BTECH 451 FINAL REPORT

iSpyHorses.com

Author:
Zhang Luo
5917643
zluo619

Academic Supervisor:
Dr. Ulrich Speidel
Course Coordinator:
Dr. S. Manoharan

October 2015

ABSTRACT

This report describes all the details related to my BTECH 451 project. The goal of this project is to build an online community and a listing platform for international equestrians. All programming work are based on PHP and HTML language. Apart from coding, also included are web server and hosting. The main topics I focused on for this project are web optimisation and web security.

The first chapter introduces the details of this project, about the requirements and what am I going to achieve. There is also a section for the key people who are involved this project.

Chapter two gives a brief outline for all related knowledge for this project and is separated into two parts. The first part introduces the tools and applications that is need for website hosting while the second part briefly describes the background of programming language.

Chapter three then provides the establishment of the system structure and an explanation of how the website processing. Some code is given in this section to help clarify the programming.

The following chapter, chapter four, looks at the main work I have done for this project.

And chapter five and six discusses the main topic of website optimisation and website security.

The last chapter concludes this report introduces difficulties that were encountered. All requirements have achieved for ispyhorses.com, the next step is to study more about security in the web develop.

TABLE OF CONTENTS

CHAPTER 1

PROJECT INTRODUCTION

This project is undertaken for completing the last year of the Bachelor of Technology at the University of Auckland. BTECH 451 is intended for students to work on projects which are related to Information Technology. Normally it requires students to have real work experience with programming. This one-year-long project is worth 45 points of study which equals to three courses at the university.

This project is involves the development of iSpyHorses.com and is supervised by Ulrich Speidel. All programming work are based on PHP, HTML, CSS, and JavaScript. In addition, the topics of web performance and web security to optimise the development are addressed.

1.1 BACKGROUND OF ISPYHORSES.COM

iSpyHorses.com is designed for the community of international equestrians. It works like an online-shopping system as well as a social networking service. There are four main parts of the website. The "I" in the name of iSpyHorses is the homophone of the eye.

- A listing platform allows sales where users can list whatever they wish to sell relating to equines.
- A news page which allow users and administers to post news from the horse world.

- An infocus page, which works like blog, for users to share their personal life and showcase their lovely horses.
- A Humans n Horses page where the administer shares horse stories from all around the world.

The owner of this website spent on the majority of her life working with horses so she then decided to start an equine-related website business. Although she has no computer science background, she is passionate about her job.

This website was half-developed by Latch Digital and Webpartners. The Latch Digital spent half a year to finish the fundamental framework then left the project before I joined. Webpartners spent another half a year to continue work on this project, but they did not manage to finish this project either. Webpartners halted work on this projectin April. I, with the help of several key people, then rebuilt the whole website without changing the framework.

The framework of this system is developed by Latch Digital themself instead of using common PHP frameworks. The developer who works for Webpartners lives overseas so it was quite hard to communicate with him. Both companies' developers left no documentation for the program.

Thus, I spent a huge amount time to understand the foundation in the beginning. I then started to make visible progress from June. Also, due to some unclear concept and ideas, many functions of this website have been rebuilt many times during this process.

iSpyHorses was soft launched on August and was under testing to make more improvement. Now the website is officially launched and there is about 100 visits per day.

1.2 PEOPLE INVOLVED

There are some people involved in this project, from both the University of Auckland and iSpyHorses.

Ulrich Speidel: my academic supervisor for this project.

Sathiamoorthy Manoharan: the coordinator of the Bachelor of Technology (Information Technology).

Heather Cato: the owner of iSpyHorses who I directly work under.

Dave Wilcox: the senior PHP developer who joined this project in June.

CHAPTER 2

INTRODUCTION

This chapter provides simple explanations for all related tools and applications.

2.1 SERVER-SIDE

This website used to be hosted by Worldnet, but it has now been moved to Optimus System now. Both companies are located in Auckland. Both of them provide highly efficienct service to maintain the server when problems occur as well as taking daily care of operations.

2.1.1 VIRTUAL PRIVATE SERVER

A virtual private server is used with Optimus System web hosting service while a physical server is used with Worldnet. Virtual private server is a virtual machine which runs its own copy of an operating system [1]. Customers have high priority to access the operating system and can install almost any software that runs on that OS. A virtual private server is functionally equivalent to a dedicated physical server, but it is much simpler to create and configure. The price is lower than the physical server. However, the performance is lower than the physical server since it shares underlying physical hardware with other VPSs. A certain delay is detected with VPS during processing.

2.1.2 APACHE SERVER

The Apache server is the world's most widely used web server software [2]. The software can be used in the different operating system, including Linux, Windows, FreeBSD, Solaris, OS X, Unix, NetWare, OS/2, TPF, OpenVMS, and eComStation. Linux is the most common and popular operating system to use Apache server. In our case, we are using Linux as the operating system.

2.1.3 SMTP MAIL SERVER

Simple Mail Transfer Protocol (SMTP) is an Internet standard for electronic mail transmission. SMTP by default uses TCP port 25. It uses it for sending and receiving emails to communicate with customers.

2.1.4 DATABASE SERVER

A database server provides database services to other computer programs or computers in a client–server model. This can be described as a computer which is dedicated to run such a program. Database management systems regularly deliver the functionality of a database server. Some database management system such as MySQL rely entirely on the client–server model for database access.

Such a server can be accessed from a "front end" running on the user's computer which displays all tables and schemas of requested data and the "back end" which runs on the server and handles tasks such as data analysis and storage.

Most of the Database servers work on the basis of Query language. Each Database understands its query language and converts it to Server readable form and executes it to retrieve the results [4]. Although every server uses its own query logic and structure, the SQL query language is quite similar in all relational database servers.

MySQL is used for this application.

2.1.5 CONTROL THE SERVER

2.1.5.1 SSH

Our Linux server has been provisioned with SSH access.

Secure Shell, or SSH, is a cryptographically encrypted network protocol which is used for remotely connecting machines or servers from another machine in a secure way.

This allows a user to remotely control a machine's command prompt. It also allows a user to establish a secure channel over an insecure network in a client-server architecture, connecting an SSH client application with an SSH server [4]. Normally a remote command-line login is required.

2.1.5.2 PuTTY

PuTTY is a free implementation of Telnet and SSH for Windows and Unix platforms, along with an xterm terminal emulator. It is written and maintained primarily by Simon Tatham [5]. Use PuTTY to SSH to the server machine with command prompt.

2.1.5.3 FileZilla

FileZilla, a cross-platform FTP, SFTP, and FTPS client with an enormous list of features, supports Windows, Mac OS X, Linux, and more. We use FileZilla to transfer files between the local machine and the server. It allows comparison of files and efficient synchronism.

2.1.5.4 Plesk

Plesk software package is a commercial web hosting automation program. It is a web control panel for hosting website. It allows control of the server through a web-based interface. All file transformation or database management can be done with Plesk. Moreover, setting up email accounts and manage DNS are also available.

2.1.5.5 SVN Repository

Subversion is a free/open source version control system. That is, Subversion manages files and directories, and records the changes of files. This allows the recovering of older versions of data or examining of the history of how the has data changed [5].

Currently, there are two people who work on iSpyHorses.com. Subversion can make sure these two people do not overwrite each other's work. It helps us save time in comparison and data loss.

2. 1.6 LOCALHOST

To work more efficiently with web development without update the code to the server each time. A localhost server has been set in my local computer to develop the website. XAMPP is an Apache distribution which contains all requires for PHP development, such as MySQL, PHP, and Perl.

2.2 WEB APPLICATION

2.2.1 PHP

The iSpyHorses.com is based on PHP language. Hypertext Preprocessor (PHP) is a server-side scripting language which is designed for web applications. It is very popular in web development as well.

PHP program is easy to work with HTML code. It also can be used with different template engines and web frameworks. PHP code is usually processed by a PHP interpreter, which is usually implemented as a web server's native module or a Common Gateway Interface (CGI) executable [7]. The PHP code can produce a readable webpage after interpretation and execution.

2.2.2 MySQL

MySQL is a widely used relational database management system and open-source RDBMS. The SQL shortening stands for Structured Query Language.

In addition, MySQL is frequently used for web applications. Free-software-open source projects that require a full-featured database management system often use MySQL. As well as many high-profile, large-scale websites.

2.2.3 HTML

HyperText Markup Language is the standard markup language used to construct web pages. Web browsers read HTML files and display visible or audible web pages. Browsers translate HTML tags and scripts to present the content.

2.2.4 CSS

Cascading Style Sheets (CSS) is a style sheet language used for formatting a webpage. The language can be applied to any kind of XML document to perform a styled output.

2.2.5 jQuery and JavaScript

JavaScript is a scripting language that was designed for use in website development. It has been used for server-side programming, game development, and even creating desktop applications [8].

JQuery is simply a specific library of JavaScript code. It is the most common library for web applications because it is easy to use and extremely powerful.

It is noteworthy that both JavaScript and jQuery are JavaScript while many people confuse them as two different scripting languages. The difference is that jQuery has been optimised to perform many common scripting functions with less coding work.

