without a doubt, the Internet is the most popular. It appears that a stage has been reached where the web is accessible to the vast majority of fans (of the English leagues at least), and these online facilities offered are the ones they choose to use. PHPs will more often than not, be relatively static in comparison to the more popular sites, such as the BBC and SkySports, and as football fans constantly feel the need for new information, they will rarely flock to a PHP. However, this does not nullify the importance of the PHP as a source of information or the effort that ordinary football fans put into the creation of PHPs. In fact, PHPs often provide unique information not available elsewhere on the web. In addition, the combination of unique and accurate information, along with static PHPs, also gives these PHPs grey literature style attributes, which provide related benefits. Furthermore, PHPs are evolving. Originally there was the “standard” PHP of the early 1990s which was identified as a web genre, and now we have blogs and SN sites. In a similar way, as discovered in this project, PHP authors have started to work in groups, creating numerous forms of collaborative sites. To that end, this

study has provided the blueprint for the CCCP, which will facilitate further the maximisation of benefit derived from the work of these PHP creating fans. Individuals who are creating or who help create such sites are a benefit to Internet users everywhere, and their efforts are notable. This project aimed, broadly speaking, to help increase the attention paid to the PHP and its authors, with the idea that there is useful information on these sites which is not being utilised. This is an idea that has been strongly supported by the findings of this thesis. However, though the case study here looked at football, the implications of the findings of this thesis can almost certainly be extended to many other fields. This is the first study examining PHPs from the point of view of a single subject. However, if, in the future, it can be shown (as expected by the author), that in numerous other fields PHPs are providing valuable information (in any of the ways mentioned in this work), this could potentially be a cause to rethink the way information is sought and provided online. The findings of this thesis are certainly significant in their own right, yet this significance could be redoubled, should future research reinforce the ideas and implications mentioned here.

References and Bibliography

5-Star Support Site (2004). Glossary: B (available online at <http://www.5starsupport.com/glossary/b.htm>).

Adamic L A and Adar E (2003). Friends and Neighbours on the Web. *Social Networks*, 25(3), pp 211-230.

Alberani V and De-Castro P (2001). Grey literature from the York Seminar (UK) of 1978 to the year 2000. *INSPEL* 35 (4): 236-247.

Anderson B (2001). Grey Literature and electronic publishing. *Behavioral and Social Sciences Librarian* 19 (3/4): 57-73.

Apgar M (1998). The Alternative Workplace - Changing where and how people work. *Harvard Business Review*, May-June 1998

Argyle M (1996). The social psychology of leisure. New York. Penguin Books.

Arnold J and Miller H (2000). Gender and Web home pages. *Computers and Education*, 34, 2000, pp 335-359.

Arsenal FC Official Website (2006). Emirates Stadium (available online at <http://www.arsenal.com/thestadium>).

Auger C P (1998). Information sources in grey literature, 4e. London: Bowker, 1998.

Bachmann D and Elfrink J (1996). Tracking the progress of e-mail versus snail-mail. *Marketing Research*, 8 (2), pp.31-35.

Bains J and Patel R (1996). Asians Can't Play Football? Asian Social Development Agency.

Barjak F, Li X and Thelwall M (2006). Which factors explain the Web impact of scientists' personal homepages? *Journal of the American Society for Information Science and Technology*, 58 (2), 200-211.

Barlow J P (1995). Is there a there in cyberspace? *Utne Reader*, March-April pp50-56.

Bates M J (1996). Learning About the Information Seeking of Interdisciplinary Scholars and Students. *Library Trends*, v45 n2, p155-64, Fall 1996.

Bates M J and Lu S (1997). An exploratory profile of Personal Home Pages: Content, Design, Metaphors. *Online and CDROM Review*, 1997, Vol. 21, no. 6.

BBC Website (2004). Top flight soccer boosts economy. 13th April 2004 (available at <http://news.bbc.co.uk/1/hi/england/norfolk/3621571.stm>).

BBC Website (2005a). League agrees improved TV deals. 10th June 2005 (available at <http://news.bbc.co.uk/sport1/hi/football/4079956.stm>).

BBC Website (2005b). Blues fans to scupper Bowyer move. 16th June 2005 (available online at http://news.bbc.co.uk/sport1/hi/football/teams/b/birmingham_city/4092468.stm).

BBC Website (2006a). Setanta joins Premiership action. 5th May 2006 (available at <http://news.bbc.co.uk/1/hi/business/4975632.stm>).

BBC Website (2006b). Dyke becomes Brentford chairman. 20th Jan 2006 (available at <http://news.bbc.co.uk/sport1/hi/football/teams/b/brentford/4627114.stm>).

