National College of Ireland

MSc Web Technologies – 2010/2012

Dissertation

RDFa Can Help An Online Marketing Strategy in Terms of Search Engine Optimisation

Orla Butterly

x06350682

I hereby certify that this material, which I now submit for assessment of the programme of study leading to the award of Master of Science in Web Technologies is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work

Signed

Date

Student Number

Abstract

This document discusses the benefits of using RDFa and how it can help a companys' online marketing strategy The web is a big part of the world today and plays a big part in e-commerce and online marketing When searching for a particular company, product or service the end users/consumers tend to use a search engine such as Google or Yahoo to find what they are looking for With the large amount of data now available on the web it can be difficult for companies to get their information to the right people at the right time

In this study the technology RDFa is being used to enhance web document mark-up by making to semantically meaningful and machine-readable The majority of people in the web development industry haven't heard of this technology or are not implementing it Through the research I conducted, primary and secondary, I found that where companies did implement this technology they received an increase in their website traffic and their organic search results Even though the technology is easy to implement developers have either not become aware of the technology or are simply not implementing it because they don't fully understand how it works and what benefits there are in terms of Search Engine Optimization

Keywords RDFa, Semantic Web, Online Marketing, Search Engine Optimization

Table of Contents

1 Literature Review

1.1 Introduction

This dissertation is to research how marketing information can be presented in a semantically meaningful way so that it creates aids a business marketing strategy This review provides information from various journals, books, conferences and websites The research covers various topics such as Web 2 0, the move to the Semantic Web, The Semantic Web and Search Engine Optimization Businesses are aware that their presence on the web is vital and the following Literature Review highlights ways to improve on the exsiting type of online marketing From the research I have conducted the overall trend is that the web is becoming more semantically based and will continue to do so over the coming years

1.2 Web 2.0

Web 2 0 was brought to our attention approximately ten years ago and in 2004 had its first conference held by O'Reilly Media It provided a better user experience, more controlled data sources and a better scalability of services [13] Web 2 0 uses folksonomies to classify what content users are viewing by using open-ended tags or labels [14] Web 2 0 became known as the social web What we have seen here is many social connections being made via social networks to keep in touch with others, blogging to present your views in an online fashion or wikis where tools are provided for users to read and/or edit online content The Semantic Web shows a more refined and advanced construction of the web services available than Web 2 0 [15]

1.3 The Semantic Web

The Semantic Web was first introduced by Tim Berners-Lee in 2001 He states that this "new form of web" will allow computers to understand web content more meaningfully It will extend on to what we already know as Web 2 0 [16] Web 2 0 has its limits to what it can search and what information it can retrieve The Semantic Web will be able to provide the Internet with a better understanding of the content on the web [17] Meaning that the information will be deciphered by the search engines omre easily and they will provide the end user with more semantically correct information The Semantic Web still isn't entirely evident in today's Internet What we are seeing is what is also known as the Social Semantic Web This is by integrating semantic web technologies into the Social Web [18] Within the last few years it has been used in different areas of research [19]

The Semantic Web declares that access to information will be able through "machine-understandable representation of knowledge" [20] (Heß, Maaß, Dierick, 2011) Due to the large amount of data that is that is now readily available on the web, the traditional web isn't sufficient enough to go through it To access the data at a semantic level the machine needs to understand and organise the knowledge that is sought after in the right context [21] It give the Internet meaning and provide its users with meaningful information Even though the structure of the information may change, their meanings, also known as semantics, won't [22] Resource Description Framework (RDF) is a general-purpose language It is part of the Semantic Web technologies that are used to create a semantic web The development of this type of web uses ontologies [23] Jespen states in 2009 that "An ontology is a method of representing items of knowledge (ideas, facts, things—whatever) in a way that defines the relationships and classifications of concepts within a specified domain of knowledge " [25] Getting to "The Semantic Web" will take a vast amount time and effort Already there are some stepping stones placed in that direction Two of these are known as of Microformats and RDFa [24]

1.4 Web Marketing

Today, the vast majority of compames are expected to have some sort of web presence [26] The web has transformed how companies market themselves and how consumers hear about products and services It provides users with greater flexibility and a longer time to browse [3] How consumers use the web nowadays is very different to how it was originally used It is important to allow consumers to engage in market research and to supply the company with marketing content This is achievable over the web by using company blogs, wikis or web-based customer communities Consumer generated content on the web can be either a major help or hindrance to a companys service or product [31] Web marketing or e-marketing is very competitive and companies need to try get their information out there faster than others They also need to ensure that they are resenting the right information to the right people [32] Optimizing web marketing is a very important part of a company's campaign and needs to ensure the right type of advertising is used in the right way [33] Marketing needs to be measured in a way so that the information delivered is more timely and effective [34] There are four factors that companies need to consider when developing their marketing strategies the scope, the site, the synergy with physical processes and the system [35] To stay competitive and still make a profit in their market, companies marketing strategies are important The internet today can be accessed from almost anywhere using almost every electronic device from a PC or laptop to a mobile phone or MP3

player [36]

In marketing you need to understand who your customers are The user experience looks after the gap between the people, the technology and the business [37] "The ideal user experience for many routine tasks is the one that ultimately makes the tools disappear" (Spohrer, Strein, 2000) To enhance the user experience the information presented needs to be attentive to each individual user [38] Contextual advertising means showing advertising information in relation to what a user is surfing on the web [39] Marketing on the web has evolved from the conventional online advertising such as DoubleClick/AdSense to a better model, known as contextual advertising This type of advertising provides textual and multimodal relevance matching [40] It makes it easier to measure user interaction and their reactions with the product or service [41] It has evolved from this to a more semantically developed approach where the adverts and information become more relevant to the users and have a higher "exposure rate" [3] Using semantic technologies provides a more flexible framework for the existing advertising and optimises the online market [12]

1.5 Semantic Personalization for Online Marketing

"Web personalization can be described, as any action that makes the Web experience of a user personalized to the user's taste" (Mobasher, Cooley, Srivastava, 2000)[42] Personalizing web pages is a great way to develop a users interest by predicting their preferences [43] Due to the large expansion of information on the web it is necessary to focus attention on information personalization [44] Personalizing web pages allows for web users to have more useful and appropriate information returned to them by using their preferences It's effectiveness and efficiency help to improve the data that is presented [45] From a marketing point of view the each web page needs to be sure it supports the users references to obtain the correct information [46] Displaying personalized information can reduce a users browsing time and go with the information that is being recommended to them [47]

As it stands at the moment, when searching for certain information the Web does not understand the meaning behind the keywords and a lot of the time presents the wrong marketing information to the user [48] The usual way to find out users preferences is by using rated content or by user inputting their preferences This means of collecting information can be out of context or the information might be longer relevant as users and their preferences can change over time Semantically personalizing this data can eliminate these factors and present more accurate information [49] The Semantic Web technology will aid in this process by retrieving individual

information automatically by using RDF metadata and ontologies The technology allows for inference to retrieve personalized content [50] To advertise semantically each website or webpage needs to be appropriately reviewed to ensure its "true meaning" is found It increases the chance of that webpage being viewed if the advertisements are relevant to each user [51] The Semantic Web creates meaning behind the words in the advertisements It is beneficial for advertisers as it can ensure a more precise target audience and hopefully a better sales margin for the product or service [52]

1.6 Search Engine Optimization

A search engine is a tool on the internet that trawls the web to return results (webpages) related to the information that is inputted by the user Search Engine Optimization, or SEO, is about building these webpages in such a way that it improves their placement and visibility in search engine results The higher up the better, so long as it is relevant to the information that is being searched for [57] Every search engine has their own algorithm to crawl through databases to retrieve the webpages requested What the crawlers are looking for is optimized content which include appropriate keywords in URLs, titles and content, metatags, links to and from website, well-written HTML [58] Nowadays there is a large amount of websites and data on the internet and each site wants to make sure they are on top of the ranking or as close to as possible Companies are also beginning to change the way they present their information by changing their content type to a media source (audio or video) or increasing their already existing media Search Engines crawlers need to be altered to be able to understand this type of content [59] According to Jerri Ledford in 2009 from 2005 to 2010 there is an expected growth of 34% in online marketing, with nine out ten companies going to implement SEO into their websites [60] Before starting to optimize a site each search engine should be reviewed to see how they search for and display information It is essential that your website can be accessed by a web-crawler [75] Search Engine Optimization requires various parts to help websites achieve their ranking They need to decide on a specific target market, scope out the competitors and ensure keywords are relevant and not overused The following are two concepts of SEO

- On-page Optimization
- Off-page Optimization

On-page optimization refers to the optimization of the content that is on the page This includes factors such as using proper HTML elements, using meta descriptions and keywords, providing a sitemap and using internal linking

Off-page optimization is in relation to data about your site that is not directly in your page content Types of this are external linking and using anchor text [61]

1.7 Privacy and Trust Issues

Personalization can be taken the wrong way and seen as an invasion of a users privacy [53] There has been decrease in the extent privacy features on the web and users are wary about what information they share Due to this the W3 Consortium has issued what is called the P3P Project to enable a standard privacy practice on the web [54] Personalization and privacy go hand-in-hand while trying to advertise to users with privacy concerns being at the forefront of people's minds [55] "Personalization is the strategy of establishing an individualized approach to servicing customers through products and services and through your marketing strategy" Providing this interaction with web users a company can benefit from receiving a loyal group of consumers and be able to provide them with the information they want to enable the increase of sales (Virtual Advisor Interactive, 2009)[56] Trust is a major issue This can be seen not only in personalizing the web but also within the Semantic Web People are wary of the information they are given already With the development of the Semantic Web may be less tolerant of believing the information as the philosophy, of this web and the World Wide Web states "anyone can be an information provider or consume anyone else's information"[11] Trust is especially important in the area of e-commerce as money has to be transferred over the web It has been found that there are six factors that help reassure trust in the websites These are

- Brand – The companies reputation
- Seals of Approval – Certified security measures
- Navigation – Ease of use
- Fulfilment – Good user experience
- Presentation – Good design and presentation of content
- Technology – Good technologies [74]

Trust can be assigned based or credentials or on reputation With the Semantic Web, being able to write the rules and mark-up documents with actual meaning adds some weight to the data that is being presented [77] There is an ontology in the Semantic Web called The Trust Ontology This allows people, site authors, to show who else and what data on the web they trust [76]

1.8 Microformats and RDFa

As mentioned above Microformats and RDFa are two of the main technologies providing a

bridge between the existing Web and the Semantic Web Both of these takes on semantic mark-up are very similar and both have the same end goal

- Microformats

"Designed for humans first and machines second, microformats are a set of simple, open data formats built upon existing and widely adopted standards "(Dan Cedarholm, 2005) They are a way of adding semantic meaning to your HTML data It is used by web developers to present better, more meaningful information [68] They are like design patterns and rely on a pre-existing agreement between the author and the parser These agreements are known as vocabularies, e g hCard for contact information Once you are covering data supported by these vocabularies microformats are easy to use [69] Microformats is not a new language but a new technology that makes use of the features and elements that already exist in HTML It gives structure and meaning to HTML and is easy to implement and maintain [70] Microformats has been developed from Technoratis' developer community and is not standardized The developers decided that they would maintain the vocabulary of Microformats which means that it is not always sufficient enough for what is being described [62]

- RDFa

Resource Description Framework in Attributes, more commonly known as RDFa, is a specification that enables you to add structured meaning to the data in your XHTML/HTML document It is a Semantic Web technology that comes from the RDF language that enables RDF "triples" to be embedded into the attributes of XHTML and HTML The technology can find specific information in relation to what the user is searching for by creating more structured and meaningful code to make it machine-readable as well as human-readable [72] It also provides the ability to embed vocabularies into a document, either by building their own or use pre-defined ones [63] It easily allows vocabularies to be combined RDFa using prefix specified namespace [62] It is a World Wide Web Consortium recommendation and is now a working group [73] In terms of Search Engine Optimisation, the structured data, RDFa, can help bring up your search rankings which in turn can lead to a higher click-through-rate which from a marketing point of view is a big priority

Both Microformats and RDFa are supported by search engine Google and Yahoo! Google uses Rich Snippets and Yahoo! Uses Yahoo! Search Monkey

- Google Rich Snippets

Rich Snippets provide end users with extra information in the search results Google has an

algorithm that looks for RDFa and Microformats mark-up. [64]

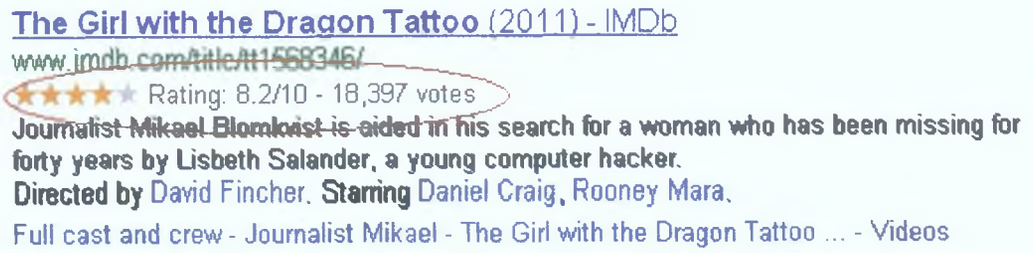


Figure 1: Google Rich Snippets showing a film rating.