JavaScript allows developers to show different actions for webpage, such as animations and event handler.

2.2.6 FRAMEWORK

A web application framework supports the development of dynamic websites, web applications, web services and web resources. The framework aims to improve the overhead associated with common activities performed in web development. For example, many frameworks provide libraries for database access, template frameworks, and session management, and they often promote code reuse. Therefore, developers can easily code with existing stable and powerful structure.

CHAPTER 3

PROGRAM STRUCTURE

This chapter describes the programming structure of iSpyHorses.com in order to introduce the main operation of this system. Although the framework of this website is developed by the previous developer, some fundamental work is based on Symfony.

3.1 MAIN MODULES

Firstly I will introduce the main components in the web application layer. Each of them play a different role to process this website.

3.1.1 VIEW

A folder which is called View contains all the html files to manage the layout and presentation of iSpyHorses.com.

3.1.2 ASSETS

This folder contains all stylesheets and JavaScript files to format the website. Furthermore, it stores the pictures that is updated by users.

3.1.2 MODEL

PHP classes are written in this folder to connect and to query the database.

3.1.3 CONTROLLER

PHP classes in this folder control the communication of database and webpage.

3.2 FRAMEWORK

3.2.1 TWIG

Twig is a template engine for PHP. A handbook which is given by [9] has explained how twig works and will be discussed in more detail later.

3.2.2 APERTURE-CORE FOLDER

The fundamental work is defined in this folder. Four main PHP files control the process of calling files in core structure.

- apertureControl.php
  Define the classes in the Control folder to load model classes and html pages as well as mailing functions.
- apertureModel.php
  Define the classes in Model folder.
- apertureEndpoint.php
  Define the path of HTML files.
- apertureRoutes.php
  Define the way to read routing file.

3.2.3 APERTURE.PHP AND APERTURE.JSON

The environment of the server and the website is set in aperture.json file. When the browser reads the index.php file, it requires aperture.php to check the environment with aperture.json. If the environment matches then it will start to process the website by calling routing files.

More details will be given in next section.

3.3 PROCESSING

This section will explain the processing with example code.

3.3.1 ROUTE.PHP

```
route('GET /index', $this->to("index", "index"));
```

In route.php, urls are read like above. This means a path with /index executes indexcontroller.php (the first index inside bracket)'s index() function(the second index).

3.3.2 INDEXCONTROLLER.PHP

```
public function indexAction() {
    $focus = Profile::getInFocus(4);
    $highlights = Highlight::getHome();
    $horses = Horse::getFeatured(6, 1, 'rand()', 'feature = 1');
    $slides = Horse::getFeatured(6, 1, 'rand()', 'superfeature = 1');

    echo $this->view("index/index.html", array(
        'slides' => $slides,
        'highlights' => $highlights,
        'focus' => $focus,
        'horses' => $horses
    ));
}
```

Variables are read from the Model file. For example, Horse.php returns values for the variable, horses. The view method was defined in apertureController.php which outputs horses' value to index.html.

3.3.3 MODEL/HORSE.PHP

```
static function getFeatured($limit = 6, $page = 1, $order = 'id desc', $type = 'feature = 1 or superfeature = 1') {
    $extraConditions = array(
        'and (' . $type . ') and expiry_date > now() and is_slider=1'
    );
    return self::search($extraConditions, $limit, $page, $order);
}
```

Horse.php then queries the database to return the values to the controller file.

3.3.4 INDEX.HTML

```
{% for horse in horses %}
<div class="horsepic">

<a href="{{horse.link()}}"><img width="100%" src="{{horse.image()}}" /></a>
<p>{{horse.name}}</p>
</div>
{% endfor %}
```

Twig applies an HTML file to read the database value from controller file. In the webpage this would output the name, link, and image of all the horses from the query.

CHAPTER 4

THE WEBSITE

This chapter presents all the work I have done with iSpyHorses.com in the year 2015. Although the website was 70 percent finished when I joined, I was the only developer who worked on this project from early April to late May. Also this website has been rebuilt due to the requirements of the owner so most functions are completely different from the previous work. The main problem for me in the beginning was to understand the entire system which was established by previous developers. It was not easy in the beginningsince there was no documentation or comments in the code.

The major achievements of the first semester were understanding of the application and communication with the owner along with the implemention of some functions to the website.

I then finished this project in second semester, the website works fine in many aspect. This chapter provides the details of the website.

4.1 OVERALL LAYOUT

The first job was to change the overall layout from an entirely different design to another. The overall layout has changed three times during development. I implemented my own CSS stylesheet along with a widely used online stylesheet, bootstrap, to modify the performance. To make the website responsive to different screen sizes, viewport size and percentage width are commonly used in this stylesheet. Float and overflow properties are also quite important to format the website.

I have noticed that some styles cannot exactly match the design because it is not only a stylesheet issue. It is also related to functions behind the code. Consequently, some parts of the design seemed hard or even impossible to plug-in.

Another problem was with the design as the design comes after development. The designer had not communicated with developers before he output the layout.

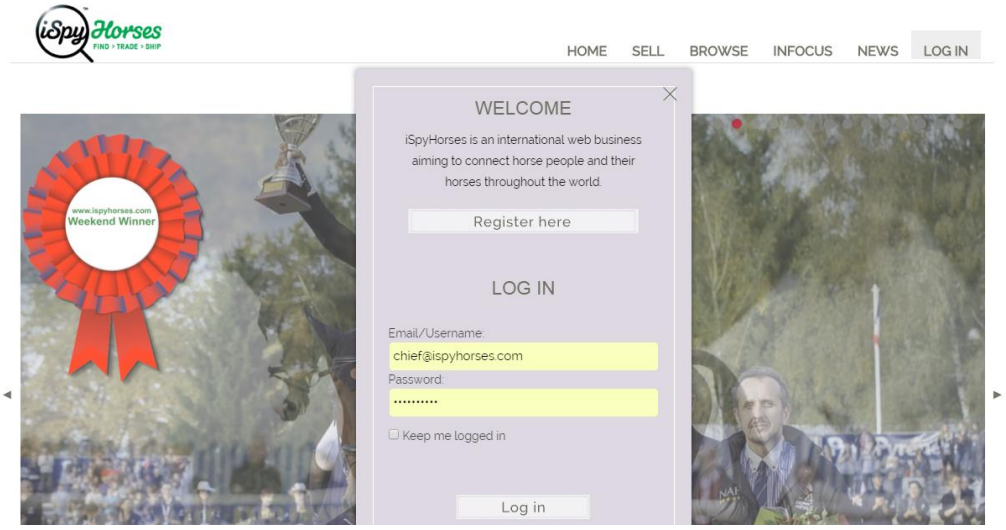
Thus, many functions are not in the application but it is required by the design.

Heather, the owner, also had changed her mind for the layout part several times. Therefore, it took quite long to finish. All styles have been tested under all different resolutions. More details will be given in later section.

4.2 JAVASCRIPT AND JQUERY

Some JavaScript files and jQuery files have also changed or been added when applying the new design to the website. The picture 4.2.1 is an example of the pop-up window.

- In the homepage, the design requires a section for a set of slide banner pictures and a carousel show for members' listings.
- A pop-up window without address bar or title bar, also called modal window, is requested to load a log-in page and some form functions.
- Some DIVs need to be hidden when a click event occurs in the navigation bar.
- Enlargement of the picture when an click event of image is on fire, as shown in picture 4.2.2
- Preview image support when users upload images.
- Other general parts such as equal the height of different div and image affection.



PICTURE 4.2.1

PICTURE 4.2.2

4.3 DATABASE

Some functions of the website did not work because some variables in the controller file read non-existent values from the database.

Some values are linked into a specific ID as a default value in the database. Therefore, every time a form function is submitted, there is a default value for such table.

Most work to do with the database to clean up all the useless schema are either related or not related to model and controller files. This website has changed lots of functions during development, hence, there were lots of redundant schemas in the database as well as some expired listings which were created for testing listing functions.

All the fake data has been deleted when the website was launched. Many new tables and columns had been added when processing the final website.

4.3 SETTING UP THE SERVER

iSpyHorses.com has changed its server once during development. The new one is VPS which is hosted by Optimus System.

I did not have that much server related knowledge in the beginning, but I was able to helpto choose the one that we needed.

Most configurations are done by Optimus, but it was my job to connect the website and the server. Many problems arose during the first time updating the application to the server. The environment of mail server and database has changed due to change of website hosting. I created a new database in the new one in order to produce further development with this server.

SMTP are registered with iSpyHorse.com domain. Therefore, I have connected the domain to the provided IP address. Heather does not want the website to go public before everything is finished, therefore, an htpassword file for authorization access has been produced for security earlier. It has since been removed as it is public now.

It was important to connect the domain name to the server because the register function usesthe mail server to send new members a confirmation email.

Moreover, I have noticed that the live server is case-sensitive for file names and model classes.

4.4 FORM FUNCTION OF PHP

The listing platform is the main goal of iSpyHorses.com. Similar to trademe.co,nz, it allows users to sell their products online, but iSpyHorses are only focused on equine. I have developed one category of listing to the application.