BBC Website (2006c). Keane triggers city tourism boom. 9th October 2006 (available at <http://news.bbc.co.uk/1/hi/england/wear/6032905.stm>).

BBC Website (2006d). Premier clubs enjoy summer spree. 1st September 2006 (available at <http://news.bbc.co.uk/1/hi/business/5305326.stm>).

BBC Website (2006e). Footballer Cole sues newspapers. 3rd March 2006 (available at <http://news.bbc.co.uk/1/hi/uk/4769776.stm>).

BBC Website (2006f). Blackmailers target \$1m website. 18th January 2006 (available at <http://news.bbc.co.uk/2/hi/technology/4621158.stm>).

BBC Website (2007). Foreign cash 'can boost Arsenal'. 31st August 2007 (available at <http://news.bbc.co.uk/1/hi/business/6971124.stm>).

Belkin N J, Marchetti P G and Cool C (1993). BRAQUE: Design of an Interface to Support User Interaction in Information Retrieval. *Information Processing and Management* 29 (3):325-344.

Bishop I D (1997). Testing perceived landscape color-difference using the Internet. *Landscape and Urban Planning*, 37 (3-4), pp.187-196.

Blood R (2002). *The Weblog Handbook: Practical Advice on Creating and Maintaining Your Blog*. Cambridge MA. Perseus Publishing

Boyd D M and Ellison N B (2008). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication* 13 (1) 210–230.

Bradbury S (2001). *The New Football Communities: A Survey of Professional Football Clubs on Issues of Community, Ethnicity and Social Inclusion*. Sir Norman Chester Centre for Football Research.

Bradford N and Reeve W (1997). *Net Profits*. Fletcher Research, London.

Brentford FC Official Website (2006). Club History section (available online at <http://www.brentfordfc.premiumtv.co.uk/page/ClubHistory/0,,10421,00.htm>).

Broadband TV News Website (2007). Landlady loses Sky pub appeal. 21st December 2007 (available at <http://www.broadbandtvnews.com/?p=3103>).

Brown C M (1999). Information seeking behavior of scientists in the electronic information age: Astronomers, chemists, mathematicians, and physicists. *Journal of the American Society for Information Science*, 50 (10) 929 – 943.

BSkyB Website (2006). Press Release. British Sky Broadcasting Group PLC. 5th May 2006 (available at <http://phx.corporate-ir.net/phoenix.zhtml?c=104016&p=irol-newsArticle&ID=852973&highlight=>).

Buchanan T (1985). Commitment and leisure behavior: A theoretical perspective *Leisure Sciences*, 7, 401-420.

Bulls News Website (2007). Turner lifts Ban on Watson. 6th September 2007 (available at <http://bullsnews.blogspot.com/search?q=loveibull>).

Buten S (1996). First World Wide Web Personal Home Page Survey (available at <http://www.asc.upenn.edu/usr/sbuten/survey1.htm>).

Calantone R J, Gudmundson D, Lang J R (1997). Small Firm Information Seeking as a Response to Environmental Threats and Opportunities. *Journal of Small Business Management*, Vol. 35.

Chandler D (1998). Personal Home Pages and the Construction of Identities on the Web (available at <http://www.aber.ac.uk/media/Documents/short/webident.html>).

Chaudhary V (2003). Forty factors fuelling football inflation. *The Guardian*, Thursday July 31, 2003 (available online at <http://football.guardian.co.uk/comment/story/0,9753,1009392,00.html>).

Choo C W, Detlor B and Turnbull D (1999). Information Seeking on the Web: An Integrated Model of Browsing and Searching. *ASIS Annual Meeting* (available online at <http://www.ischool.utexas.edu/~donturn/papers/asis99/asis99.html>).

Collins M, Henry I, Houlihan B and Buller J (1999). Sport and Social Inclusion: A Report to the Department for Culture, Media and Sport. Loughborough University: Institute of Sport and Leisure Policy.

Constant D, Sproull L and Kiesler S (1996). The Kindness of Strangers: The Usefulness of Electronic Weak Ties for Technical Advice. *Organization Science* 7 (2).

Cox A M, Clough P H and Marlow J (2008). Flickr: a first look at user behaviour in the context of photography as serious leisure. *Information Research*, 13 (1) (available online at <http://informationr.net/ir/13-1/paper336.html>).

Craig Murray Website (2007). Alisher Usmanov, potential Arsenal chairman, is a Vicious Thug, Criminal, Racketeer, Heroin Trafficker and Accused Rapist. 2nd September 2007 (available at http://www.craigmurray.org.uk/archives/2007/09/alisher_usmanov.html).