- Yahoo! Search Monkey

In 2008 Yahoo! Started to extract RDFa data amongst its many other technologies, such as microformats, from webpages. Using Search Monkey it was able to provide the users with more structured information. [66] As of October 2010 Search Monkey was shut down but Yahoo! Still shows semantically marked-up information. [67]

1.9 Observation

Marketing your company or product online is a sure way to get your information out to existing and prospective clients. Most web users, unless looking for a specific site, will use a search engine to find the information they are looking for. Each site needs to make sure they are building websites that create a way to optimize their position in the search rankings. To do this they need to personalize their data to their target audience and mark-up their code appropriately. The Semantic Web is going to give a huge advantage to these businesses by providing them with new technology that will give meaning to their information. This meaning will allow machines to understand and in turn provide end users with more relevant, detailed and appropriate information. Of course with any new technology there are issues regarding security and trust issues regarding the informtaion being received. There are options in place to ensure that the data is credible. It may take time for people, developers and end users alike, to move in this direction but inevitably with the amount of data that exists on the web combined with the amount of people looking for this data it seems like the best way forward.

2 Research Background

2.1 Hypothesis

The research surrounding the hypothesis will be discussed throughout the document The hypothesis is as follows

"The technology RDFa is beneficial to companies in terms of increasing their presence on the web and improving their online marketing strategies"

2.2 Introduction

The idea behind this topic is to help enhance a companys position on the web and how easily it can be done? These days a lot of business is done online trying to increase their awareness and profit This is not only being achieved by large or technology driven businesses but also a vast amount of small companies regardless of sector have decided to put themselves on the World Wide Web in order to maintain a competitive advantage [1]

How a company promotes itself or its' product on the web is important Businesses need to receive a vast amount of numbers accessing their websites to ensure a return According to CRR research by Kelkoo as of 2010 10 7% of trade in the UK was from online retail [2] As a way to increase their chances they must try to heighten their "exposure rate" over the web [3]

What is coming to the forefront now is The Semantic Web At the moment there are various steps being undertaken to achieve the goal of reaching a fully semantic web The research documented here is going to exhibit one of these steps known as Resource Description Framework in Attributes (RDFa)

2.3 Scope

The main goal of this research is to show how semantic meaning in the web can improve the type of information that is presented to end users There a two systems that are being developed alongside this research is to show how RDFa can be implemented into a web application to enhance the mark-up and in turn improve online businesses One incorporates a sample of different types of

marketing that businesses use online, such as videos, blogs and images that is developed myself The second is a website named The Galway Roast, which allowed me to access their site and implement the technology to see what impact it would have Both systems are developed using RDFa and HTML or XHTML

2.4 Area of Proposed Contribution

The area that this dissertation is based on is Search Engine Optimization and the move to The Semantic Web The topic that this dissertation is focusing on is how RDFa can provide a semantically meaningful web and enhance a company's online profile RDFa is one of many stepping stones from Web 2 0 to The Semantic Web The hope is that it will create the links needed in the existing, and new, data to accomplish that step [4] In 2008 Tim Berners-Lee says "it's time to just go do it", that the potential is out there to reach the world of machine-readable data [5] Development in the semantic web is on-going and it will be years before the web as we know it will be defined as a Semantic Web At the moment the technology surrounding it is still seen as relatively new with many parts still under development There are several applications and websites out there that are semantically meaningful but they are not fully approved by the W3c standard This is due to the fact that as it stands at the moment some of the technologies are a W3c recommendation [6] and others are still a working draft [7] Although In spite of this development is still taking place

"*The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries* "
—World Wide Web Consortium, W3C Semantic Web Activity

RDFa can provide great benefits to the structure of your documents and data that are presented on the web [27] It enables you to place structured and meaningful data directly into your HTML file [28] It allows you to link data that couldn't be linked before in a way that is understandable to humans as well as machines At the moment RDFa validates with XHTML under the XHTML1 1+RDFa DTD It is a working draft for HTML4 and HTML5 and there are people developing in this format to show it can be done [29][30] RDFa is already being supported by several companies over the web such as Google's Rich Snippets, Yahoo! Search Monkey, Facebook'

Social Graph

"RDFa helps search engines add something called "rich snippets" to appropriate items within the ten sites that come up after a search The rich snippets are small visual cues to the person who sees the search results, increasing the click-through rate of that search result over the others, often even over those which are ranked higher on the page " [Alex Genadinik, http //www devx com/semantic/Article/45052/1954]

2.5 Research Methods

This study used several types of research to retrieve information about this topic

- Secondary Research

This includes reviewing existing journals, articles and conferences on this topic It also incorporates research into the technologies by using the World Wide Web Consortium and other online resources

- Case Studies

This research consists of secondary interviews and research conducted They delve into specific companies' stories about how and why they implemented this technology and what their outcome was

- Research Interviews

This research is primary research conducting by myself It involved face-to-face interviews, phone interviews and a survey All techniques used the same format of questions, (see Appendix B)

2.6 Case Studies – Secondary Research

2 6.1 Case Study 1[8]

RDFa

"RDFa The Inside Story of Best Buy, an Interview with Jay Myers"

Conducted by Doc Sheldon

Jay Myers is the Lead Development Engineer for Best Buy Doc Sheldon, also known as Sheldon Campbell is an SEO Content Strategy Specialist Jay implemented RDFa into existing webpages as an experiment They found that in less than a few months they began to see a larger amount of traffic to the website It showed an increase of 15% in their click-through rate and a 30% increase in their search results He says that he found RDFa a "much more stable concept" than either microformats or microdata Implementing the RDFa into the source code was a challenge as it requires a different mind-set to development As the experiment was a success they have begun implementing RDFa into other web developments RDFa is what he calls "a gateway drug" to the semantic web and would recommend for others to adopt this coding

For full interview see Appendix A

2.6.2 Case Study 2[9]

Semantic Advertising for Web 3 0

Conducted by Edward Thomas, Je Z Pan, Stuart Taylor, Yuan Ren, Nophadol Jekjantuk and Yuting Zhao, 2010

Case Study Product Blog

Advertising accounts for the main source of income for the World Wide Web Each advertiser only receives payment per click and not for merely providing the advertisement This case study outlines examples of the use of semantics and RDF in advertisements It uses a blog that publishes news and reviews on electronic items An example review taken from www pocket-lint com/reviews/review phtml/3526/nikon-dslr-D90-dslr-camera phtml, a Nikon camera, was used for the study The study also goes into detail about two potential advertisers of the product Taken into consideration is which RDFa annotation should be used When this is established the RDFa

and metadata is then embedded into the blog

The first potential advertiser is an Electronic Store It decides on the constraints needed for the placement of the advertisements, which are as follows

- the advert should only appear on reviews of the same product
- the review in question must be favourable review
- the price that the store sells the product should be at least 10% less than the price quoted in the review

RDFa is then used to add meaning to the descriptions of these products which are then added to the RDFs repository A SPARQL query will provide the information that was requested, i e the adverts that have only favourable reviews at the right price Ideally what is expected is that when this item is viewed the most suitable advertisement is displayed

The second potential advertiser is the competing manufacturer, Canon There constraints are the following

- Only advertise on products which compete with ours
- Only advertise where the same website carries a review of the competing product
- Only advertise where our product has a better review than the competing product

In this query the property "related" from SKOS (Simple Knowledge Organization System) is used to determine the constraints The related advert, in accordance with the constraints, is then retrieved and outputted

The case study concluded that this "vision" for semantic advertising provides a sample approach that can be implemented with damaging the underlying method Given that though, it is still felt that RDFa is not mainstream enough for companies to undertake implementing it There is more research to be done in this field to ensure the cost of using it will balance in what is returned

2.6 3 Case Study 3[10]

Establishing an Open, Digital Media Commerce Standard Using Semantic Web Technologies

Conducted by Manu Sporny, Dave Longley, Mike Johnson, and David I Lehn of Digital Bazaar, Inc , Blacksburg, Virginia, USA December 2008

In this case study Digital Bazaar create an application called Bitmunk to allow digital media be transacted by independent agents on the web The main goal is to provide a legal peer-to-peer network that effectively buys and sells digital media so that the creators can get their compensation After much deliberation Digital Bazaar decided to use Microformats as their standard due to the fact that there was less red tape and had a simple mark-up that worked with HTML They created vocabularies named hAudio Microformat and hAlbum Microformat to aid in the descriptions and ratings of the music It took only four months to complete these but they were met with a series of problems These included

1 Mixed interpretation of the process
2 Using Microformats became difficult due to lack of scope and namespaces
3 Unable to add their own visions because they weren't supported by hard publishing data
4 The vocabularies were unable to evolve because they were not supported by the Semantic Web

To help resolve these issues Digital Bazaar decided to use RDFa as it held the Semantic Web Requirements they needed They implemented the Bitmunk website using an Audio RDF Vocabulary and RDFa developed by a joint Task Force with W3C RDFa was more light-weight and better engineered than Microformats and it supported namespaces which was a downfall for Digital Bazaar with Microformats At the moment Digital Bazaar now runs over 1 3million websites using RDFa and is continuing to support W3C in the standardizing of the technology

2.7 Observation

Comparing these three case studies we can see a pattern Each company wanted to try something different to try to enhance their website optimization Even though there are other options available to them each company felt that when they were tried and tested they didn't meet

the right criteria for them They all landed on RDFa and all found it to be exactly what they were looking for

3 Architecture and Implementation

3.1 Introduction

The following section explains the design, the requirements and the implementation of the of the two primary case studies that I have conducted The main objective of the each study is to show the ease at which RDFa can be implemented into a XHTML or HTML document Primary Case Study One uses XHMTL + RDFa and shows how different aspects of the technology can be implemented and semantically understood Primary Case Study Two focuses on a website www thegalwayroast com and how RDFa was implemented into HTML It will show the impact the technology had on the website

3.2 Requirements

These requirements apply to both Primary Case Study One and Primary Case Study Two

3.2 1 Development Requirements

The application must incorporate Search Engine Optimization techniques It must incorporate good content, keywords and semantic metadata

The application must have clear, structured, meaningful mark-up The code must be written in a manner that is understandable to both humans and computers It must adhere to the structure set out in the World Wide Web Consortium recommendation for XHTML+RDFa1 1

3.2 2 Non-Functional Requirements

The non-functional requirements determine the qualities of the application and define how the system is supposed to be