To complete the listing program, several classes need to be modified and added.

- The router.php file

  Add a get and a post calling for server calls the control when a page is loaded.
- Add controller file

  This controller file will manage the methods of which one shows where.
- Add model file

  This one to store data to databases as well as query databases which is required by controller file.
- The HTML file

  This one is for users to read the website. Listing functions are stored in modal window therefore it's connected to JavaScript as well.

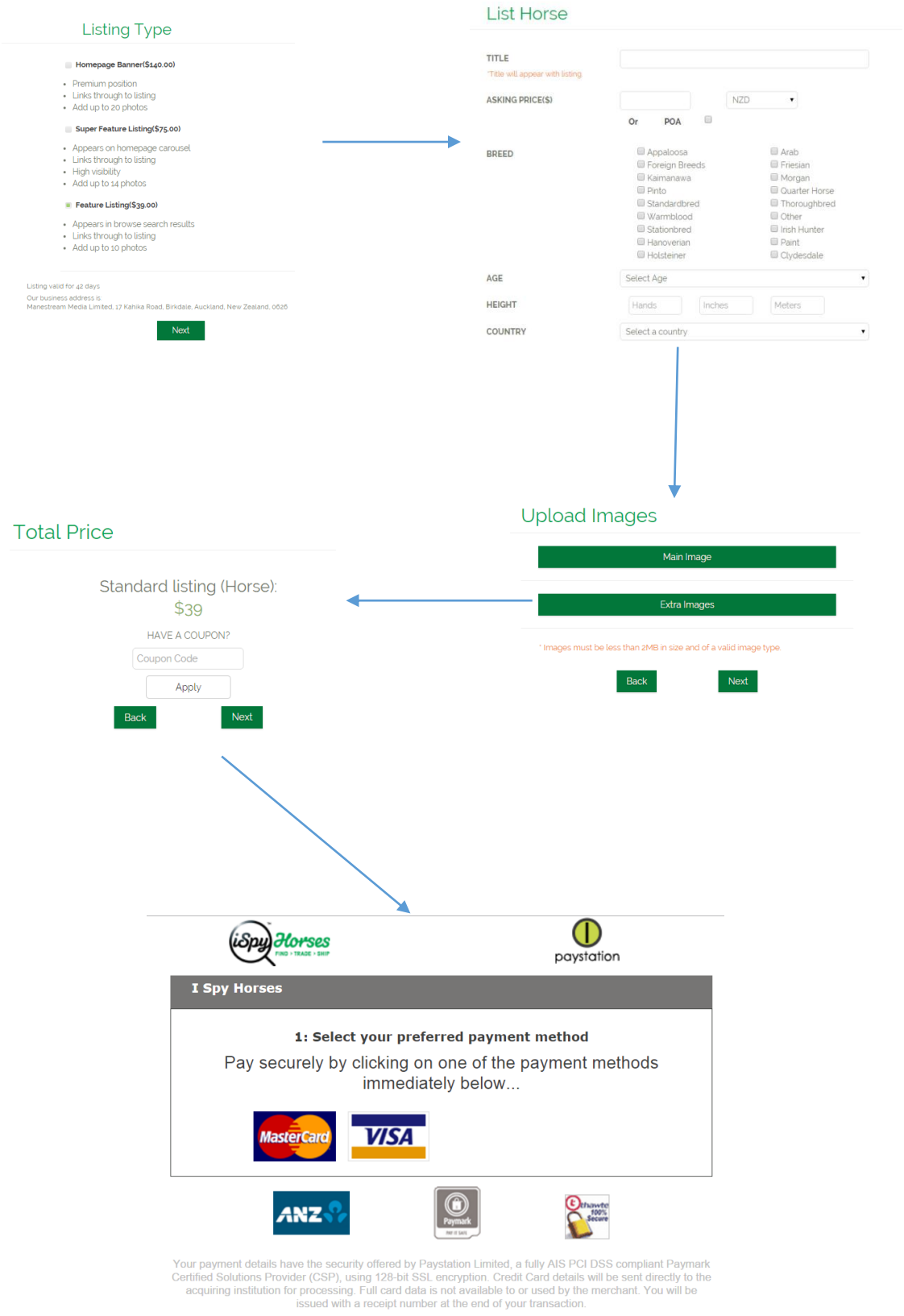
An example has been given in the earlier discussion. My implementation follows the process to add more functions to this application.

4.5 LISTING GOOD

Five sub categories to allow users list goods on the site. horses, ponies, services, vehicles, and real estate.

We do not ask users for credit card detail. All the payment processing go to paystation directly. We then store the status to analyze the income and profit.

The processing of listing are shown below:



Picture 4.5.1

4.6 SEARCH

There are search functions for each individual section. This means users can search news in news page or search listings in listing page. We allowed the worldwide search before but Heather does not want it anymore. A worldwide search means that if a user enters "test" in the search bar, the website will return all values that contain the keyword. It can be from listings, news, status, or even members who have "test" in their name.

Now the search function works individually but all information on the site is searchable with keywords.

Full-text search can be introduced in this section.

Full-text search searches for character-based data in SQL Server tables. It scans all of the words in every stored field as it tries to match search criteria. We have changed the default minimum length of characters to three incase people are trying to search “van” or “bus”.

4.6 NEWS AND INFOCUS

The news page and infocus page have same functionalities. Both of them allow users to like or comment a post.

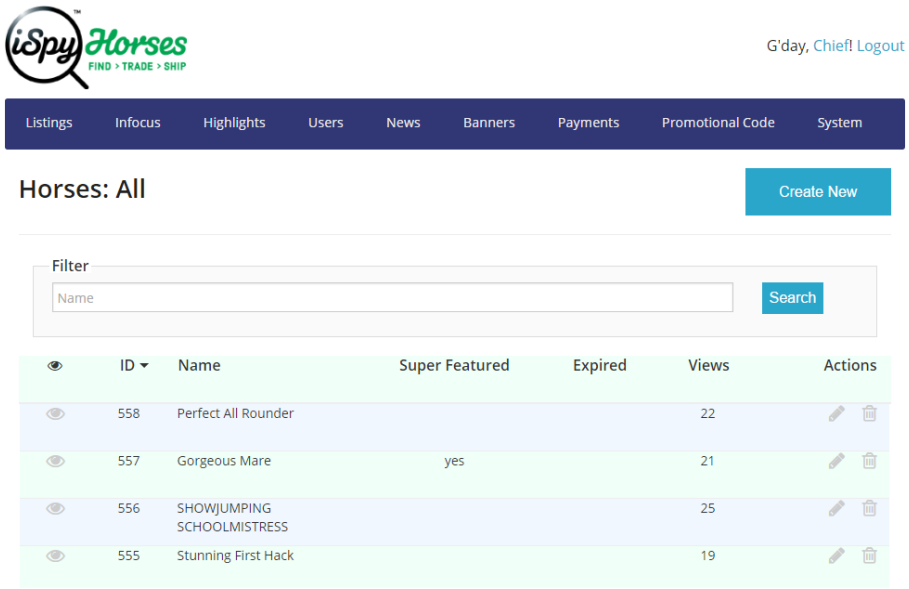
The InFocus page works more like a blog, users are able to post their daily stories and recent activities.

A ckeditor plug in text editor is used for both parts. This is to allow users style their articles and photos.

4.5 ADMIN

Letting the administer control the website without worry about codes is the proposal from the admin. In this page, admin are able to

- See all the listings made by users and edit or delete it.
- See all the news which is submitted by a user. The user should be aware that the news goes to admin first before going to the public page.
- Control the banner pictures. The banner pictures will change from time to time. Administer can edit this part without change the code.
- See all infocus posts.
- Create the promotional code for any advertising.
- A system log page for us developers to see if there is any problem such as mail sent failure, functions breakdown. It also receives all feedback from users.
- We do not store credit card detail but we store the status information we get back from PayStation. This is for evaluate profit of the website.



Picture 4.6.1

4.7 HUMANS N HORSES

It was the connect page that works like a message board or like a simple version of Facebook for each individual users. Users can follow other users and go to their communication page to leave a message. Users are able to post stories and pictures in their own page or other people's page. In addition, users can comment on single posts.