Croon A and Argen P O (1998). Fyra former av virtuella gemenskaper. Human IT (available online at <http://www.hb.se/bhs/ith/2-98/acpoaa.htm>)

Crowston K and Williams M (1997). Reproduced and emergent genres of communication on the World Wide Web. Draft of paper presented at the Hawaii International Conference on Systems Science, January (available at <http://crowston.syr.edu/papers/webgenres.html>).

CWIS Style Guidelines (2002). Murdoch University, Perth, Western Australia. CWIS Style Guidelines (available at <http://www.murdoch.edu.au/cwisad/styles/copyrite.html>).

Davenport T (1996). "Knowledge management at Hewlett-Packard", Centre for Business Innovation, Ernst & Young LLP.

de Saint-Georges I (1997). Click here if you want to know who I am. Deixis in Personal Homepages. Proceedings of the Hawaii International Conference on Systems Science, January 6-9, 1997.

Delaware Online (2007). What is a personal homepage? (available online at <http://web.dol.net/noframes.asp?t=content&f=helphwhat.html>)

Deloitte and Touche (2001). The Football Rich List. Deloitte and Touche, December 2001 (available online at <http://www.footballfinance.co.uk/publications/richlist.pdf>).

Deloitte and Touche (2006). The Football Money League. Sports Business Group, Deloitte and Touche, Manchester February 2006.

Dillon A and Gushrowski B (2000). Genres and the Web: Is the personal home page the first uniquely digital genre? Draft of paper published in the Journal of the American Society for Information Science, 51,2, pp202-205 (available at <http://www.gslis.utexas.edu/~adillon/publications/genres.html>).

Dominick J R (1999). Who Do You Think You Are? Personal Home Pages and Self-Presentation on the World Wide Web. Journalism Quarterly, 76 #4, Winter 1999, pp 646-658.

Döring N (2002). Personal Home Pages on the Web: A Review of Research. Journal Computer Mediated Communications, 7 (3) April.

Dunning E G, Murphy P, Williams J, Maguire J (1984). Football Hooliganism in Britain Before the First World War. International Review for the Sociology of Sport, Vol. 19, No. 3-4.

Ellis D (1989). A Behavioural Approach to Information Retrieval Systems Design. Journal of Documentation 45 (3).

Ellis D (1991). New horizons in information retrieval. Library Association, London

Ellis D (1996). Progress and Problems in Information Retrieval. Library Association, London.

Ellis D (1997). Modelling the information seeking patterns of engineers and research scientists in an industrial environment. *Journal of Documentation* 53 (4):384-403.

Ellis K (2001). "Sharing the best practices globally", *Training*, July, pp.32-8.

England and Wales High Court (2007a). Sheffield Wednesday Football Club Ltd & Ors v Hargreaves. EWHC 2375 (Qb). 18th October.

England and Wales High Court (2007b). Murphy v Media Protection Services Ltd. EWHC 3091 (Admin). 21st December.

England and Wales High Court (2008a). Football Association Premier League Ltd & Ors v QC Leisure & Ors. EWHC 1411 (Ch). 24th June.

England and Wales High Court (2008b). Murphy v Media Protection Services Ltd. EWHC 1666 (Admin). 16th July.

FA Premier League (2005). National Fan Survey (available online at <http://www.premierleague.com/fapl.rac?command=forwardOnly&nextPage=enPublications&offset=10>)

Farace D J (2000). The Grey Link in the Information Supply and Demand Chain. In *National Online Meeting proceedings*, edited by Martha E. Williams. Medford, N.J.: Information Today.

Fishwick N (1989). English football and society, 1910-1950. Manchester University Press, Manchester.

Flanagin A J and Metzger M J (2003). The perceived credibility of personal Web page information as influenced by the sex of the source. *Computers in Human Behavior*, 19(6), 683-701.

Football Association Official Website (2006). History of the FA Cup (available at http://www.thefa.com/TheFACup/TheFACup/History/Postings/2006/01/FACup_History.htm).

Football DataCo Ltd Email (2008). Football league rights (inc Premier League, Football League, Scottish Premier League and Scottish Football League). Received on 10th October 2008.

Football League (2006). Supporter Survey (available online at <http://www.football-league.premiumtv.co.uk/staticFiles/7f/5f/0,,10794~24447,00.pdf>).

Football League Official Website (2006). History of the Football League (available at <http://www.football-league.premiumtv.co.uk/page/History/0,,10794,00.html>).

Forbes.com (2007). #142 Alisher Usmanov. 8th March 2007 (available at http://www.forbes.com/lists/2007/10/07billionaires_Alisher-Usmanov_GIPI.html).

Friendster Site (2006). About Friendster (available online at <http://www.friendster.com/info/index.php?statpos=footer>).

Giulianotti R (1999). Football : a sociology of the global game. Polity Press, Cambridge.