- Extensibility

The application must be designed to ensure that new or modified functionality does not require a massive change to the infrastructure RDFa is designed for the Semantic Web and the application must be developed to be to grow and meet the requirement for the future

- Effectiveness

This means the application much provide a good performance in relation to the effort taken to produce the performance It will provide an increase in traffic coming to website

- Compliance

The code must be compliant with the XHTML+RDFa1 1 standard from the World Wide Web Consortium

- Semantic Interoperability

The XHTML documents must be able to be interpreted meaningfully and accurately by systems that are retrieving the information

3.3 Primary Case Study One

3.3 1 Overview

This section will discuss the details of each of the technologies used in the application

3.3.2 Technologies

- XHTML

XHTML is EXtensible HyperText Mark-up Language It takes HTML (HyperText Mark-up Language) and combines it with XML to create stricter version HTML that is defined as an XML application It consists of a series of strict rules which include, elements must be properly nested, always be closed, be in lowercase and each document must have only one root element

- RDFa

RDFa stands for Resource Description Framework-(in)-attributes It is part of the Semantic Web technologies and comes from the RDF language The 'in-attributes' means that the RDF can be easily embedded into XHTML or HTML documents It allows the web to understand more about web pages and applications by adding structured meaning to the documents It is used to improve their visibility and meaning to computers and the web It acts as a bridge between the documents and the data

RDF consists of triples and RDFa allows these to be embedded into the web documents These are embedded using vocabularies which provide defined lists of elements and terms that can be used within the mark-up The vocabularies are usually stated at the beginning of the web document in the <html> tag as an XML namespace Each vocabulary is given a prefix which is used to call a term or element from the vocabulary RDF triples are three-tiered statements that include a subject, a predicate and an object The subject and object are two resources in the statement and the predicate is the relationship between them

Using RDFa, these triples are embedded into webpages to make the existing code, such as HTML, to become machine-readable The RDF triples are extracted from the RDFa by a parser This then means that computers can understand the meaning behind the code and improve the web visibility of webpages and applications

- XHTML+RDFa

XHTML+RDFa is an extension of the XHTML mark-up language It is one of the processes being used to develop the Semantic Web The XHTML is embedded with the RDFa mark-up to make the original document semantically meaningful

- CSS

CSS stands for Cascading Style Sheets and they are used to develop the layout and design of the application It defines how HTML elements are displayed in a webpage They usually stored externally which means that they are a separate file to the HTML It gives you the ability to style an entire website using only one stylesheet

3.3.3 Design

The design of the application is straightforward It consists of two pages, the 'home' page and the 'about' page The 'home' page shows a product that is for sale and a promotional video The item is a Seagate Hard drive, which can be found at http //www seagate com/www/en-us/products/external/expansion/expansion_desktop/ and a Seagate promotional video which can be located at http //www youtube com/watch?v=CZ5OUODYDrQ The 'About' page feature a small paragraph of information and contact details

3.3.4 Implementation

3.3.4.1 Overview

This section will describe how the technology RDFa was implemented with XHTML

3.3.4.2 Implementation

XHTML is the main language for developing the application RDFa is implemented in the application by embedding the metadata into the XHTML document The metadata comes from vocabularies from the RDF language These vocabularies define sets of elements, also known as terms that can be used in an application The address of the vocabulary is defined within a XML namespace declaration These vocabularies are then given a prefix by the developer and are used throughout the document by "calling" the vocabulary through the prefix (See Example 1 below)

3.3.4.3 DublinCore

The Dublin Core is a set pre-defined metadata terms used for describing resources such as web pages, images and books It consists of fifteen elements The elements can be reused throughout the document These elements are

1 Title
2 Creator

3 Description
4 Publisher
5 Subject
6 Contributor
7 Date
8 Type
9 Format
10 Identifier
11 Source
12 Language
13 Relation
14 Coverage
15 Rights

```
<html xmlns dc="http //purl org/dc/elements/1 1/">
    <head></head>
    <body>
        <h1 property="dc title"><a href="#">Orla's Info</h1>
        <p property="dc creator">Orla Butterly</p>
    </body>
</html>
```

Example 1: Dublin Core vocabulary being called by using the prefix 'dc'

There are various examples of the use of these terms in the application The XML namespace for Dublin Core 'http //purl org/dc/elements/1 1/' and the prefix is 'dc'

3.3.4.4 FOAF

Friend Of A Friend (FOAF) vocabulary is designed to describe people and the links between them These links include things such as their documents, their photos, and their webpages It is defined as "a dictionary of terms"[http //xmlns com/foaf/spec/] that are known as classes or properties The terms can be grouped into three categories, Core, Social Web and Linked

Data Utilities. This application implements some terms from the 'Core' and 'Social Web' categories.

Figure 2: FOAF terms. Source: http://xmlns.com/foaf/spec/

From the 'Core' category the terms that are used are:

- Person
- Name
- Title
- img

From the 'Social Web' category the terms used are:

- mbox
- homepage

The XML namespace for this vocabulary is 'http //xmlns com/foaf/0 1/' The prefix used here is 'foaf'

3.3.4.5 GeoLocation

Latitudes and longitudes are the positions of places on a map GeoLocation is used to represent this data within a web document The properties associated with GeoLocation are geo lat for latitude, geo long for longitude, geo alt for altitude and geo lat_long to define latitude and longitude in the one content

Geo lat_long is the term used in the application The values are in decimal degrees and held inside the 'content' attribute The XML namespace is declared as 'http //www w3 org/2003/01/geo/wgs84_pos#'

3.3.4.6 Vcard

Vcard is a vocabulary used to represent a business card It holds information such as names, addresses, numbers and emails The XML namespace for this vocabulary is 'http //www w3 org/2006/vcard/ns#' 'Vcard' is used as the prefix for these terms The vocabulary is implemented by adding the vcard mark-up to information about a business to a website

3.3.4.7 GoodRelations

This vocabulary is known as the "The Vocabulary for E-Commerce" [http //www heppnetz de/projects/goodrelations/] It is used to display information about a company, its products and its services The XML namespace is declared as 'http //purl org/goodrelations/v1#' The prefix used here is 'gr' GoodRelations has a large amount of elements that can be used for various different industries The most common elements that are used in this vocabulary are for a company, a shop or restaurant and an offer/product Some of the terms used in the application are for reviews, pricing and description

3.3.4.8 Review

A review is a judgement about a product or a service by a critic or a consumer

who had an interaction with it The review vocabulary is used to add meaning to reviews and ratings The XML namespace for review is 'http //purl org/stuff/rev#' The prefix for this is defined as 'rev' The vocabulary has sixteen terms associated with writing a review These include

1 Comment
2 Review
3 Feedback
4 commenter
5 has Review
6 hasComment
7 hasFeedback
8 max Rating
9 min Rating
10 positiveVotes
11 rating
12 reviewer
13 text
14 title
15 totalVotes
16 type

3.3.4.9 Media

At present Google recognises two types of video mark-up, 'Facebook Share' and 'Yahoo! Searchmonkey' For the purpose of this application the media mark-up being used is the Yahoo! Searchmonkey format The XML namespace for this is 'http //search yahoo com/searchmonkey/media/' and the prefix is 'media' The purpose of this mark-up is to help increase the chances of your video receiving better web visibility

<object width="500" height="250" rel="media video" resource="http //www youtube com/watch?v=CZ5OUODYDrQ">
<param name="movie" value="http //www youtube com/watch?

```
v=CZ5OUODYDrQ"></param>
                    <embed src="http //www youtube com/watch?v=CZ5OUODYDrQ"
type="application/x-shockwave-flash" width="500" height="250">
                    <a rel="media Thumbnail" href="http //www youtube com/watch?
v=CZ5OUODYDrQ"></a>
                    <a rel="dc license" href="http //www youtube com/t/terms"></a>
                    <span property="dc description" content="Seagate Advertisement"/>
                    <span property="media title" content="Seagate channel 10
AD wmv">Title  Seagate Channel </span>
                    <span property="media type" content="application/x-shockwave-
flash"></span>
                    <span property="media region" content="*"></span>
                    <span property="media duration" content="6"></span>
                </object>
            </div>
```

Example 2 Yahoo! Searchmonkey mark-up for video

3.4 Primary Case Study Two: The Galway Roast.

3.4.1 Overview

Digital Eire is one of the companys I approached during my surveys It is a successful digital marketing company in the midlands specialising in web design, graphic design and online marketing strategies I met with the owner Ray Carolan and the director of marketing Michael Moran (*see 5 3 4, Interview Four*) Ray and Michael were both very interested in the technology RDFa and my research, that they offered me the opportunity to use one of there websites to implement the technology After consultation with the website owner it was agreed that I could use the website "www thegalwayroast com" For a week at the end of February 2012 they gave me a weeks work and full access to The Galway Roast

3.4.2 Technology

The Galway Roast is a Wordpress powered website Wordpress is a content management system that uses PHP, HTML and MySQL In order to integrate RDFa into this site I needed to activate a plugin A plugin is a piece of software that can be added to a site to enable more functionality [79] The plugin used with The Galway Roast is called RDFaCE It is a content editor for RDFa based on the Javascript WYSIWYG* TinyMCE editor The plugin allows the developer or content manager to annotate and edit Semantic content within Wordpress The following technologies are what Wordpress uses to build their systems

- HTML

HTML stands for HyperText Mark-up Language It is defined by a set of mark-up tags that are used to display the content of a webpage These tags also define how a webpage should be structured and layed out e g <head>Defines the header of a page </head> , <title> Defines the title of a page </title> , and <body> Defines the main content of a page </body>

* What You See Is What You Get (WYSIWYG) is the term for an editor or program that allows a developer to see what the user interface looks like as it is being developed.[78]

- PHP

Hypertext Preprocessor (PHP) is a server-side scripting language. PHP is used to process any HTML on the server before the content is displayed on a webpage. In Wordpress PHP is mainly used in functions and plugins. In order for PHP to be read and deciphered from the HTML correctly the code must be within this tag, <?php ...Enter Code... ?>.

- MySQL

MySQL is a database that stores its data in objects known as tables. For Wordpress MySQL stores information such posts, terms, users, links and comments. See Figure 8 for an image of a database structure in Wordpress. Image taken from The Galway Roast database.

Figure 8: Image of the structure of a MySQL database in Wordpress.

3.4.3 Implementation

The main request from Digital Eire and The Galway Roast was to help optimise the search for "Coffee Beans" and "Coffee Roasted Beans". Once the plugin RDFaCE was activated I could then procede with adding the RDFa to the HTML code within the appropriate pages. The pages included are listed below along with their URLS:

- Home – http://www.thegalwayroast.com
- Our Coffee, The Galway Roast – http://www.thegalwayroast.com/our-coffee/the-galway-roast
- Our Coffee, Peruvian/Costa Rican – page is no longer public
- Our Story, About Us - http://www.thegalwayroast.com/our-story/about-us
- Our Story, Testimonials – http://www.thegalwayroast.com/our-story/testimonials
- Roastery – http://www.thegalwayroast.com/the-roastery
- Cafe – http://www.thegalwayroast.com/cafe

The plugin RDFa already has built in URI's such as http://xmlns.com/foaf/0.1/ (FOAF, *see: 3.3.4.4*) and http://www.iptc.org/std/rNews/1.0/" (rNews). In this case, FOAF was not implemented into the final code as the plugin was able to determine that it wasn't relevant. The vocabulary rNews was used automatically by the plugin and was placed within the HTML code. It acknowledged where there were names of Persons and Places giving it semantic meaning. Figures 6 & 7 show the RDFa annotation as a visual in Wordpress.