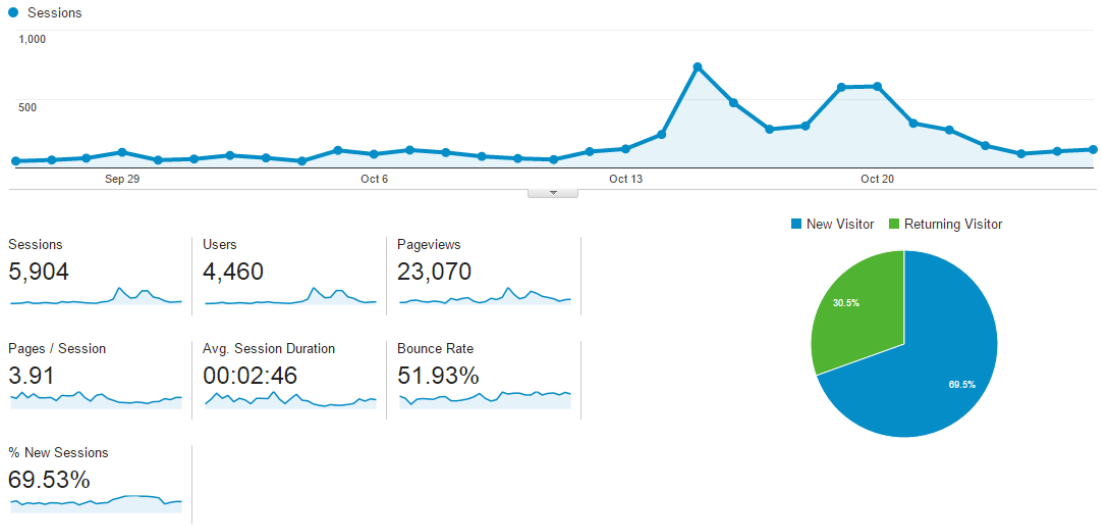
But before the website went to public, the owner decided the connect page is not needed since many people will just use Facebook.

Now this page use old function of connect page but only the admin are allowed to post. Other users can only see the posts of admin.

4.8 GOOGLE ANALYTICS

To analysis the website better, a freemium web analytics service offered by Google is embedded to the system.

The website is launched on 26th August, after that we can see picture 4.8.1 for all traffic we have so far.



Picture 4.8.1

For better performance, we found that most of our customers use mobile or tablet to browse the website. Hence, we value the responsive layout on the phone more.

	Browser	Sessions	% Sessions
1.	Safari (in-app)	2,061	34.91%
2.	Chrome	1,721	29.15%
3.	Safari	1,241	21.02%
4.	Internet Explorer	379	6.42%

	Operating System	Sessions	% Sessions
1.	iOS	2,965	50.22%
2.	Windows	1,174	19.88%
3.	Android	1,102	18.67%
4.	Macintosh	615	10.42%

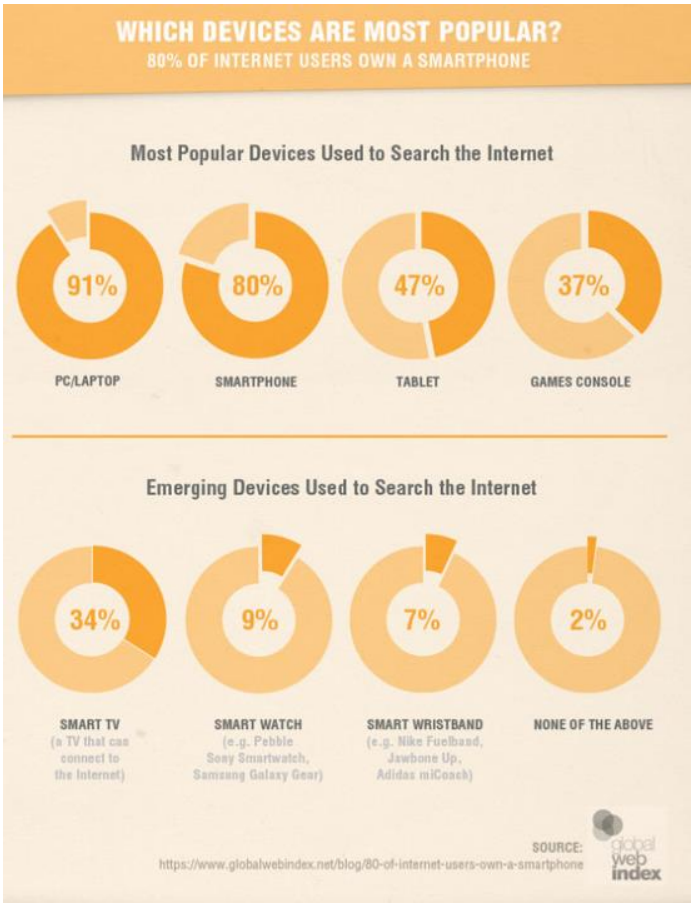
	Screen Resolution	Acquisition		
		Sessions ↓	% New Sessions	New Users
		4,120 % of Total: 69.78% (5,904)	74.61% Avg for View: 69.53% (7.31%)	3,074 % of Total: 74.88% (4,105)
	1. 320x568	1,300 (31.55%)	70.46%	916 (29.80%)
	2. 768x1024	695 (16.87%)	75.68%	526 (17.11%)
	3. 375x667	689 (16.72%)	73.58%	507 (16.49%)
	4. 360x640	664 (16.12%)	77.41%	514 (16.72%)
	5. 320x480	213 (5.17%)	85.92%	183 (5.95%)

Figure 4.8.2

CHAPTER 5

PERFORMANCE

Nowadays, a growing number of new technologies appear on the market. People are not only using PC or laptop to browse the internet but many use also use smartphones, tablet, etc. A recent market survey shows that 80% of internet users using smartphones to search internet and 47% of internet users using tablets. Figures as shown below [10]:



Picture. 5.0.1

As we can see, smart devices become very important in the daily access and usage of the internet. However, many websites do not look good in devices other than PCs and laptops.

In this chapter, I will talk about the responsive layout for different resolutions and other performance problems like browser support and Website Optimisation.

5.1 RESPONSIVE LAYOUT

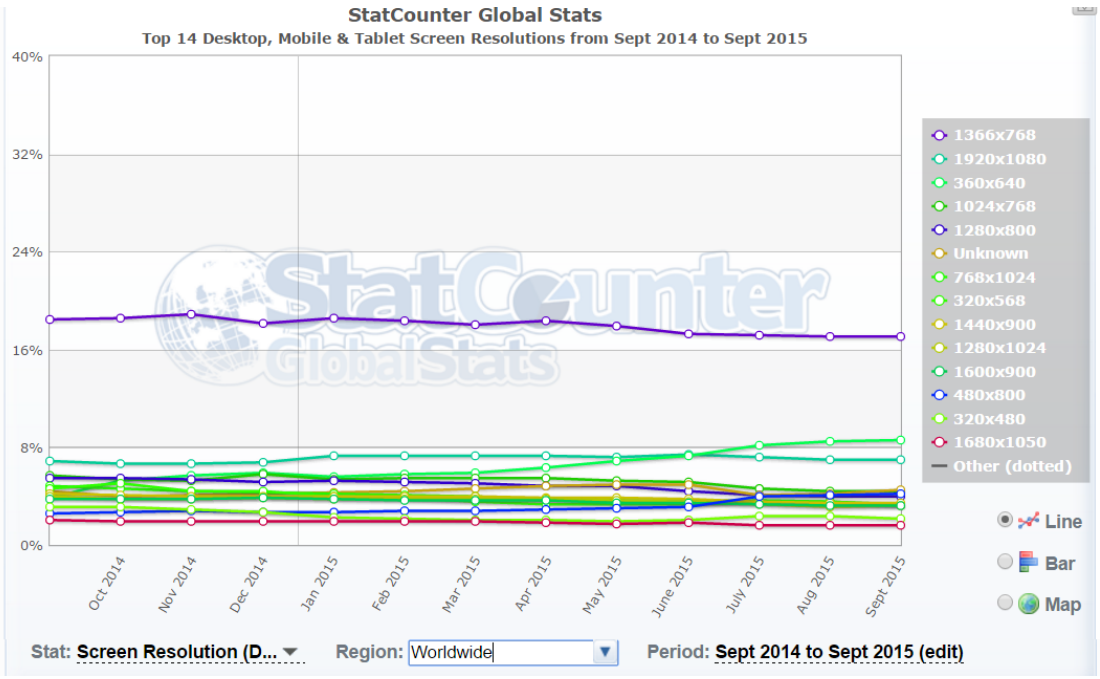
"Starting April 21, we will be expanding our use of mobile-friendliness as a ranking signal. This change will affect mobile searches in all languages worldwide and will have a significant impact in our search results. Consequently, users will find it easier to get relevant, high quality search results that are optimized for their devices."

------ GOOGLE [11]

Google indicated that their search result will show mobile-friendly website first from google searches. The mobile-friendly layouts are becoming more important for web development.

Also, there are more and more different screen resolutions coming from new technologies. To have a satisfactory display of a website in different devices, a website has to cope with these differences. How to make a website responsive to all PC, laptop and smart devices is becoming the first question we need to think about during development.

For our website, we decided that treating mobiles and tablets as different resolutions of window screen would be the best solution to make a responsive website.



Picture. 5.1.1

Picture 5.1.1 shows the most popular resolutions in the world. Although iSpyHorses.com is applicable to all resolutions, we mainly focused on 1920x1080, 1366x786, 768x1024 (iPad) and 480x800 (phone) during development.

To make a responsive website, the grid system, bootstrap, can be introduced to solve the problem:

A webpage is divided into many rows and columns to layout the content.