Giulianotti R (2002). Supporters, Followers, Fans, and Flaneurs: A Taxonomy of Spectator Identities in Football. *Journal of Sport and Social Issues* 26; 25 (available online at <http://jss.sagepub.com/cgi/content/abstract/26/1/25>).

Giulianotti R and Williams J (eds) (1994). Game Without Frontiers: Football, Modernity and Identity. Avebury Press, 1994.

Granello D H and Wheaton J E (2004). Online data collection: Strategies for research. *Journal of Counseling & Development*, 82, 387-393.

Green R (1999). Football information services: fanzines, "Match of the Day" and the modem. *Aslib Proceedings: new information perspectives*, Volume 51, Number 1, 1999, pp. 20-29(10).

Gunter B et al. (2002). Online versus offline research: implications for evaluating digital media. *Aslib Proceedings*, 54 (4), pp.229-239.

Haimila S (2001). "Shell creates communities of practice", *KM World* February 19, pp.1-2.

Haines A (1999). Librarians' Personal Web Pages: An Analysis. *College & Research Libraries* pp543-550. November 1999.

Hall H and Smith N (1997). You'll wish it was all over: the bibliographic control of grey literature with reference to print football fanzines. *Serials*, Volume 10, Number 2, July 1997, pp189-194.

Hannabuss S (1995). Approaches to research, *Aslib Proceedings*, 47 (1), pp.3-11.

Harding D (2003). Getting a financial foothold. *Accountancy Age*, 7th August (available at <http://www.accountancyage.com/Features/1134432>).

Harris R and Dewdney P (1994). Barriers to information: How formal help systems fail battered women. Greenwood Press, London.

Hartel J (2003). The serious leisure frontier in library and information science: Hobby domains. *Knowledge Organization*, 30(3/4). 228-238.

Hartel J (2006). Information activities and resources in an episode of gourmet cooking. *Information Research*, 12(1) paper 281 (available at <http://InformationR.net/ir/12-1/paper282.html>).

Haynes R (1995). The Football Imagination: The Rise of Football Fanzine Culture. Arena, Aldershot.

Herring S C, Scheidt L A, Bonus S and Wright E (2004). "Bridging the Gap: A Genre Analysis of Weblogs," hicss, p. 40101b, Proceedings of the 37th Annual Hawaii International Conference on System Sciences (HICSS'04) - Track 4.

Hoff G and Mundhenk M (2001). Finding scientific papers with HomePageSearch and MOPS. pp201-207 Proceedings of SIGDOC 2001 (The Nineteenth Annual International Conference of Computer Documentation, Communicating in the New Millennium), October 21-24, 2001, Santa Fe, New Mexico, USA.

Holt M, Michie J, Oughton C, Shailer L and Walters G (2005). The State of the Game: The Corporate Governance of Football Clubs. Football Governance Research Centre, Birkbeck, University of London.

Hourihan M (2002). What We're Doing When We Blog. O'Reilly Web DevCenter (available at <http://www.oreillynet.com/pub/a/javascript/2002/06/13/megnut.html>)

Huddersfield Town Supporters Trust (2005). News (available online at <http://www.htst.info/news>).

Huntington P, Williams P and Nicholas D (2002). 'Age and gender user differences of a touchscreen kiosk: a case study of kiosk transaction log files'. Informatics in Primary Care 10(1), 2002, pp3-9

Internet Football Ground Guide Site (2006). Football Ground Guide (available online at <http://www.footballgroundguide.co.uk>).

Instone K (2002). Usability Heuristics for the Web (available online at <http://nt.media.hku.hk/webcourse/references/HEURISTICS.htm>).

IRON (2003). The Personal Page of IRON (available at <http://www.geocities.com/ironman070277/myfirstpersonalhomepage.html>).

Jaquiss K (2003). Model Rules for a Football Community Mutual Football Governance Research Centre, Birkbeck, University of London.

Jeffery KG (2000). An architecture for grey literature in a R&D context. International Journal on Grey Literature, Volume 1, Number 2, 2000 , pp. 64-72(9)

Joinson A and Banyard P (2002). Psychological aspects of information seeking on the Internet. Aslib Proceedings, 54(2): 95-102.

Kiesler S and Sproull L (1986). Response effects in the electronic survey. Public Opinion Quarterly, 50, pp.402-413.

Klang M, Olsson S (1999). Commercializing Online Communities: From Communities to Commerce. In: Proceedings of the 2nd International IeC99. Manchester, 1999

Kling R (1996). Social Relationships in Electronic Forums: Hangouts, Salons, Workplaces and Communities. Computerization and Controversy: Value Conflicts and Social Choices (2nd Ed.) by Rob Kling. San Diego, CA: Academic Press, 1996.