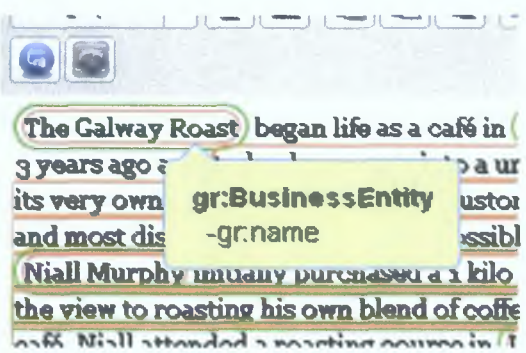


Figure 6: Visual showing how GoodRelations is annotated

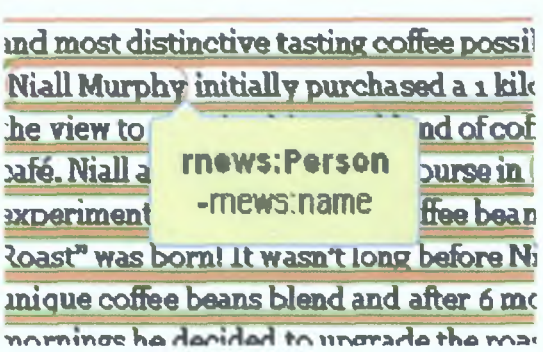


Figure 7: Visual of how rNews is annotated.

3.4.3.1 RDFaCE

The plugin RDFace was developed by Ali Khalili and Dr Sören Auer It is a content editor for RDFa for the TinyMCE editor It allows for semantic annotation and mark-up within the TinyMCE editor in Wordpress It provides API suggestions to ensure correct URIs are used

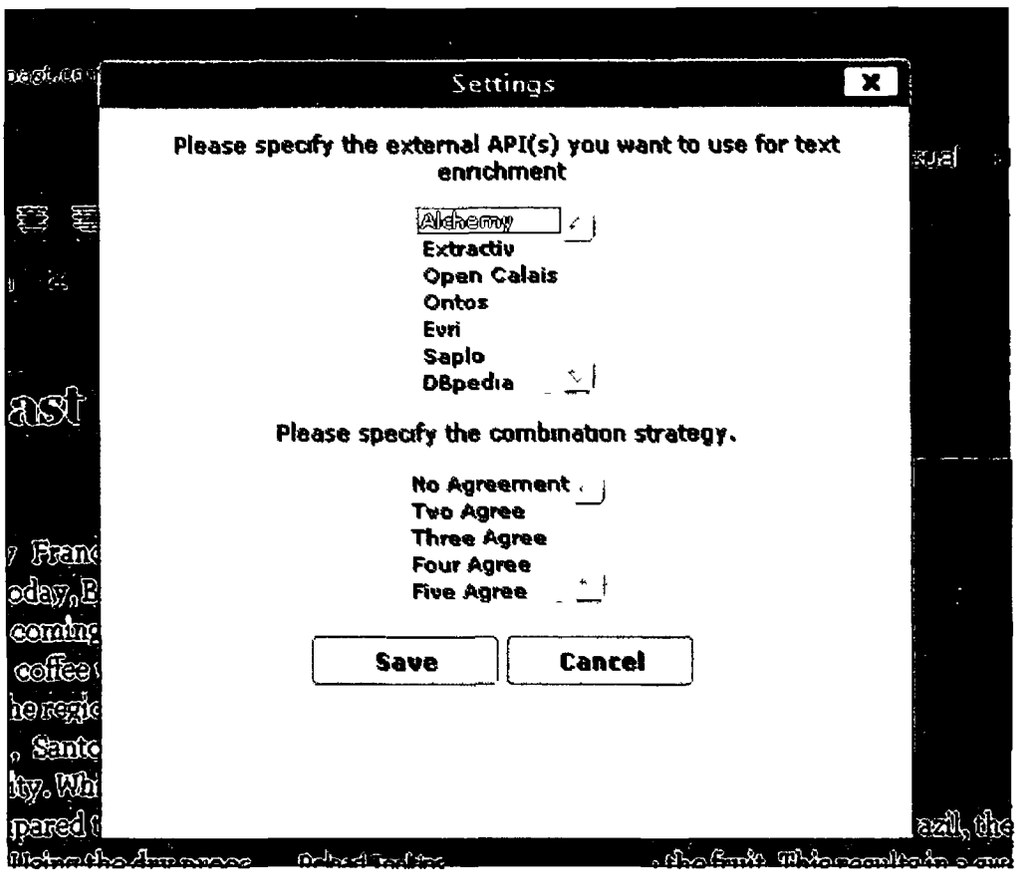


Figure 13: API suggestions for RDFaCE

3.4.3.2 rNews

The vocabulary rNews was automatically embedded into the HTML of the The Galway Roast RNews is used in the publishing industry to embed machine-readable data into web documents It is implemented in two formats, within articles and media All the subject and object classes are listed in the table below along with the definition for each.

Subject Class	Verb	Object Class	Definition
NewsItem	about	Concept	Indicates that the NewsItem is specifically about a concept.
NewsItem	accountablePerson	Person	Specifies the person that is legally accountable for the NewsItem.
NewsItem	comment	UserComments	Comments, typically from users, on this NewsItem.
NewsItem	contributor	Person | Organization	A secondary contributor to the NewsItem.
NewsItem	copyrightHolder	Person | Organization	The party holding the legal copyright to the NewsItem.
NewsItem	creator	Person | Organization	The author of the NewsItem.
NewsItem	editor	Person | Organization	Specifies the person or organization who edited the NewsItem.
NewsItem	mentions	Concept	Indicates that the NewsItem contains a reference to, but is not necessarily about a concept.
NewsItem	provider	Person | Organization	Specifies the person or organization that distributed the NewsItem.
NewsItem	sourceOrganization	Organization	The organization on whose behalf the creator of the NewsItem was working.
Article	associatedMedia	"ImageObject | AudioObject | VideoObject"	An image or audio or video object associated with an Article.
ImageObject	associatedArticle	Article	An Article associated with the Media Object
ImageObject	associatedMedia	"ImageObject | AudioObject | VideoObject"	An image or audio or video object associated with an Article.
AudioObject	associatedArticle	Article	An Article associated with the Media Object
AudioObject	associatedMedia	"ImageObject | AudioObject | VideoObject"	An image or audio or video object associated with an Article.
VideoObject	associatedArticle	Article	An Article associated with the Media Object
VideoObject	associatedMedia	"ImageObject | AudioObject | VideoObject"	An image or audio or video object associated with an Article.
UserComments	creator	Person	The creator of the comment.
UserComments	discusses	NewsItem	Specifies the NewsItem associated with this Comment.
Place	address	PostalAddress	A real-world postal address associated with this entity.
Place	geoCoordinates	GeoCoordinates	The geo coordinates of the location.
Person	address	PostalAddress	A real-world postal address associated with this entity.
Organization	address	PostalAddress	A real-world postal address associated with this entity.

Figure 9: Table of metadata for the vocbulary rNews. Source:http://dev.iptc.org/rNews-10-Introduction-to-rNews.

In relation to The Galway Roast rNews was able to define the people involved in the company, the locations of the coffee shops and the places where the coffee is sourced. This works with rNews data model (see: Figure 10)as it relates information about the organization and coincided with the information I was implementing using the vocabualry GoodRelations.

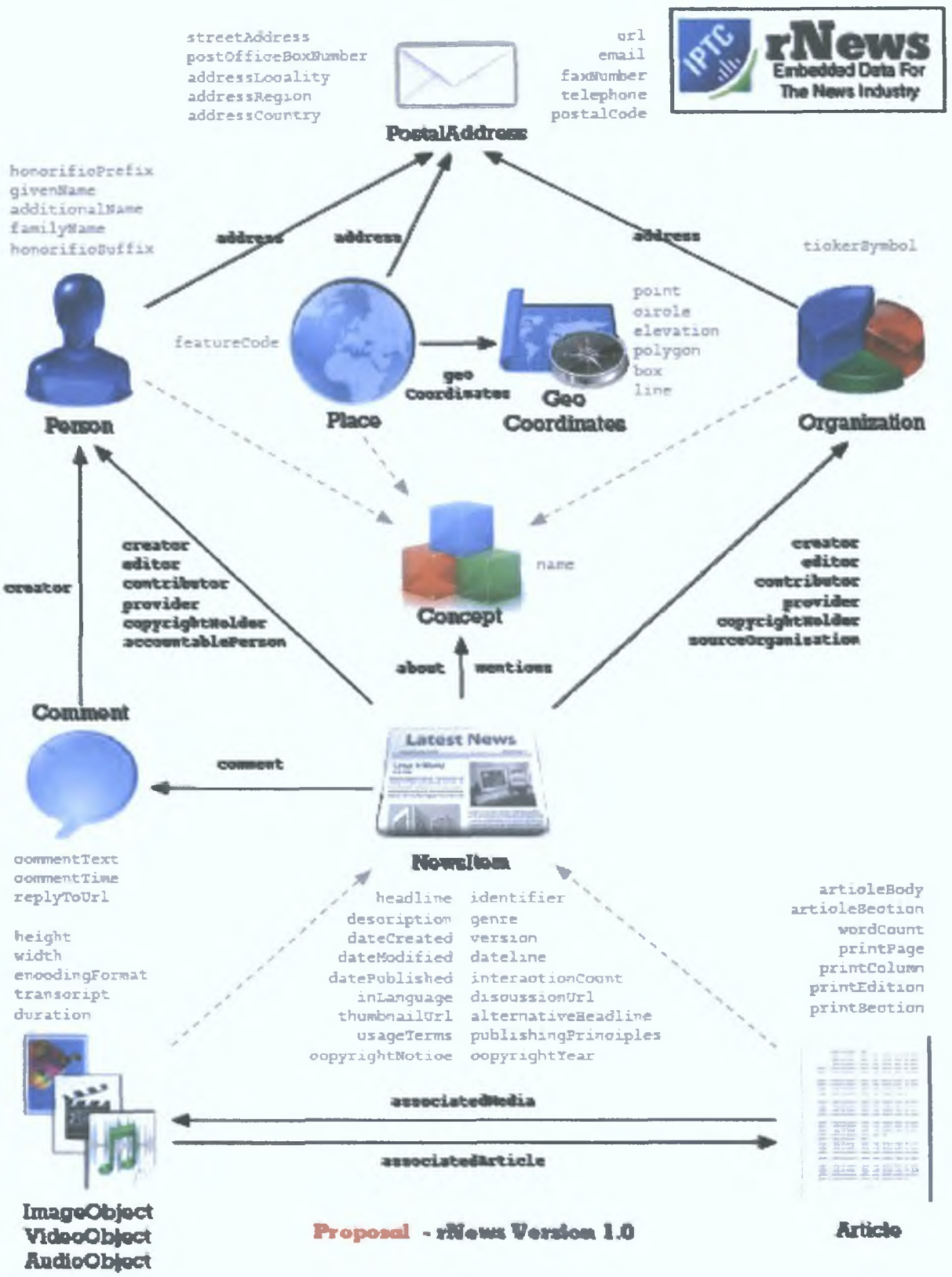


Figure 10: rNews data-model. Source: http://blog.semantic-web.at/wp-content/uploads/2011/10/rnews-data-model.jpg.

The vocabularies that I implemented in this website are GoodRelations (*see: 3.3.4.7*) and Review (*see: 3.3.4.8*). I applied the use of the vocabulary GoodRelations to add extra meaning to the name and the description of the website and products. The Review vocabulary was used to associate the words on the webpage as a Review. All pages with the exception of Testimonials was implemented with GoodRelations. The vocabulary Review was used within the page Testimonials. In the image below you will see a screenshot of the Wordpress editor view "About Us". It uses the metadata terms gr:BusinessEntity, gr:Name, gr:Brand and gr:Description. Using these terms I was able to add extra metadata to each page. This consisted in ensuring that the terms outlined by The Galway Roast and Ray Carloan were used, which were "Coffee", "Coffee Beans" and "Roasted Coffee Beans".