Several rules are defined for the grid system [12]:

- Rows must be wrapped within a fixed-width container or full-width of screen for proper alignment and padding.
- Use rows to horizontally align columns.
- Content should be placed within columns, and only columns may be immediate children of rows.
- One row is split to 12 columns with equal width, grid columns are created by specifying the number of columns that it wishes to span. For example, three equal columns would use three .col-xs-4.

- If more than 12 columns are placed within a single row, each group of extra columns will, as one unit, wrap onto a new line.

The grid system uses the media queries in CSS files to create the key breakpoints:

```
/* Extra small devices (phones, less than 768px) */

/* No media query since this is the default in Bootstrap */


/* Small devices (tablets, 768px and up) */

@media (min-width: @screen-sm-min) { ... }


/* Medium devices (desktops, 992px and up) */

@media (min-width: @screen-md-min) { ... }


/* Large devices (large desktops, 1200px and up) */

@media (min-width: @screen-lg-min) { ... }
```

We break all the different screen sizes to four particular ranges, for some overlarge resolutions we set a maximum width in CSS stylesheet to make sure the web page does not look too enormous. Table 5.1.1 below gives some basic information of the breakpoints:

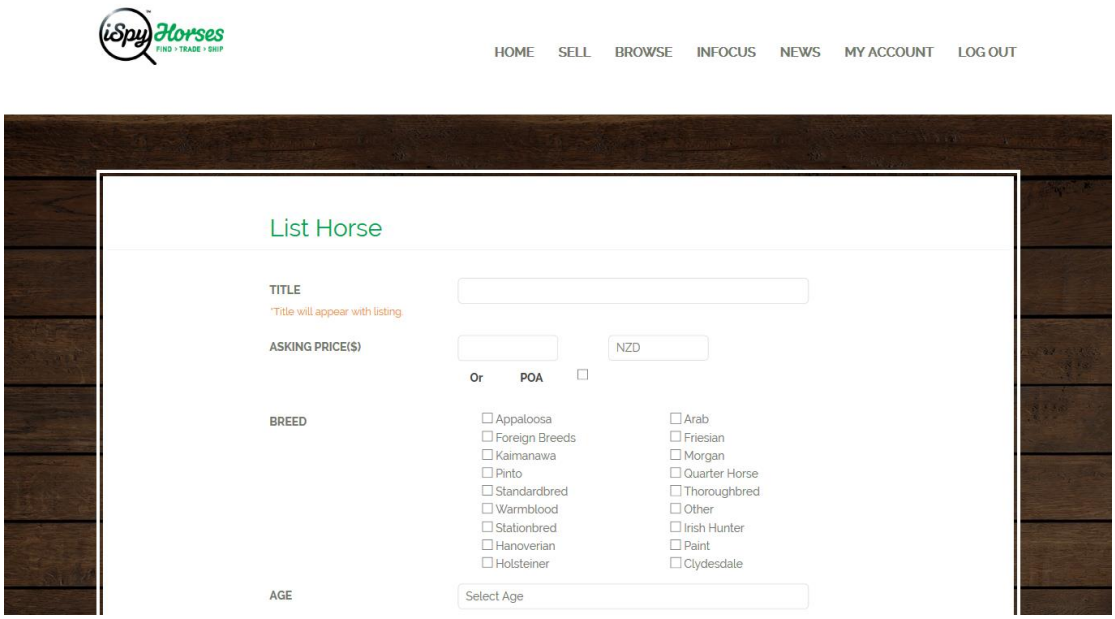
	Extra small devices Phones (<768px)	Small devices Tablets (≥768px)	Medium devices Desktops (≥992px)	Large devices Desktops (≥1200px)
Grid behavior	Horizontal at all times	Collapsed to start, horizontal above breakpoints		
Container width	None (auto)	750px	970px	1170px
Class prefix	`.col-xs-`	`.col-sm-`	`.col-md-`	`.col-lg-`
# of columns	12			
Column width	Auto	~62px	~81px	~97px
Gutter width	30px (15px on each side of a column)			
Nestable	Yes			
Offsets	Yes			
Column ordering	Yes			

Table 5.1.1

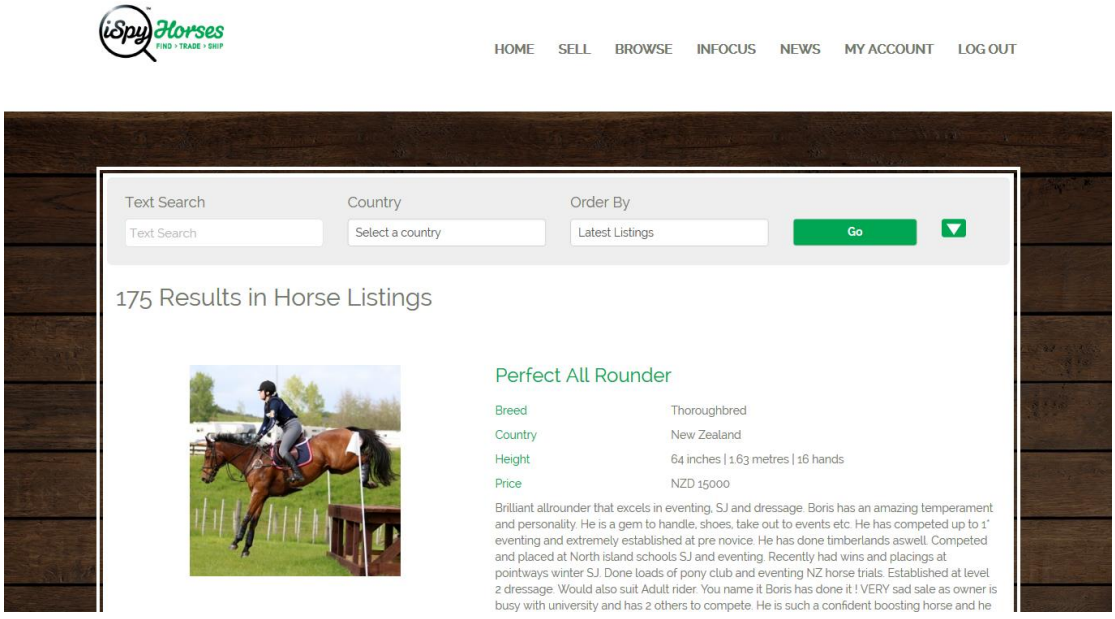
The grid system, Grid classes apply to devices with screen widths greater than or equal to the breakpoint sizes, and override grid classes targeted at smaller devices. Therefore, .col-md-* class can only affect the style of medium devices and large devices if a .col-lg-* class is not present [12].

The following gives a few examples of layout in resolution 1366x768 (this layout applies to all widths which are bigger than 768px):