Koch T (1993). Traugott Koch's Personal Homepage, created 1993 (<http://www.lub.lu.se/netlab/staff/koch.html>).

Koehler W (2002). Web page change and persistence---a four-year longitudinal study. Journal of the American Society for Information Science and Technology, v.53 n.2, p.162-171, January 15, 2002.

Kollock P (1999). The Economies of Online Cooperation: Gifts and Public Goods in Cyberspace. In: Marc Smith and Peter Kollock (editors). 1999. Communities in Cyberspace. London: Routledge

Konami.com Site (2006). News (available online at http://www.konami.com/edealinv/servlet/ExecMacro?url=control/News.vm&ctl_nbr=3810).

Kuhlthau C C (1991). Inside the Search Process: Information Seeking from the User's Perspective. Journal of the American Society of Information Science 42 (5):361-371. Belkin.

Lave J and Wenger E (1991). Situated Learning. Cambridge University Press, Cambridge, UK.

Lawrence S and Giles C L (1999). Accessibility of information on the web. Nature, vol 400, 8th July, pp107-109.

Li D (2005). Why do you blog: A uses-and-gratifications inquiry into bloggers' motivations. MA Thesis, Marquette University, Milwaukee, Wisconsin.

Manchester United Official Website (2006a). Club (available online at <http://www.manutd.com>).

Manchester United Official Website (2006b). Old Trafford (available online at <http://www.manutd.com>).

Marcella R and Baxter G (1999). The information needs and the information seeking behaviour of a national sample of the population in the United Kingdom, with special reference to needs related to citizenship. Journal of Documentation, Vol. 55, No. 2, March 1999.

Maruyama G (1999). The Web Crusade Part 1: Motives for Creating Personal Web Pages by Individuals. pp 774-777 Proceedings of HCI International '99 (the 8th International Conference on Human-Computer Interaction), Munich, Germany, August 22-26, 1999, Volume 1.

Maurer D (2006). Four Modes of Seeking Information and How to Design for Them. Boxes And Arrows. March 2006 Issue (available online at http://www.boxesandarrows.com/view/four_modes_of_seeking_information_and_how_to_design_for_them).

McKimmie T and Szurmak J (2002). Beyond grey literature: how grey questions can drive research. *Journal of Agricultural and Food Information* 4(2):71-79.

Mehta R and Sivadas E (1995). Comparing response rates and response content in mail versus electronic mail surveys. *Journal of the Market Research Society*, 37 (4), pp.429-439.

Michie J and Oughton C (2004). Competitive Balance in Football: Trends and Effects. Football Governance Research Centre, Birkbeck, University of London.

Middleton I, McConnel M and Davidson G (1999). Presenting a model for the structure and content of a university World Wide Web site, *Journal of Information Science*, 25(3), pp 219-227.

Morning Advertiser (2007). Host cleared in Sky football case. 7th November 2007 (available at http://www.morningadvertiser.co.uk/news.ma/article/53188?Ns=PublicationDate%7C0&N=598291&PagingData=Po_0~Ps_200~Psd_Asc).

Morning Advertiser (2008a). Two TV football cases put on hold. 25th September 2008 (online at <http://www.morningadvertiser.co.uk/news.ma/article/80791?Ns=ArticleTitle%7C0&N=598291&Ne=598327>).

Morning Advertiser (2008b). Sky can intervene in landmark case. 14th November 2008 (available online at <http://www.morningadvertiser.co.uk/news.ma/article/81258?Ns=ArticleTitle%7C0&N=598291&Ne=598327>).

Morris M and Ogan C (1996). The internet as mass medium. *Journal of Communication* 46 (1), pp39-50.

MySpace Site (2006). About Us (available online at <http://www.myspace.com/Modules/Common/Pages/AboutUs.aspx>).

Nardi B, Schiano D and Gumbrecht M (2004a). Blogging as social activity or would you let 900 million people read your diary? *Proceedings of Computer Supported Cooperative Work 2004*.

Nardi B, Schiano D J, Gumbrecht M, Swartz L (2004b). Why we blog. *Communications of the ACM*, v.47 n.12, December 2004.

Narsesian S (2004). Personal home pages as an information resource. *Webology*, 1(2), Article 5 (available at <http://www.webology.ir/2004/v1n2/a5.html>).

Narsesian S and Nicholas D (2005). Personal Home Pages as a resource: Evaluating the informational value of the PHP on the web. *Aslib proceedings 2005* 57 (1)63-84.

National Statistics Omnibus Survey (2005a). Office Adults who have ever used the Internet by sex/age (Great Britain) (available at <http://www.statistics.gov.uk/StatBase/Expodata/Spreadsheets/D6928.xls>)

National Statistics Omnibus Survey (2005b). Internet Access (available at <http://www.statistics.gov.uk/CCI/nugget.asp?ID=8&Pos=&ColRank=1&Rank=374>).