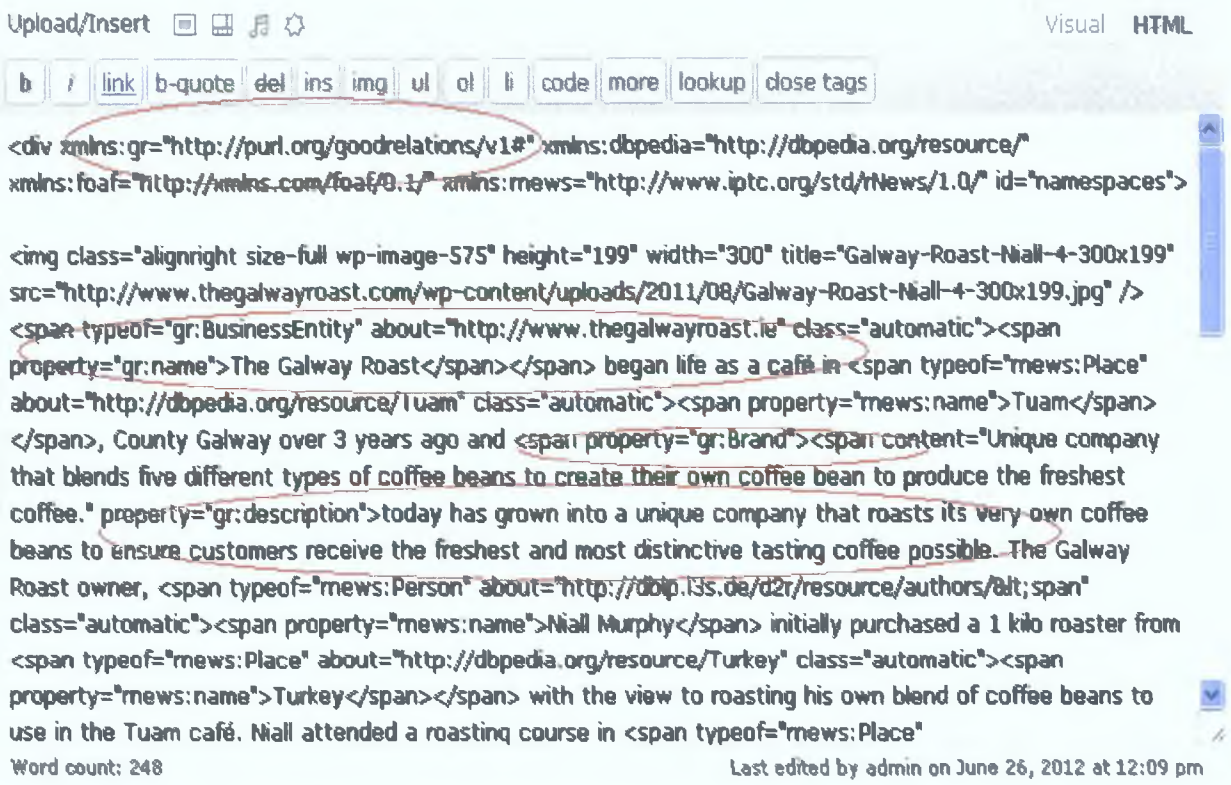


Figure 11: Wordpress editor view of the "About Us" page.

4 Testing

4.1 Overview

This section outlines the testing procedures of the application and the RDFa mark-up

4.2 Methodologies for Testing

4.2.1 Testing Methodolgy - Primary Case Study Two

The only way to test if RDFa does enhance your search engine visibility is to monitor the analytics of the webpage and growth in the company As this was previously built website, the original developers had already created a Google Analytics account to monitor the the sites taffic

Google Analytics provides information about the traffic to your website and its' marketing effectiveness Each site will have the Google Analytics code embedded into the <head> tag of the pages It can provide information regarding

- Site visits – which is broken down into
  - Unique site visits
  - Page views
  - Ratio of pages viewed per visit
  - Average time spent on the site
  - Percentage of Bounce Rate
  - Percentage of New Visits
- Demographics – which is broken down into
  - Location
  - Language
- Technology – which is broken down into
  - Browser
  - Operating System
  - Service Provider
  - Mobile Operating System
  - Mobile Service Provider

- Mobile Screen Resolution

The analytics will be monitored over a period three months and an evaluation of its performance over that period of time will be recorded The Rich Snippets will also be tested using Googles Rich Snippets Testing Tool from Googles webmasters tools (http //www google com/webmasters/tools/richsnippets) This will show whether the RDFa can be interpreted as rich snippets

4.2 2 Testing Methodology for RDFa Testing - Primary Case Study One

This method is used to help ensure that the RDFa syntax being used is correct and is based on the Primary Case Study One RDFa is easy to implement into any web document To test the validity of the RDFa it must be extracted by an RDF parser To do this W3C have created an extractor and validation service These are known as the 'W3C RDFa Distiller and Parser Service' and the 'W3c RDF Validation Service' These services can be found 'http //www w3 org/2007/08/pyRdfa/' and 'http //www w3 org/RDF/Validator/' respectively The procedure has a series of steps, which are as follows

1. Open the W3C RDFa distiller service
    1. Use the 'distill by file' upload option and click Go!
    2. Open the extracted file in a web browser
    3. View and copy the page source
2. Open the W3C RDF validation service
    1. Paste the copied page source into the text box under 'Direct Input'
    2. Select triples and graphs
    3. Click Parse RDF

The extracted RDF can be viewed in three different formats, either as a plain triple, a graph representation of the triple or an XML document (See figures 3, 4 and 5 below)

Home.html

Number	Subject	Predicate	Object
1	http://www.w3.org/RDF/Validator/run/1326661812418#productSeagate	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://purl.org/goodrelations/v1#Product
2	genid:A12872	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://purl.org/goodrelations/v1#UnitPriceSpecification
3	http://www.w3.org/RDF/Validator/run/1326661812418#productSeagate	http://purl.org/goodrelations/v1#hasPriceSpecification	genid:A12872
4	genid:A12872	http://purl.org/goodrelations/v1#hasCurrency	"EUR"@en
5	genid:A12872	http://purl.org/goodrelations/v1#hasCurrencyValue	"100"@en
6	http://www.w3.org/RDF/Validator/run/1326661812418#productSeagate	http://purl.org/goodrelations/v1#hasBusinessFunction	http://purl.org/goodrelations/v1#Sell
7	http://www.w3.org/RDF/Validator/run/1326661812418#productSeagate	http://purl.org/goodrelations/v1#hasEAN_UCC-13	"013803123784"@en
8	genid:A12873	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://purl.org/stuff/rev#hasReview
9	http://www.w3.org/RDF/Validator/run/1326661812418#productSeagate	http://purl.org/stuff/rev#Review	genid:A12873
10	genid:A12873	http://purl.org/stuff/rev#comment	"This is a good harddrive"@en
11	genid:A12873	http://purl.org/stuff/rev#rating	"2"@en
12	genid:A12873	http://purl.org/stuff/rev#reviewer	"Kim Lynam."@en
13	genid:A12873	http://purl.org/stuff/rev#count	"5"@en
14	http://www.w3.org/RDF/Validator/run/1326661812418#productSeagate	http://purl.org/goodrelations/v1#Description	"Seagate Expansion HardDrive 500BG"@en

15	http://www.w3.org/RDF/Validator/run/1326661812418#productSeagate	http://purl.org/goodrelations/v1#Name	"HardDrive"@en
16	http://www.w3.org/RDF/Validator/run/1326661812418#productSeagate	http://xmlns.com/foaf/0.1/depiction	http://www.w3.org/RDF/Validator/run/images/seagate.png
17	http://www.w3.org/RDF/Validator/run/1326661812418#video	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://search.yahoo.com/searchmonkey/media/video
18	http://www.w3.org/RDF/Validator/run/1326661812418#video	http://search.yahoo.com/searchmonkey/media/video	http://www.youtube.com/watch?v=CZ5OUODYDrQ
19	http://www.youtube.com/watch?v=CZ5OUODYDrQ	http://search.yahoo.com/searchmonkey/media/region	"*"@en
20	http://www.youtube.com/watch?v=CZ5OUODYDrQ	http://purl.org/dc/elements/1.1/description	"Seagate Advertisement"@en
21	http://www.youtube.com/watch?v=CZ5OUODYDrQ	http://purl.org/dc/elements/1.1/license	http://www.youtube.com/t/terms
22	http://www.youtube.com/watch?v=CZ5OUODYDrQ	http://search.yahoo.com/searchmonkey/media/duration	"6"@en
23	http://www.youtube.com/watch?v=CZ5OUODYDrQ	http://search.yahoo.com/searchmonkey/media/title	"Seagate channel 10 AD.wmv"@en
24	http://www.youtube.com/watch?v=CZ5OUODYDrQ	http://search.yahoo.com/searchmonkey/media/type	"application/x-shockwave-flash"@en
25	http://www.youtube.com/watch?v=CZ5OUODYDrQ	http://search.yahoo.com/searchmonkey/media/Thumbnail	http://www.youtube.com/watch?v=CZ5OUODYDrQ