Homepage

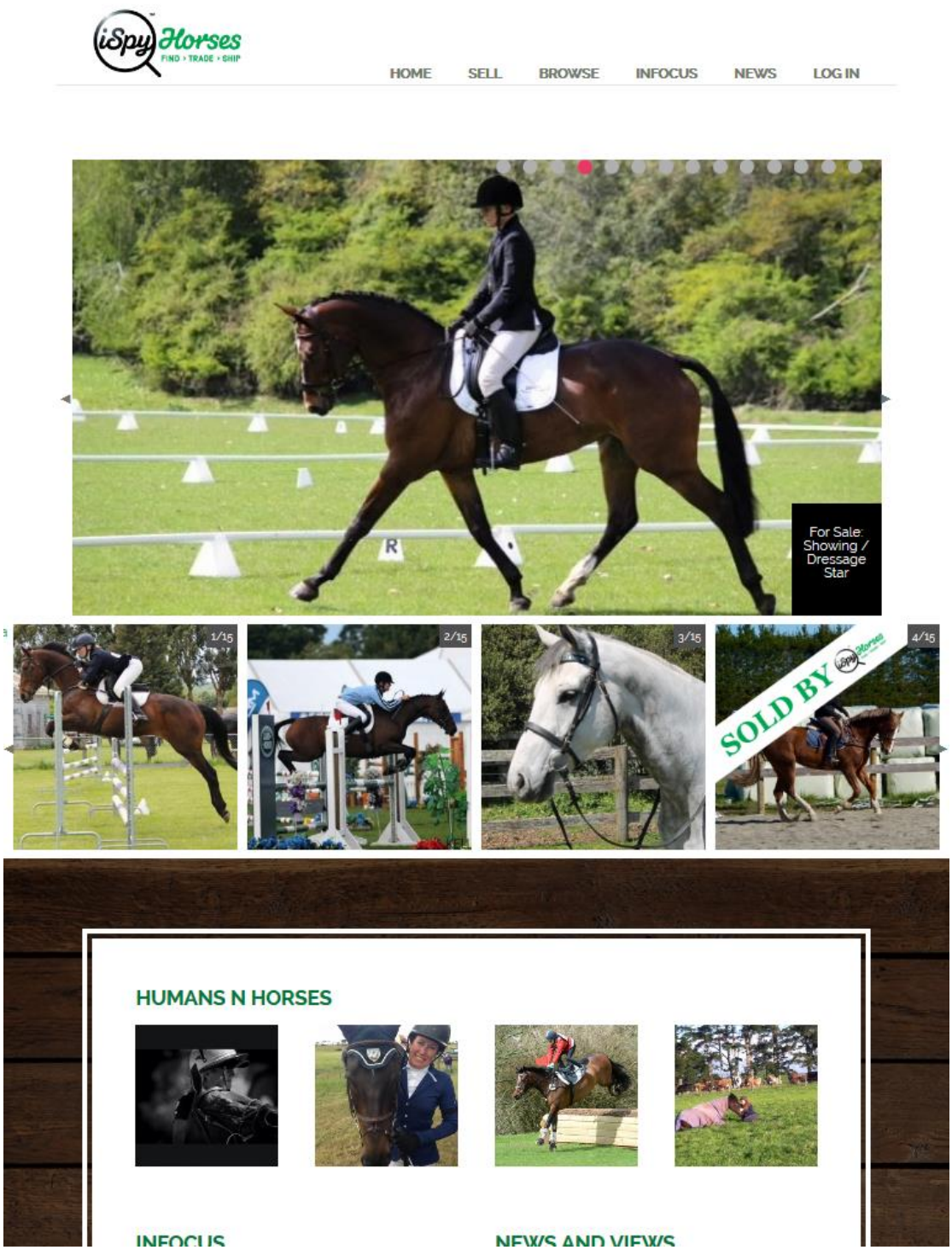


List Horse



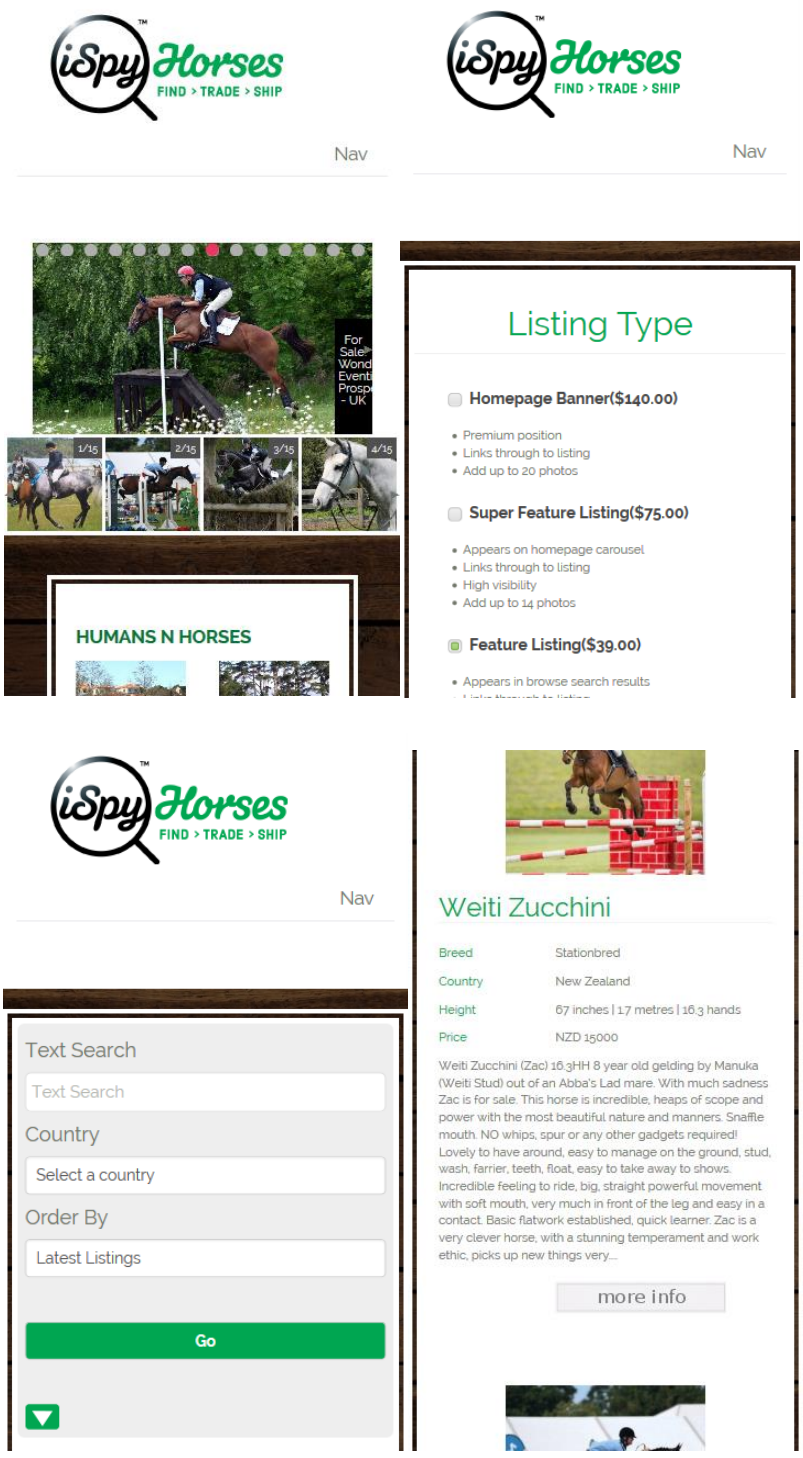
Browse

For tablet screens, we set all the layouts to be the same as PC/laptop but with smaller font size, therefore there is not much difference between 1366x768 and 768x1024:

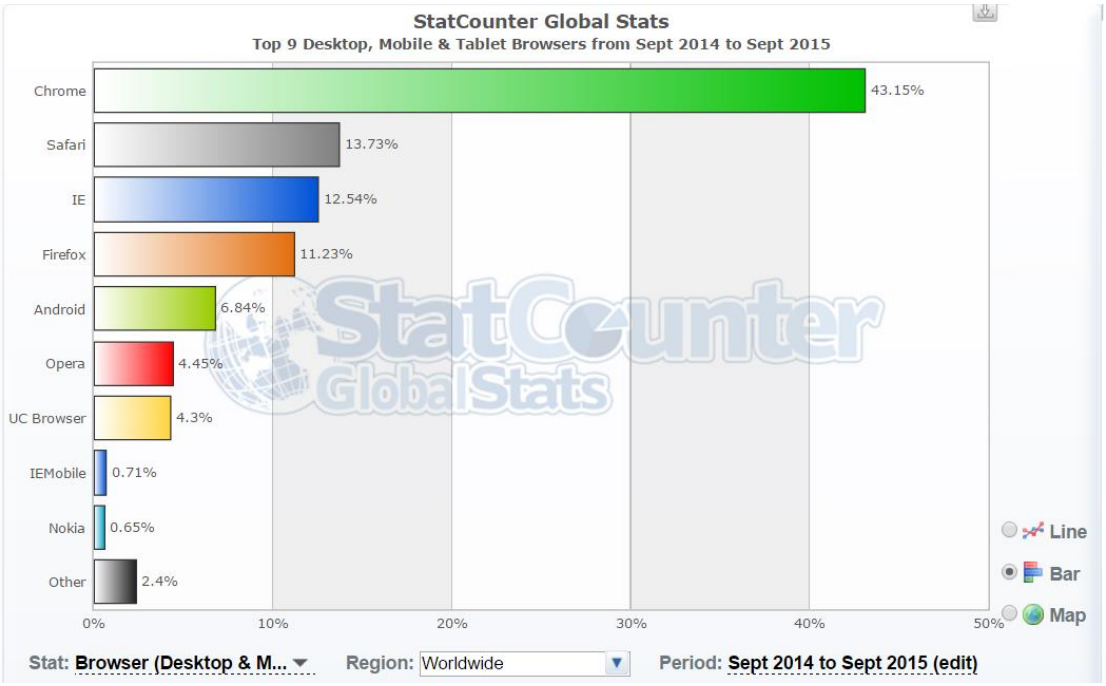


Homepage in 768x1024

In phone layout, the homepage did not change much due to the requirement of owner. The following gives the website layout under resolution: 360x640 (from left to right, top to bottom: homepage, List Horse, search section of Browse and result section of Browse):



5.2 BROWSER SUPPORT



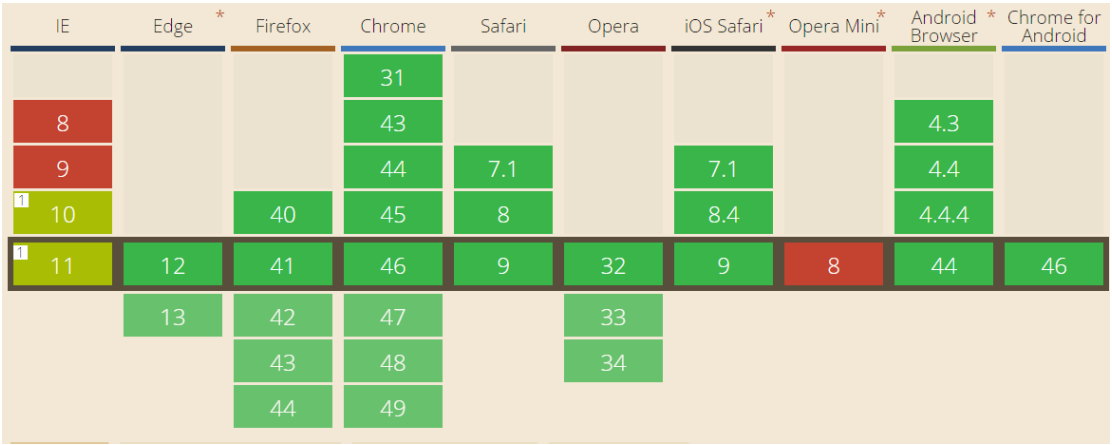
Graph 5.2.1

The above Graph 5.2.1 indicates that Chrome is the most popular among all browsers. But Safari, IE, and Firefox are also commonly used. In the beginning of this project, I was focused on developing the website on Chrome therefore there were many problems in other browsers.