Newcastle United Official Site (2006). Stadium Information (available online at <http://www.nufc.premiumtv.co.uk/page/Stadium/0,,10278,00.html>).

NexBand Communications (2006). What is a personal homepage? (available online at http://nexband.com/faq_answers.php?question_number=6)

Nicholas D (2000). Assessing Information Needs: tools, techniques and concepts for the Internet age. Aslib, London.

Nicholas D, Dobrowolski T, Withey R, Russell C, Huntington P and Williams P (2003). Digital information consumers, players and purchasers: information seeking behaviour in the new digital interactive environment. Aslib Proceedings Volume 55 Number 1/2 2003 pp. 23-31.

Nielsen J (1994). Heuristic evaluation (1994) in Nielsen, J. and Mack, R.L. (Eds), Usability Inspection Methods, John Wiley & Sons, New York, NY.

Nomensa Site (2006). New online community for football lovers, 20 March 2006. (available online at <http://www.nomensa.com/news/industry-news/2006/3/new-online-community-for-football-lovers.html>).

Nomura S, Ishida T and Yokozawa M (2001). International Comparative Study of Identity Presented on the Internet. 4th International Conference on Computational Intelligence and Multimedia Applications (ICCIMA01), October-November 2001.

Ocholla D N (1996). Information-seeking Behaviour by Academics: A Preliminary Study. The International Information & Library Review, Volume 28, Number 4, December 1996, pp. 345-358(14).

Oliver K, Wilkinson G and Bennet L (1997). Evaluating the Quality of Internet Information Sources. Annual Convention of the Association for the Advancement of Computing in Education (AACE) Calgary, AB, Canada.

Online Computer Library Center (OCLC) (2002). Web Characterisation Project (available online at <http://wcp.oclc.org>).

Oppenheim A N (1992), Questionnaire design, interviewing and attitude measurement. Pinter Publishers, London.

Orton R, Marcella R and Baxter G (2000). Information seeking behaviour of Members of Parliament in the United Kingdom. Aslib Proceedings Vol 52, No.6, June 2000 – 207.

Oughton C, McClean M, Mills C, Hunt P (2003). Back Home - Returning Football Clubs to their Communities. *Mutuo*.

Pace A K (2002). Black, white, and shades of gray (Literature) on the Web. *Computers in Libraries* 22 (4): 44-47.

Papacharissi Z (2002a). The presentation of self in virtual life: Characteristics of personal home pages. *Journalism and Mass Communication Quarterly* 79(3), 643-660.

Papacharissi Z (2002b). The Self Online: The Utility of Personal Home Pages. *Journal of Broadcasting and Electronic Media* 46(3), 2002, pp 346-368.

Paquet S (2003). Personal knowledge publishing and its uses in research. *Knowledge Board*, 10 (available online at <http://www.knowledgeboard.com/item/253>).

Parker, L (1992). Collecting data the e-mail way. *Training and Development*, July, pp.52-54.

PEW / Internet and American Life Project (2005). Internet Activities. (http://www.pewinternet.org/trends/Internet_Activities_8.05.05.htm).

Pickavet C (2006). Balancing The Business of MySpace. *Internet News* (available online at <http://www.internetnews.com/ec-news/article.php/3582941>)

Premier League Press Release (2007). National Music Publishers' Association, Robert Tur, sports bodies, others to join copyright action against YouTube/Google; interim class counsel appointed. 6th August 2007.

Punch K F (1998). Introduction to Social Research: Quantitative and Qualitative Approach. Sage Publications, London.

Ramos K, Linscheld R and Schafer S (2003). Real-time information-seeking behavior of residency physicians. *Family Medicine*, 35(4), 257-260.

Rheingold H (1992). A Slice of Life in My Virtual Community. *Editor, Whole Earth Review* 27, June 1992, Gate Five Road Sausalito, CA 94965 (available online at http://www.nectec.or.th/net-guide/bigdummy/bdg_255.html).

Rheingold H (2000). The Virtual Community: Homesteading on the Electronic Frontier. MIT Press. London.

Rick, J (2007) AniAniWeb: a wiki approach to personal home pages. *International Symposium on Wikis, Proceedings of the 2007 international symposium on Wikis*, pp. 99 – 118.

Rieh S Y (2004). On the Web at Home: Information Seeking and Web Searching in the Home Environment. *Journal of the American Society for Information Science and Technology*, 55(8):743–753.