Figure 3: RDF Triple for Home.html

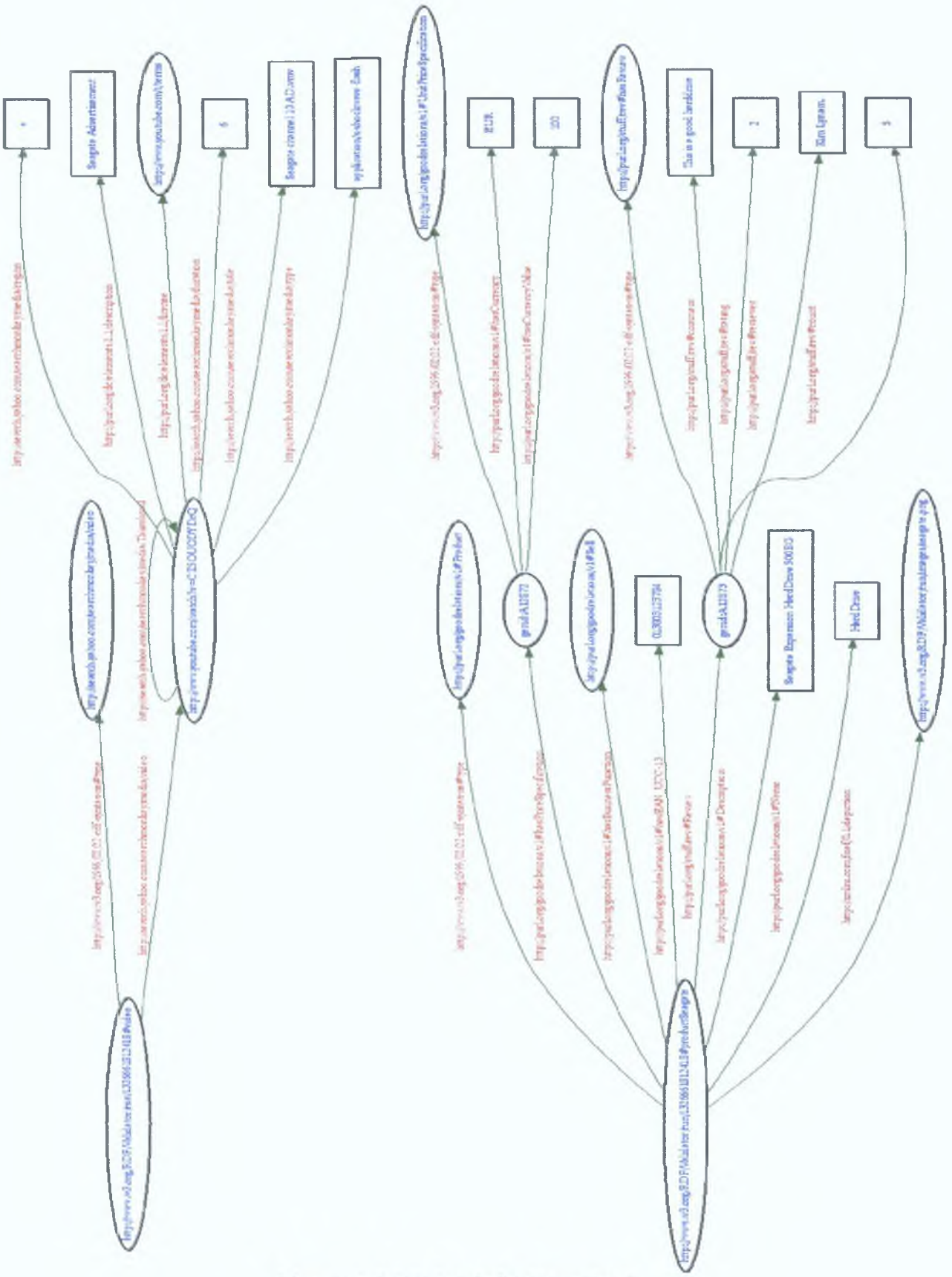


Figure 4: RDF Graph for Home.html

<rdf RDF xmlns dc="http //purl org/dc/elements/1 1/" xmlns foaf="http //xmlns com/foaf/0 1/"
xmlns gr="http //purl org/goodrelations/v1#"
xmlns media="http //search yahoo com/searchmonkey/media/"
xmlns rdf="http //www w3 org/1999/02/22-rdf-syntax-ns#"
xmlns rdfs="http //www w3 org/2000/01/rdf-schema#"
xmlns rdfs1="http //www w3 org/2000/01/rdfs-schema#"
xmlns rdfs2="http //www w3 org/2000/01/rdfs-schema#" xmlns rev="http //purl org/stuff/rev#"
xmlns vcard="http //www w3 org/2006/vcard/ns#"
xmlns xhv="http //www w3 org/1999/xhtml/vocab#">
<gr Product rdf about="#productSeagate">
<gr hasPriceSpecification>
<gr UnitPriceSpecification>
<gr hasCurrency xml lang="en">EUR</gr hasCurrency>
<gr hasCurrencyValue xml lang="en">100</gr hasCurrencyValue>
</gr UnitPriceSpecification>
</gr hasPriceSpecification>
<gr hasBusinessFunction resource="http //purl org/goodrelations/v1#Sell"/>
<gr hasEAN_UCC-13 xml lang="en">013803123784</gr hasEAN_UCC-13>
<rev Review>
<rev hasReview>
<rev comment xml lang="en">This is a good harddrive</rev comment>
<rev rating xml lang="en">2</rev rating>
<rev reviewer xml lang="en">Kim Lynam </rev reviewer>
<rev count xml lang="en">5</rev count>
</rev hasReview>
</rev Review>
<gr Description xml lang="en">Seagate Expansion HardDrive 500BG</gr Description>
<gr Name xml lang="en">HardDrive</gr Name>
<foaf depiction rdf resource="images/seagate png"/>
</gr Product>
<media video rdf about="#video">
<media video>

<rdf Description rdf about="http //www youtube com/watch?v=CZ5OUODYDrQ">
<media region xml lang="en">*</media region>
<dc description xml lang="en">Seagate Advertisement</dc description>
<dc license rdf resource="http //www youtube com/t/terms"/>
<media duration xml lang="en">6</media duration>
<media title xml lang="en">Seagate channel 10 AD wmv</media title>
<media type xml lang="en">application/x-shockwave-flash</media type>
<media Thumbnail rdf resource="http //www youtube com/watch?v=CZ5OUODYDrQ"/>
</rdf Description>
</media video>
</media video>
<rdf Description rdf about="uploaded ">
<dc title xml lang="en">The Tech Shop</dc title>
<dc creator xml lang="en">Orla Butterly</dc creator>
<xhv stylesheet rdf resource="main css"/>
</rdf Description>
</rdf RDF>

Figure 5: XML/RDF Document

5 Evaluation

The following section will go through all the information and provide a conclusion Information was collected from various types of research These included face-to-face interviews, phone interviews, case studies and articles regarding the technology

5.1 Secondary research

The secondary research I conducted consisted of reading articles and journals relevant to the topic It also included reading case studies about companies who had experience in using the technology in their websites The research showed that RDFa provided several benefits to the businesses, namely an increase to their organic search results and an increase in their click-through rate Even though there are other options available, such as mircoformats, RDFa was found to be easy to use and implement It also provides a choice to develop your own ontology or vocabulary whereas mircoformats only allow you to use what has already been defined Companies such as Best Buy and Digital Bazaar both gained from implementing RDFa

Best Buy embedded the technology within an existing site The results showed that when people were presented the information with the rich snippets it increased the traffic to the website by 15% They said the technology wasn't too difficult to implement and has less constraints than other options they had considered

Digital Bazaar took a different approach and actually worked with the World Wide Web Consortium to create a vocabulary that suited their needs They started off working with microformats but because of the constraints in the mark-up they couldn't implement some of the features they wanted For them RDFa was the best option because they got to incorporate what they needed It was more light-weight and better engineered Their application was implemented using the base of the Microformats already in use and incorporate it with RDFa

5.2 Primary Research

The primary research I undertook included conducting interviews and implementing the code using two seperate studies

The first study Primary Case Study One was conducted on the original submission of this document Its research is based on implementing RDFa into a XHTML file using sample content for the purposes of testing This was to see if it was easy to embed RDFa into XHTML and be able to be understood by an RDFa distiller and parser, in turn meaning that it can be understood by a search engine/web crawler As can be seen from figure 4 the RDFa embedded within the Home html was fully understood by the W3C Distiller and Parser used to validate this information

The second study I performed began in February 2012, Primary Case Study Two – The Galway Roast I implemented RDFa into an existing website that is powered by Wordpress and the sites performance was monitored The Galway Roast made some significant improvements in the few months after the technology was implemented Ray Carolan, Owner of Digital Eire, Developer of The Galway Roast said that the technology RDFa helped the site in terms of its search engine optimisation "*Yes the RDFa definitely made a differance The Galway Roast moved up to page one in 90% of its search terms, with Coffee Roasted beans and Coffee beans performing the best*" - Ray Carolan

Googles Rich Snippets Testing Tool was used to demonstrate whether or not the search engine to extract the rich snippets The results from three pages "About Us", "The Galway Roast" and "Testimonials", screenshots taken from the testing tool, can be seen below

- About Us Page

Figure 13: About Us Page – You can see the information it has extracted.

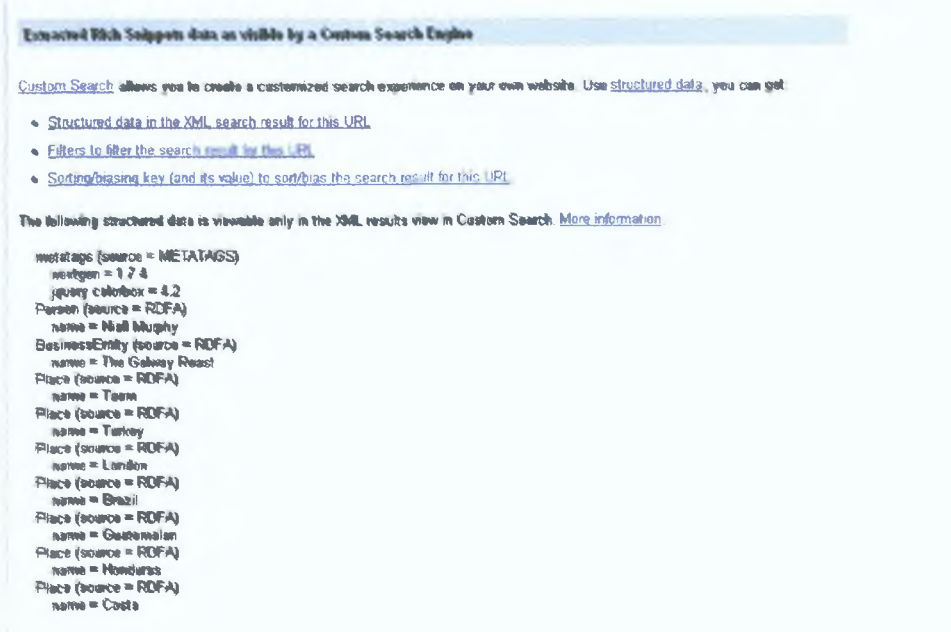


Figure 14: About Us Page – You can see the XML results view.

- The Galway Roast

Extracted Author/Publisher for this page

author
linked author profile = http://twitter.com/?_escaped_fragment_=/galwayroast

Extracted rich snippet data from the page

Offering
name = Brazilian Coffee

Person
name = Franciscode Mello

Offering
name = Guatemala Coffee

Offering
name = Honduras Coffee

BusinessEntity
legalName = The Galway Roast

description = Coffee beans from Brazil are processed in two ways which give the coffee a a fruitier, cleaner coffee
description = There are seven varieties of Coffee beans from Guatemala
description = Honduras coffee beans have an unremarkable quality

Place
name = Cayenne

Place
name = French Guiana

Place
name = San Paulo

Place
name = Santos

Figure 15: The Galway Roast – Here is the Rich Snippet data that was extracted.

Figure 16: The Galway Roast – XML view of the structured data.

As you can see from Figures 13-16, Google was able to deciphere the RDFa that was placed in the HTML code and could extract the rich snippets from it. It was able to define The Galway Roast as a business, the products that are offered, any description available and the place names.

- Testimonials

Figure 17: Testimonials – Here you can see the exrtacted Rich Snippets.

The page Testimonials unfortunately didn't work out as expected As you can see from figure 17, it extracted some of the rich snippets correctly, the Reviewer(s) but only one Review description After some investigation as to why this happend I found it to be an issue with how the code was implemented and can easily be rectified by changing the placement of certain metadata

The Google Analytics also showed an increase of new visitors to the site in the time after the RDFa was implemented The performance can be seen below in a download directly from The Galway Roast Google Analytics account From the analytics tool we were able to compare two seperate time periods to see the differences The two time periods are from December 2nd 2011 – March 3rd 2012 and March 4th 2012 – June 4th 2012 You can see from the graph that there was a large increase in the number of visits between the two periods, an increase of 377 27% That is 105 visits in the period March – June 2012 verses 22 in the period December 2011 – March 2012 The number of new visits was also up by 27 33%