Some CSS merging and padding display differently in different browsers, and so does line height. We went through all the details after we observed the browsers differences. Many problems can be solved by redefining the CSS stylesheet.

Not only were theproblems due to different browsers, but also browser version differences. For example, in our image system, a preview image shows after the user updates their image. We use FileReader API in jQuery to support this function. However, it does not work in older versions of Safari and IE.

Picture 5.2.2 gives the overall information of FileReader API in different browser. Green means it is supported while yellow means partial support. Red means not supported at all. The numbers represent the version.



Picture 5.2.2

We can see that most commonly used browsers such as Chrome, Firefox, Safari and IE support the FileReader API. However, we received feedback from one customer that she cannot see updated images and indicated that she was using Safari.

After testing the website in our local computer with higher version of Safari, we noticed it worked fine. We then noticed that she was using an old version 5.1 which does not support FileReader at all.

To solve this kind of problem, we have suggested customers to use the latest version of their browser.

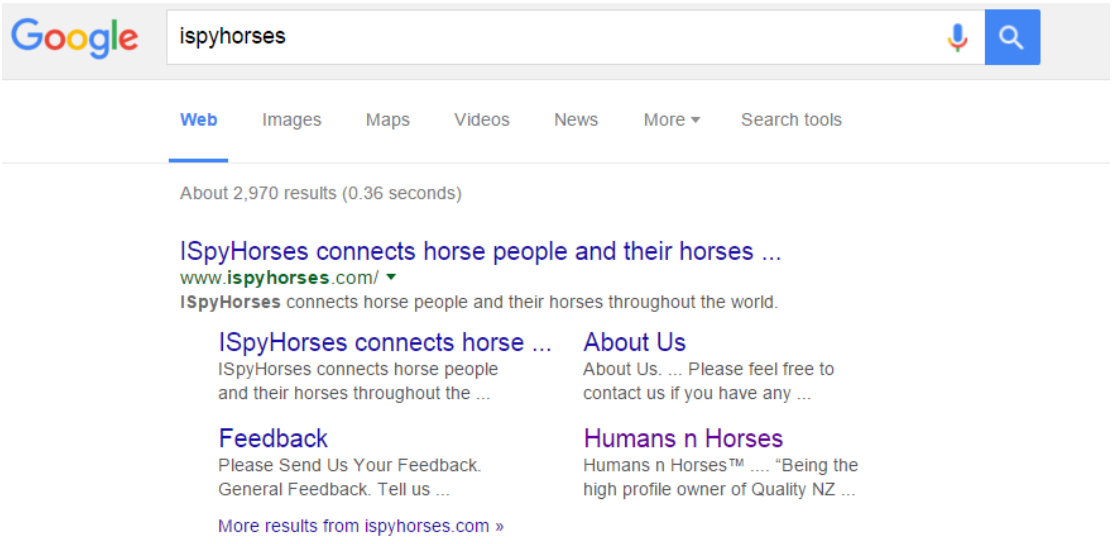
5.3 To the Top of Google

Search engine optimisation can have a big impact on the success of an online business. The higher the website is listed in the search result page, the more people a website can get.

Lots of people only look for results on the first page, hence, to be among the top listings become very significant for a website. To do this, Search Engine Optimization (SEO) can be very useful to web development.

Firstly, let's see how the search engine works. Once the search engine knows a site exists, they then scan the website, index the information, and analyse the content to decide how and where the website should display on the results page.

Therefore, we d optimised iSpyHorses.com by a few ways in order to increase the ranking form the search engine.



Picture 5.3.1

SEO tools like Google Analytics are used to identify keywords.

The title tag provides a brief summary of iSpyHorses to tell people what they see from the website. It displays the browser's title bar as the title in search engine results. Different pages have different tags.

Description Tag, "iSpyHorses connects horse people and their horses throughout the world", describes more details of the page. Every page includes a unique description using the keywords for that page.

Each page includes only one header tag. The search engine read the largest or most prominent text on the page to tell visitors what they're reading or viewing.

Sitemap guides search engines throughout the website with the names and locations of pages. They can speed up indexing and, in some cases, increase site traffic by indexing previously buried pages.

Image Tag with an "alt" attribute uses keywords for the page. Search engines cannotsee images therefore they depend on alt attributes to appropriately catalog and index the image. In our image system, we resize the updated images of users to make sure there are no overly large images to slow down loading.

5.4 JUNK MAIL FILTER

In the first few weeks after the website went to public, we received many complaints from customers saying they were not able to get the registration email from us.

Most of the auto-send emails went straight to their junk mailbox instead of inbox. In addition, some of them set their junk box to automatically delete spam emails.

As a new website, it was not a total surprise to us. Most websites experience similar issues in the beginning when they are not yet well-known.

To reduce the chance of our emails going to spam, we have re-designed the layout of the mail so it looks more professional.

We checked the many known blacklists to see if our server is blacklisted. Fortunately, we are not in one of them.

We have also tested the SPF record, MX record and Reverse DNS to see if there is any irregular problem.

To date, the number of people who cannot receive emails from us has significantly reduced.

CHAPTER 6

SECURITY

A website, such as iSpyHorses.com, an online listing platform, can be the number one target for malicious online attacks. The most important factor about a business is money. Hackers will try to get all information from a website in order to steal what they want. Webpages attacked by mass SQL Injection are around 1.5 million in the last year alone, including those belonging to the Department of Homeland Security, the United Nations, and Sony, among others [13].

All the security loopholes can cause massive losses of traffic, and more importantly, money. Therefore, web security has become the number one topic in the web development.

Figure 6.0.1 shows the top ten vulnerability classes in web security.

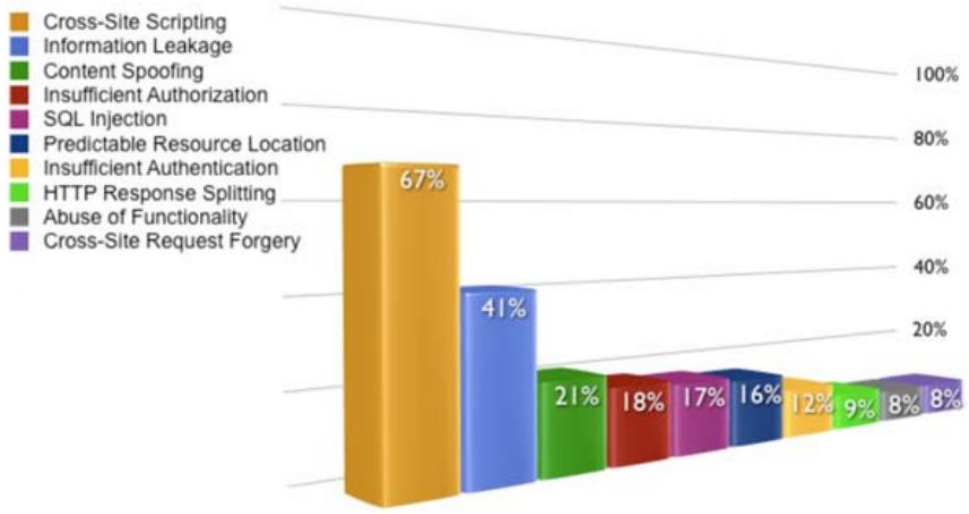


Figure 6.0.1

In this chapter we mainly discuss popular problems related to iSpyHorses.com.