Riley D (2005). Blog Count for July, 70 million blogs. Blog Herald (available online at <http://www.blogherald.com/2005/07/19/blog-count-for-july-70-million-blogs/>)

Rochester Institute of Technology (2007). Opening up your RIT Computer Account for Web Access (available online at http://www.rit.edu/its/services/personal_web_pages/personal-home-pg.html)

Rubio S (1996). Home Page. Bad Subjects. Issue #24, February 1996 (available at <http://eserver.org/bs/24/rubio.html>).

Ruddock A (2005). Let's Kick Racism Out of Football— and the Lefties Too!: Responses to Lee Bowyer on a West Ham. *Journal of Sport and Social Issues* 29; 369 (available online at <http://jss.sagepub.com/cgi/content/abstract/29/4/369>).

Safran C and Kappe F (2008). Success Factors in weblog community. *Journal of Universal Computer Science*, 14 (4), 546-556 (available online at http://www.jucs.org/jucs_14_9/success_factors_in_a/jucs_14_4_0546_0556_safran.html)

Schaefer D R and Dillman D A (1998). Development of standard e-mail methodology: Results of an experiment. *Public Opinion Quarterly*, 62 (3), pp.378-397.

Schmidt W C (1997). World-Wide Web survey research: Benefits, potential problems, and solutions. *Behavior Research Methods, Instruments and Computers*, 29, pp.274-279.

Schreiber T and Moring C (1997). The communicative and organisational competencies of the librarian in networked learning support: a comparative analysis of the roles of the facilitator and the librarian. *Information Research*, 3(1) (available online at <http://informationr.net/ir/3-1/paper29.html>).

Shakes J, Langheinrich M, and Etzioni O (1997). Dynamic reference sifting: A case study in the homepage domain. In *Proceedings of the 6th International World Wide Web Conference*, pages 189-200, 1997.

Shepherd M and Watters C (2004). *Proceedings of the WWW2004 Conference Workshop*. New York.

Sifry D (2006). State of the Blogosphere, April 2006 Part 1: On Blogosphere Growth (available online at <http://technorati.com/weblog/2006/04/96.html>).

SkySports.com Website (2004). More Live Football Than Ever Before!, SkySports.com, Monday 26th January 2004, (available at <http://www.skysports.com/skysports/article/0,,123-1100512,00.html>).

Smith A G (1997). Testing the Surf: Criteria for Evaluating Internet Information Resources, *The Public-Access Computer Systems Review* 8, no.3 (available online at <http://info.lib.uh.edu/pr/v8/n3/smit8n3.html>).

Soccer-stats.com Site (2006). Attendances (available online at <http://www.soccer-stats.com/divisions/attendances.asp?>).

Sonnenwald D H, Wildemuth A and Harmon G L (2001). A research method to investigate information seeing using the concept of information horizons: an example from a study of lower socio-economic students' information seeking behaviour. *The New Review of Information Behaviour Research*, 2.

Spink A and Ozmutlu HC (2001). Sexually-related information seeking on the web. *Proceedings of the Annual Meeting of the American Society for Information Science and Technology*, 545-554.

Sproull L S (1986). Using electronic mail for data collection in organizational research. *Academy of Management Journal*, 29, pp.159-169.

Stebbins R A (1982). Serious leisure: A conceptual statement. *Pacific Sociological Review*, 25, 251-272.

Stebbins R A and Hartel J (2006). The Serious Leisure Perspective (available at <http://www.soci.ucalgary.ca/seriousleisure>).

Sunday Mirror/NPower (2001). Fans Survey (available online at <http://bisa.afcb.co.uk/afcbisanews/detail.asp?story=45>).

Tacon R (2005). Football and Social Inclusion: Evaluating Social Policy. Football Governance Research Centre, Birkbeck, University of London.

Tashakkori A and Teddlie C (1998). Mixed methodology. Sage, London.

Tenopir C, King D and Bush A (2004). Medical faculty's use of print and electronic journals: changes over time and in comparison with scientists. *Journal of the Medical Library Association*, 92(2), 233.

The Football Task Force (1999). Investing in the community, a submission by the Football Task Force to the Minister for Sport. January 11th 1999. London.

The Independent (2005). Football club chairman wins £250,000 libel case against 'Times' columnist. 27th October 2005 (available at <http://www.independent.co.uk/news/uk/crime/football-club-chairman-wins-163250000-libel-case-against-times-columnist-512750.html>).

The Observer (2008). Big win for Moyes in Rooney libel row. 23rd March 2008 (available at <http://www.guardian.co.uk/sport/2008/mar/23/football.newsstory>).

The Register Website (2008). Premiership tackles Ukrainian football streaming site. 27th June 2008 (at http://www.theregister.co.uk/2008/06/27/premiership_domain_wipo/).

The Sheffield Telegraph (1935). 29 April 1935.

The Swindon Evening Advertiser (1929). 29 January 1929.