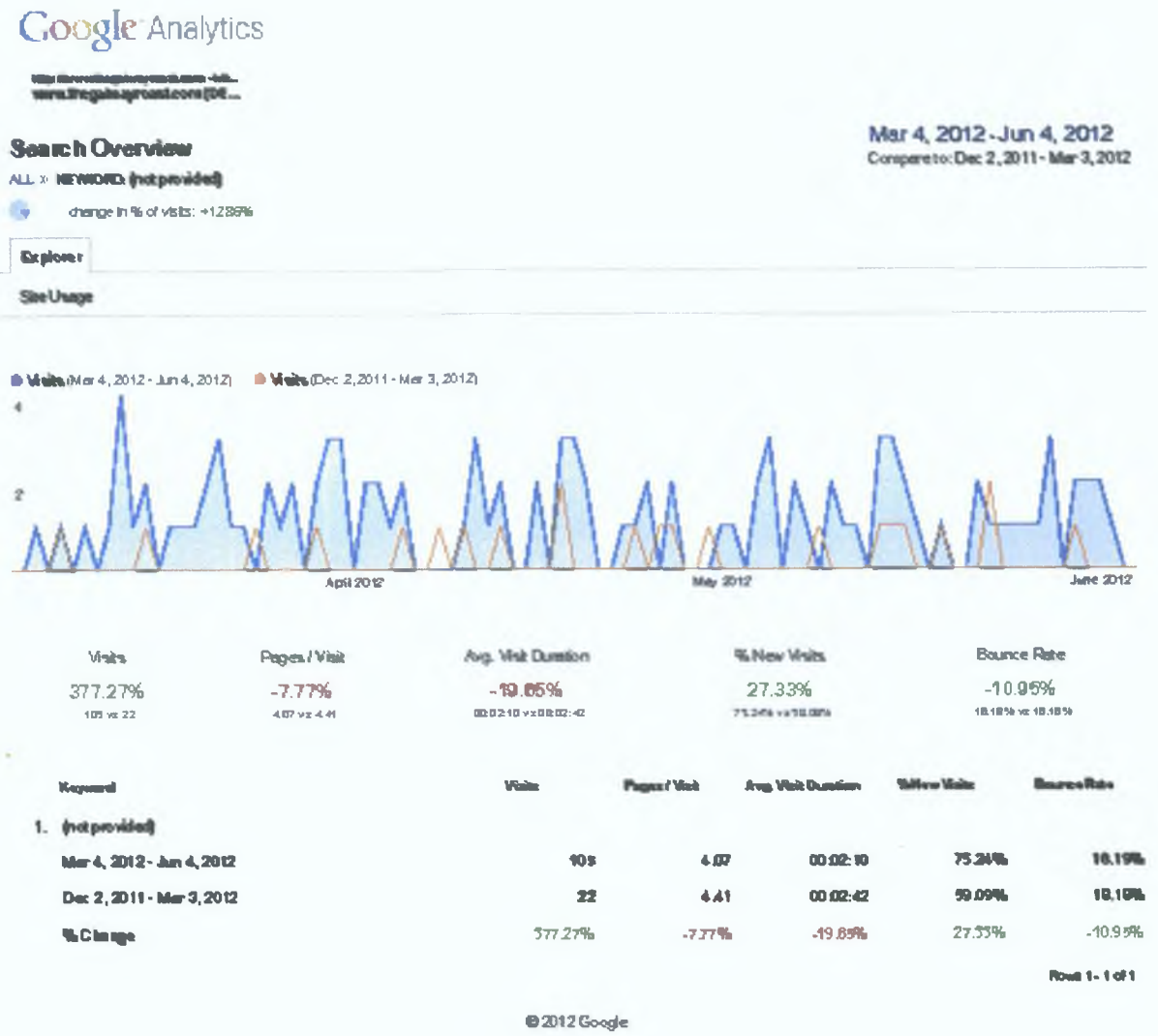


Figure 12: Google Analytics for The Galway Roast.

5.3 Research Interviews

I conducted a number of interviews during this study. I found that the majority of people I spoke with had not heard of the technology or simply were not implementing it as they did not know much about how it worked or its benefits. In my own implementation of the technology, once it was researched properly that I found it is easy to implement and understand. The interviews are shown below:

The following section details some of the feedback from the interviews that were conducted. They include information on the services offered, technologies used and their knowledge of and views on RDFa.

5 3.1 Interview One

Company Croan ie – Conducted over the phone

Services include

Web Design, Web Development, Search Engine Optimization

Technologies they use

HTML, PHP

SEO Techniques

Good Content, Inbound Links

Croan ie had never heard of RDFa After the technology was explained to them they stated that if it does have a definite impact in SEO they would strongly consider using it

5.3.2 Interview Two

Company Web Page Design Company, http //www webpagedesign ie – Conducted over the phone

Services Include

Web Design, Web Hosting

Technologies they use

HTML, CSS, PHP, MySql, Ajax, Adobe All code is OpenSource

SEO Techniques

Using defined keywords Also use Google adwords

Web Page Design Company had heard of RDFa but don't use it and know very little about it They are hesitant to use it because they are not sure how the web is going to take it They are afraid that it won't be picked up by anyone else except Google and don't want to take that risk

5 3.3 Interview Three

Company Stewart Curry, Web Developer

Services Include

Web Development, mainly front-end design

Technology used

HTML, PHP, CSS, JavaScript

SEO Techniques

Semantic mark-up, good content, og data for content sharing

Stewart Curry didn't know the name of the technology RDFa but was aware of some of its elements He has used the Dublin Core vocabulary for accessibility purposes and association between pages He did not conduct any analytics

5.3 4 Interview Four.

Company Digital Eire, Ray Carolan Owner, Michael Moran

Marketing Director - Face to Face Interview

Services include

Web Design, Web Development, Search Engine Optimization

Technologies used

HTML, jQuery, PHP, JavaScript, MySql, Linux based servers, wordpress

SEO Techniques

Metadata, good content, adwords campaigns, keywords, 301 redirects, robots txt, Article submissions and inbound links

Digital Eire had never heard of RDFa They were very interested in finding out more about it They would possibly implement the technology once they researched the topic more and thought it would have an impact They offered me the opportunity to use a website of theirs for my research if I needed

5.3.5 Interview Five

Company Primary Position – Conducted over the phone

Services Include

Specialist technical SEO, Social media and digital strategy marketing

Technologies used

Web development is outsourced but deal mainly with HTML and PHP

SEO techniques

Technical SEO, Internet PR, good content, keywords, Pay Per Click, Social media LinkedIn, Facebook and Blogs have huge impact

Primary Position had never heard of RDFa After a brief explanation they said they didn't totally agree with the technology and the Semantic Web and wouldn't consider using it in the near future They said that it was invented by academics who are trying to solve a problem that doesn't exist

5 3.6 Interview Six

Company RedFly Marketing, Sinead Cochrane, Operations Manager – Conducted over the phone

Services include

Web development, SEO

Technologies used

XHTML, CSS, PHP, HTML5

SEO techniques

Link building, Content marketing, Video marketing

RedFly have heard of RDFa and implement it in some of their websites They chose to use it because it gives extra visibility to listing, especially in Google by means of "Rich Snippets" They noticed an increase in organic Click-through Rate and a perceived trust of our sites and clients that use them Some sites that don't incorporate RDFa use XML metadata because they feel they don't need them They will however be continuing to consider using the technology in future projects

5.3.7 Interview Seven

Company Terminal Four Paul Kelly, Senior Software Architect Paul was unable to take a phonecall at the time but filled out the survey and sent it back to me as follows

1 Is Web Development your line of work?

I'm the senior software architect working on the TERMINALFOUR Site Manager WCMS product While Web Development is not actually my "work", my work involves the design and development of functionality within Site Manager to enable our customers to easily develop large scale websites, mobile sites, intranets and extranets

2 What type of languages do you use to develop your websites?

Our WCMS is developed in Java and plugins to extend functionality can be developed in Java With regard to sites that we develop, the language used depends on the customer and their own preferences Generally, we'd use Java (JSP / Servlets), PHP and ASP

3 What SEO techniques do you incorporate?

All URLs published using the WCMS can be tailored for SEO In addition, internally within the WCMS, there is an SEO checker, which can be used to rate every page within the site based on keyword usage and density as well as the location of keywords within the page (H1-H6, first hundred words, anywhere within page, etc)

4 Have you heard of RDFa?

Yes

5 Do you use RDFa?

I don't think that we have used RDFa I've asked around internally and we have at least one customer that provides data using RFD feeds One of the reasons our customers buy a WCMS is to enable them to easily manage large amounts of content, while having full control over the published site Our WCMS allows for RDF and other data formats to used utilised, with some setup work

6 If you use RDFa

6 1 Why did you choose to use it?

N/A

6 2 What differences have you noticed?

N/A

7 If you have heard of RDFa but don't implement it

Why not?

In general our product and website development is driven by our customer requirements

RDFa is not something that appears to have gained traction within the markets that we target

7 1 What do you use instead?

Many of our customers in the UK higher education sector are currently using XCRI to enable them to exchange course related material XCRI is inspired by RDF design principles, although it does not use the same binding approach

7 2 Would you consider using it in the future?

Our future use of RDFa will likely depend on customer / market requirements We are more likely in the short term to more closely integrate the development of XCRI formatted data At the moment, we do not believe that there is a sufficient need within our target markets to warrant its tighter integration with TERMINALFOUR Site Manager

5.3.7.1 Observation on Interviews

After talking to the several companies about the technology RDFa I found that there was very little knowledge about it Once it was explained to the companies who were unfamiliar with the concept the majority seemed to find it promising and wanted to research it further themselves, especially if it would have a significant impact on the SEO of their sites One company (Primary Position, Interview Five) that I spoke to was not convinced about the technologies benefits I think that this is because, even after a brief description, they didn't fully understand how the technology works In comparison from my research I have found multiple companies that have benefited from

this technology, such as RedFly LTD (Interview Six) It shows that the implementation of RDFa has worked but because of either a lack of awareness or companies afraid to make the change it isn't being used as it should be

6 Conclusion

The hypothesis of this study states that "The technology RDFa is beneficial to companies in terms of increasing their presence on the web and improving their online marketing strategies" I feel that the information that has been provided in this study proves the research statement to be true With the rate at which the web and how it is used is developing, it is essential that businesses need to use the best and most efficient way of getting their information to the end user first For a company to make sure that they are noticed on the web they need to ensure that their website contains great Search Engine Optimisation techniques In addition to this they should use semantic mark-up, such as RDFa, to improve user experience, greatly increase search results and click-through-rate In the end it doesn't make sense for a company with an online profile to not try implement these semantic techniques because without it they will fall behind their competition

7 Future Works

The research that has been conducted in this study has a strong basis for future research The broader domain of the Semantic Web is a huge area with many different routes to follow This study can contribute to how web developers can begin to implement this technology and start bridging the gap between Web 2 0 and the Semantic Web

Future works would include

- Incorporating a SPARQL database and evaluate the findings

A SPARQL database is specifically designed to work with RDF and could provide a number of advantages

- Development of more semantic meaningful applications

The more semantically meaningful that data becomes in more applications will increase the capabilities in building a Semantic Web

- Development of linked data

Linked data is about connecting and sharing related data on the web

- All of the above will be providing more steps to reaching the Semantic Web

8 Bibliography

[1] Prof Eduard Babulak (2010)My 21st century cyberspace *in* 8th IEEE International Symposium on Applied Machine Intelligence and Informatics

[2] http //www retailresearch org/onlineretailing php
Accessed 28th December 2011

[3] Chyong-Ling Lun, Jin-Tsann Yeh (2010) Marketing Aesthetics on the web Personal Attributes and Visual Communication Effects *in* 2010 IEEE International Conference on Management of Innovation and Technology

[4] Ben Adida, ,Mark Birbeck and Ivan Herman (2011) Semantic Annotation and Retrieval Web of Hypertext – RDFa and Microformats *in* Handbook of Semantic Web Technologies

[5] http //www readwriteweb com/archives/tbl_calls_for_semweb php
Accessed 15th November 2011

[6] http //www w3 org/TR/2004/REC-rdf-primer-20040210/
Accessed 1st November 2011

[7] http //www w3 org/TR/rdfa-m-html/#xmlns--prefixed-attributes
Accessed 1st November 2011

[8] http //searchnewscentral com/20110207129/Technical/rdfa-the-inside-story-from-best-buy html
Accessed 28th December 2011

[9] Edward Thomas, Jeff Z Pan, Stuart Taylor, Yuan Ren, Nophadol Jekjantuk, and Yuting Zhao (2010) Semantic Advertising for Web 3 0 *in* FIS'2009 Proceedings of the Second Future internet conference on Future Internet

[10] http //rdfa digitalbazaar com/bitmunk-case-study/
Accessed 31st Dec 2011

[11] Matthew Richardson, Rakesh Agrawal, Pedro Domingos (2003)Trust Management for the Semantic Web *in* The Semantic Web – ISWC 2003

[12] Norbert Walchhofer, Milan Hronský, Michael Poettler, Robert Baumgartner, and Karl A Froeschl (2010) Semantic Online Tourism Market Monitoring *in* Information and Communication

Technologies in Tourism

[13] http //oreilly com/web2/archive/what-is-web-20 html,
Accessed 17th November 2011