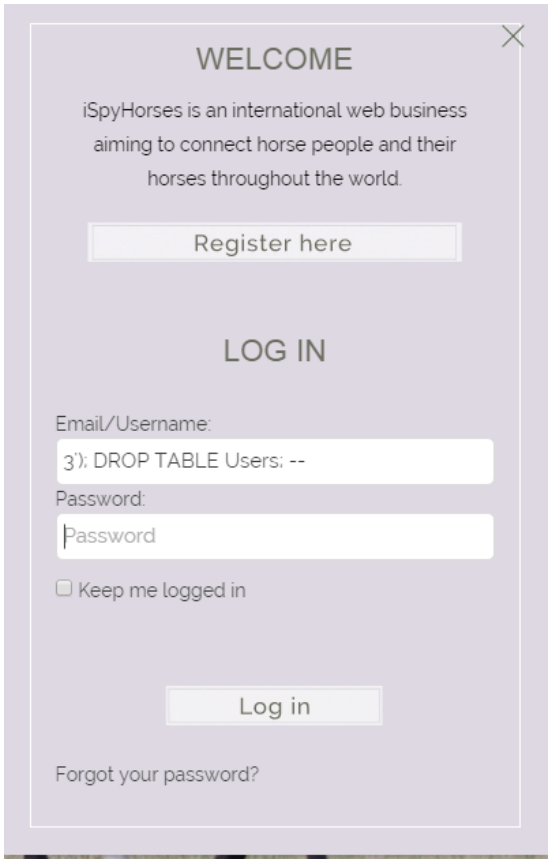
6.1 Cross-Site Scripting and SQL Injection

SQL injection is a code injection technique, used to attack data-driven applications, that allow attackers to obtain unrestricted access to the databases underlying the applications and to the potentially sensitive information these databases contain [14]. For example, when a user input contains SQL query and the system execute the query without escaping the illegitimate characters. SQL injection is not only the most known attack vector for websites but it also can be used to attack any type of SQL database.

In iSpyHorses.com, a query is used to look up the users in database while users log in:

SELECT * FROM Users WHERE UserName = $input of username;

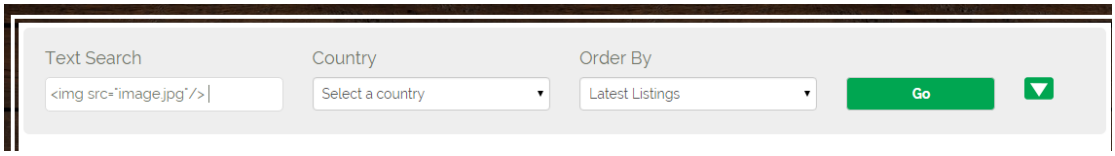
In picture 6.1.1, if a user inputs their name as “3'); DROP TABLE Users; --“then the original query will drop the table of users after execution.



Picture 6.1.1

Cross-Site Scripting (XSS) attacks are another type of injection, which allow injection of code into the client side of a web application. XSS attacks occur when an attacker uses a web application to send malicious code, generally in the form of a browser side script, to a different end user [15]. A user can attack a website by input script or html code to the site in order to execute the script or HTML. The malicious script can access any cookies, session tokens, or other sensitive information retained by the browser and used with that site. These scripts can even rewrite the content of the HTML page.

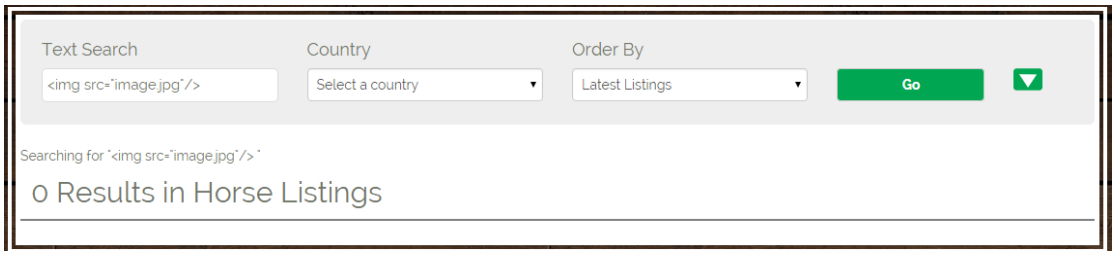
For example in picture 6.1.2, an image with name image.jpg with full path of source could display after clicking the Go button.



Picture 6.1.2

To ensure such problems will not occur in iSpyHorses.com, we made sure that the system read the string only for nput. Therefore, the system needs to escape all the special characters and HTML tags.

Picture 6.1.3 shows the returned result after a completed search.



Picture 6.1.3

6.2 CODE UPLOAD

Users are able to upload their own pictures onto iSpyHorses.com in order to advertise their goods. However, in an uploading system, attackers can upload any files, such as PHP file, JavaScript, or any others, to the website server. It will certainly leave a "backdoor" for the hacker to access to the web server in order to steal or destroy the data stored on the server.

These kinds of backdoors are not easy to find once they have been placed to the server. Nevertheless, it is easier if the system can stop it in the beginning.

To guarantee the system does not store odd files to the server, we defined that the uploading system can only upload files with correct image extension: jpg, jpeg, png and gif.

6.3 FAIL2BAN

Fail2ban scans log files and bans IPs that have malicious signs. The application can capture who is trying to guess a password with many failures and who is seeking to exploit the server. Generally Fail2Ban is used to update firewall rules to reject the IP addresses for a specified amount of time, although any ordinary actions, for instance, sending an email, could also be configured. .

Fail2Ban can decrease the rate of incorrect authentications attempts, though it cannot reduce the risk when weak authentication is presents.

6.4 THIRD-PARTY PAYMENT

Definition [16]: A third-party transaction is a business deal involving a buyer, a seller and a third party. The third party's involvement varies with the type of business transaction. For example, an online payment portal, such as PayPal, acts as a third party in a retail transaction. A seller offers a good or service, and a buyer uses a credit card

entered through the PayPal payment service. The payment is run through a third party, and is therefore a third-party transaction.

The third-party involved in this business is Paystation, a New Zealand local online payment processing service.

Using a third-party payment can confirm that iSpyHorses.com does not store users' credit card or any financial details.

6.5 AUTHENTICATION

A plaintext passwords should never be stored on the server therefore we keep a hash of the password to ensure we do not store customers' personal information.

CHAPTER 7

CONCLUSION

Up to now this project has finished and the maintenance of the website is under control. The main goal in the first half semester was to understand the processing of the framework. I then started to contribute my own work to implement the website after Easter. I have learnt a lot more than just programming from this project. For instance, I have improved my communication skills to appropriately communicate with people with different levels of website development knowledge. I am glad that I have achieved progress for this project by the end of the first semester.

In the beginning I had some difficulties trying to understand the real ambition of iSpyHorses.com because of confusion with some concepts and ideas which are required by the owner as the owner, herself was unclear of what she wanted. Therefore, we have rebuilt many functions to achieve the final goal. Finally the project managed to be completed in October. I have noticed that the objective of this project is not enormously hard to achieve, but it does require enormous effort to make the website perfect.

To optimise the website performance we did many additional implementations to present the website in a better to the user. We also checked the common security problems in case of possible attacks to iSpyHorses.com. Latch Digital did a good job with their fundamental framework thus the website is secured and stable.

We have received much feedback from real users and many suggestions are valuable and feasible. They have also helped with improvements such as finding problems that we were not aware of.

Overall this is a good project and I really enjoyed being a part of this experience.

ACKNOWLEDGEMENTS

Many thanks to Ulrich and Mano for their useful comments and suggestions.

Many thanks to my mentor, Dave Wilcox, who helped me a lot in last few months.

Many thanks to Tania Tian for her support.

BIBLIOGRAPHY

[1] Virtual Private Server

http://en.wikipedia.org/wiki/Virtual_private_server

[2] Apache HTTP Server

http://en.wikipedia.org/wiki/Apache_HTTP_Server

[3] Database Server

http://en.wikipedia.org/wiki/Database_server

[4] Network Working Group of the IETF, January 2006, RFC 4252, The Secure Shell (SSH) Authentication Protocol

http://tools.ietf.org/html/rfc4252

[5] PuTTY: A Free Telnet/SSH Client

http://www.chiark.greenend.org.uk/~sgtatham/putty/

[6]What is Subversion?

http://svnbook.red-bean.com/en/1.6/svn.intro.whatis.html

[7] Why PHP development is very popular in the web development industry?

http://www.sircltech.com/category/web-development-with-php/

[8] JQuery vs. JavaScript: What's the Difference Anyway?

https://blog.udemy.com/jquery-vs-javascript/

[9] Twig is a modern template engine for PHP

http://twig.sensiolabs.org/

[10] Mobile Marketing Statistics 2015

http://www.smartinsights.com/mobile-marketing/mobile-marketing-analytics/mobile-marketing-statistics/

[11] Finding more mobile-friendly search results

http://googlewebmastercentral.blogspot.co.nz/2015/02/finding-more-mobile-friendly-search.html

[12] Bootstrap CSS

http://getbootstrap.com/css/

[13] Grossman, Jeremiah. "Whitehat website security statistics report." Retrieved March 8 (2007): 2010.

https://community.whitehatsec.com/assets/WPstats0808.pdf

[14] Halfond, W. G., Jeremy Viegas, and Alessandro Orso. "A classification of SQL-injection attacks and countermeasures." Proceedings of the IEEE International Symposium on Secure Software Engineering. Vol. 1. IEEE, 2006.

http://www.cc.gatech.edu/fac/Alex.Orso/papers/halfond.viegas.orso.ISSSE06.pdf

[15] Cross-site Scripting (XSS)

https://www.owasp.org/index.php/Cross-site_Scripting_(XSS)

[16] Third-Party Transaction

http://www.investopedia.com/terms/t/third-party-transaction.asp