The Telegraph Online (2007). Premier League to take action against YouTube. 23rd May 2007 (available at <http://www.telegraph.co.uk/sport/football/2312532/Premier-League-to-take-action-against-YouTube.html>).

Thelwall M (2002). Extracting Macroscopic Information from Web Links. *Journal of the American Society of Information Science and Technology*, vol 52, no 13, pp 1157-1168.

Thelwall M and Harries G (2004). Can personal web pages that link to universities yield information about the wider dissemination of research? *Journal of Information Science* 30 (3) pp 240 – 253.

Thompson D and Guistini D (2006). Finding the Hard to Finds: Searching for Grey Literature. CHLA/ABSC. (available online at weblogs.elearning.ubc.ca/googlescholar/Finding%20the%20Hard-to-finds_May10.ppt)

Tse A (1998). Comparing the response rate, response speed and response quality of two methods of sending questionnaires: E-mail vs. mail. *Journal of the Market Research Society*, 40 (4), pp.353-361.

Ujigawa M (1999). The Web Crusade Part 2: Creating a Web Page Design Map. pp 778-782 *Proceedings of HCI International '99 (the 8th International Conference on Human-Computer Interaction), Munich, Germany, August 22-26, 1999, Volume 1.*

Vondracek R (2007). Comfort and Convenience? Why Students Choose Alternatives to the Library. *Portal-Libraries and the Academy*, 7(3), 277-294.

Waghorn R, Downer J and Munby R (2005). FA Premier League National Fan Survey 2004/2005 (available online at <http://www.premierleague.com/en/files/publications/>).

Walvin J (1975). *The people's game : a social history of British football.* Allen Lane, London.

Washington Post (2007). Viacom Sues YouTube Over Copyright. By Frank Ahrens, Washington Post Staff Writer, 14th March. Page D02 (available online at <http://www.washingtonpost.com/wp-dyn/content/article/2007/03/13/AR2007031300595.html>).

Weaver A E (2000). Personal web pages as professional activities: an exploratory study. *Reference Services Review*, 28(2), 171-177.

Weber R P (1990). *Basic Content Analysis.* 2nd ed., Newbury Park, CA: Sage.

Weintraub I (2000). The role of grey literature in the sciences (available online at <http://library.brooklyn.cuny.edu/access/greyliter.htm>)

Wellman B and Gulia M (1999). Net Surfers Don't Ride Alone: Virtual Communities as Communities. Networks in the Global Village. Boulder, CO: Westview.

Wenger E (1998). Communities of Practice. Cambridge University Press, Cambridge, UK.

Whatis.com Site (2004). Glossary, Blog (available online at http://whatis.techtarget.com/definition/0,289893,sid9_gci1186975,00.html#b).

When Saturday Comes (2007). Gagged to order. November 2007 (available online at <http://www.wsc.co.uk/content/view/290/29/>).

Whitmire E (2004). The relationship between undergraduates' epistemological beliefs, reflective judgment, and their information-seeking behavior. Information Processing and Management, 40(1), 97-111.

Williams J (1988). England's Barmy Army. Social Studies Review, Vol. 4, No. 1, 1988.

Williams J and Wagg S (1991). British football and social change: getting into Europe. Leicester University Press, Leicester.

Wilkinson M A (2001). Information sources used by lawyers in problem solving - An empirical exploration. Library and Information Science Research, Volume 23, Number 3, Autumn 2001, pp. 257-276(20).

Wilson T D (1997). Information behaviour: an interdisciplinary perspective Information Processing and Management, 33, 551-572.

Wilson T D (1999). Models in information behaviour research. Journal of Documentation, 55(3) 249-270 (available online at <http://informationr.net/tdw/publ/papers/1999JDoc.html>).

Windows Live Spaces Site (2006). About Spaces (available online at <http://spaces.live.com/>).

Wired Magazine (2000). 1.6 Billion Served: The Web According to Google. Wired, December 2000, 118-119.

Wolverhampton Wanderer Official Site (2005). History – Club Honours (available at <http://www.wolves.premiumtv.co.uk/page/HistoryDetail/0,,10307~60471,00.html>).

Wood D N (1982). Grey literature - the role of the British Library Lending Division. Aslib Proceedings 34 (11/12) 459-465

Wynn E and Katz J E (1997). Hyperbole over cyberspace: Self-presentation and social boundaries in Internet home pages and discourse. The Information Society, 13(4), 297-328. (available online at <http://www.slis.indiana.edu/TIS/articles/hyperbole.html>)

Yakel E (2004). Seeking information, seeking connections, seeking meaning: Genealogists and family historians. *Information Research*, 10(1) paper 205 (available at <http://InformationR.net/ir/10-1/paper205.html>).