[14] Murugesan, Sam (2007) Web 2 0 Second Generation Technologies *in* IT Professional

[15] Mikroyannidis, A (2007)Towards the Social Semantic Web *in* Computer, Volume 40, Issue 11

[16] Berners-Lee T , Hendler James, Lassila Ora (2001) The Semantic Web *in* Scientific American

[17] Breslin, Passant, Decker (2009) The Social Semantic Web Published by Springer

[18] Bojars, Breslin, Peristeras,Tummarello, Decker (2008) Interlinking the Social and Semantic Web *in* Intelligent Systems, IEEE, Volume 23, Issue 3

[19] Carusi, Clark, Marshall (2009) Web Semantics In Action in E-Science Workshops, 5th IEEE International Conference

[20] Heß, Maaß, Dierick(2011) From Web 2 0 to Semantic A Semi-Automated Approach
Retrieved from http //citeseerx ist psu edu/viewdoc/versions?doi=10 1 1 143 1321, 8th November 2011

[21] Wen, Youkui, Wen,Hao (2011) Semantic Text Deep Mining Based Knowledge Element *in* Internet Computing & Information Services

[22] Corchuelo, Arjona, Ruiz (2002)Automatic Extraction of Semantically Meaningful Information from the Web *in* Proceedings of the Second International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems

[23] Valentine Janev, Sanja Vranes (2009) Semantic Web Technologies Ready for adoption? *In* IT Professional, Issue 99

[24] Ben Adida, Mark Birbeck, Ivan Herman(2011) Semantic Annotation and Retrieval Web of Hypertext – RDFa and Microformats *in* Handbook of Semantic Web Technologies, pages 157-190

[25] Jespen Thomas C (2009) Just what is an ontology anyway? *in* IT Professional, Issue 5

[26] Pearse Chris (2003) Web Marketing the Basics *in* Engineering Management

[27] http //rdfa info/about
Accessed 12th Decemeber 2011

[28] http //www linuxjournal com/magazine/semantic-web-pubhshing-rdfa?page=0,1
Accessed 22nd Decmember 2011

[29] http //www w3 org/TR/xhtml-rdfa-primer/

Accessed 1st November 2011

[30] http //blog 3kbo com/2010/11/10/simple-html5-rdfa-example/
Accessed 15th November 2011

[31] Parise, Guinan (2008) Marketing Using web 2 0 *in* 41st Hawaii International Conference on System Sciences

[32] Zhang, Liu, Yin (2010) The Application Research on Web Log Mining *in* E-Marketing in 2nd Internationl Conference on e-Business and Information System Security

[33] Milam (2009) Online PSO for Web Marketing Optimization *in* IEEE Internatioal Conference on e-Business Engineering

[34] Huang, Zhang (2009) A New Marketing Effectiveness Metric Based on Web Data Mining *in* IEEE Symposium on Web Society

[35] Constantmides (2002) From Physical Marketing to Web Marketing *in* 35th Hawaii International Conference on System Sciences

[36] Ogunlana, Cheng (2009) E-Marketing and Digital Communications Implementing An Effective Knowledge Based Targeted e-Marketing Campaign *in* International Conference on Machine Learning and Cybernetics

[37] Miller (2005) The User Experience *in* IEEE Internet Computing, Volume 9, Issue 5

[38] Spohrer, Stem (2000) User Experience in the Pervasive Computing Age *in* IEEE Mutimedia, pg 13

[39] Le, Nguyen, Ha, Phan, Horiguchi (2008) Matching and Ranking with Hidden Topics Towards Online Contexutal Advertising *in* IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology

[40] Mei, Hua (2010) Contextual Internet Multimedia Advertising *in* Proceedings of the IEE, Volume 98, Issue 8

[41] Niu, Ma, Zhang (2009) A Survey of Contextual Advertising *in* Procedings of the 6th International Conference on Fuzzy Systems and Knowledge Discovery, Volume 7

[42] Mobasher, Cooley, Srivastava (2000) Automatic Personalization Based on Web Usage Mining *in* Communications of the ACM, Volume 43, Issue 8

[43] Antoniou, Paschou, Sourla, Tsakalidis (2010) A Semantic Web Personalization Technique *in* ICSC '10 Proceedings of the 20th IEEE Fourth International Conference on Semantic Computing

[44] Samir K, Habiba D (2009) Multi-agent System for personalizing Information Source

Selection *in* WI-IAT '09 Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, Volume 1

[45] Yang Q, Fan J, Wang J, Zhou L (2010) Personalizing Web Page Recommendation via Collaborative Filtering and Topic-Aware Markov Model *in* IEEE 10th International Conference on Data Mining (ICDM)

[46] Patricio, Fisk, Cunha, Nunes (2004) Customer Experience Requirements for Multi-platform Service Interaction Bringing Services Marketing to the Elicitation of User Requirements *in* RE '04 Proceedings of the 12th IEEE International Requirements Engineering Conference

[47] Tam, Ho (2003) Web Personalization- Is It Effective? *In* IT Professional, Volume 5, Issue 5

[48] Mirizzi, Ragone, Di Noia (2010) Semantic Tags generation and Retrieval of Online Advertising *in* CKIM '10 Proceedings of the 19th ACM International Conference on Information and Knowledge Management

[49] Alam, Iqbal, Noll, Chowdhury (2009) Semantic Personalization Framework for Connected Set-Top Box Environment

[50] Wolowski, Ishikawa, Sumin (2007) Semantic Web Approach to Content Personalization *in* UBICOMM '07 International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, 2007

[51] http //www peer39 com/glossary/semantic-advertising/
Accessed 13th November 2011

[52] http //www 101 biz/blogs/semantic-advertising-practical-example-semantic-web
Accessed 13th November 2011

[53] Strasser, Zugenmaier (2002) Personalization through Mask-Marketing *in* Proceedings of the 36th Annual Hawaii International Conference on System Sciences

[54] Fredrikson, Livshits (2011) Re-imaging Content Personalization and In-browser Privacy *in* 32nd IEEE Symposium on Security and Privacy

[55] Lee Y (2010) Factors Influencing Attitudes Towards Mobile Location-Based Advertising *in* 2010 IEEE International Conference on Software Engineering and Science Services

[56] http //www va-interactive com/inbusiness/editorial/sales/ibt/personal html
Accessed 13th October 2011

[57] Davis, H (May 2006) SEO, Building Traffic & Money with SEO *in* Search Engine

Optimisation, Publisher O'Reilly Media

[58] Chandra D , Rewatkar P M , Kahurke S A , Rughwani V D (2011) Search Engine Optimization *in* Asian Journal of Computing Updates and Trends, Volume 2, Issue 1

[59] Wilde Tom, Morton Kyle D, Lobacheva Yuliya, Zinovienva Nina, Meteer Marie (2009) Search Engine Optimization

Accessed 6th January 2012

http //www google com/patents?

hl=en&lr=&vid=USPATAPP12405772&id=6kfKAAAAEBAJ&oi=fnd&dq=Search+Engine+Optim ization&printsec=abstract#v=onepage&q=Search%20Engine%20Optimization&f=false)

[60] Ledford Jerry (2007) SEO Search Engine Optimization Bible, Publisher John Wiley & Sons Inc

[61] Nazar N (2009) Exploring SEO Techniques for Web 2 0 Websites Department of Computer Science and Engineering, Chalmers University of Techonolgy

Accessed 6th January 2012

http //publications lib chalmers se/records/fulltext/96559 pdf)

[62] Tomberg V , Laanpere M (2009) RDFa vs Microformats Exploring the Potential for Semantic Interoperability of Mash-up Personal Learning Environments *in* CEUR Workshop Prceedings, International Workshop on Mashup Personal Learning Environments

[63] Graf A (April 2007) RDFa vs Microformats, A Comparison of Inline Metadata Formats in XHTML *in* Digital Enterprise Research Institute Technical Report

[64] http //googlewebmastercentral blogspot com/2009/05/introducing-rich-snippets html

Accessed 6th January 2012

[65] http //developer yahoo com/searchmonkey/

Accessed 6th January 2012

[66]

http //developer yahoo com/blogs/ydn/posts/2008/09/searchmonkey_support_for_rdfa_enabled/

Accessed 6th January 2012

[67] http //developer yahoo com/blogs/ydn/posts/2010/08/api_updates_and_changes/

Accessed 6th January 2012

[68] http //microformats org/wiki/what-are-microformats

Accessed 6th January 2012

[69] http //microformats org/wiki/what-are-microformats
Accessed 6th January 2012

[70] Allsopp J (March 2007) Microformats empowering your markup for Web 2 0 Publisher friendsofED

[71] http //evan prodromou name/RDFa_vs_microformats
Accessed 6th January 2012

[72] http //dev w3 org/html5/rdfa/rdfa-module html
Accessed 6th January 2012

[73] http //www w3 org/standards/techs/rdfa#w3c_all
Accessed 6th January 2012

[74] Golbeck Jennifer (2008) Trust on the World Wide Web A Survey Publisher Now Publishers Inc

[75] Thurow S (2007) Search Engine Visibility Part 2 Publishers New Riders Press

[76] http //trust mindswap org/trustOnt shtml
Accessed 7th January 2012

[77] http //www w3 org/2000/10/swap/doc/Trust
Accessed 6th January 2012

[78] http //whatis techtarget com/definition/WYSIWYG-what-you-see-is-what-you-get
Accessed 10th July 2012

[79] http //searchcio-midmarket techtarget com/definition/plug-in)
Accessed 5th July 2012

9 List of Companies

- Primary Position
  - www primaryposition com
- Digital Eire, Ray Carolan Owner
  - www raycarolanweb ie
- Sinead Cochrane, RedFly LTD
  - www redflymarketing com
- Web Page Design Company
  - www webpagedesignco ie
- Croan
  - www croan ie
- Stewart Curry, Web Designer
  - www irishstu com
- Paul Kelly, Terminal Four
  - www terminalfour com

10 Appendices

10.1 Appendix A

RDFa The Inside Story from Best Buy
Written by Doc Sheldon
Monday, 07 February 2011 14 08
An Interview with Jay Myers
Late last year, I approached Jay Myers, Lead Development Engineer for Best Buy, and asked him if he'd be willing to be interviewed about his project of implementing RDFa on Best Buy's Website He accepted, and we began exchanging information, a process that culminated in a voice interview last Monday

It's a lengthy interview, so grab some popcorn,

1 I know you've been interested in RDFa for some time Was this Best Buy effort your first effort at implementation of RDFa?
"Yes It's going on three years since our first foray into heavy duty Semantic Web coding I initially started out using microformats, but found that RDFa (and subsequently microdata) was a more powerful and stable coding methodology based on many years of work in academia and by semantic web practitioners "
Jay went on to say that in the beginning, he played a part in reviving a microformat called hProduct He was attempting to address some issues using code and the microformats he was using weren't meeting his needs That's when he realized that RDFa was more robust, offering more flexibility That's also about the time that he made contact with Dr Martin Hepp, developer of the Good Relations ontology, with whom he began to communicate closely

2 Many seem to fear that RDFa implementation is a formidable task Others question its contribution to ROI I would think that would make it difficult to convince some organizations to

make the investment of time and/or money. How difficult was it to sell the concept internally at Best Buy?

"Initially we didn't do any "selling" of the concept…it was something I worked into a project as an experiment. The great thing about RDFa is the ability to weave meaning and rich data directly into a web page without having any impact on the front-end user experience."

Expanding on this, Jay says he was simply experimenting initially, in an effort to see if they could incorporate RDFa into some of their 1,100 stores' pages, without any adverse affects. "Management doesn't read code", he said. So with no front-end impact, he didn't have to sell the concept, he just did it!

3.What medium-to-long-term benefits did you expect to see for Best Buy?

"We really didn't go into it with any expectations. We just wanted to see if it was something we might want to do. That's why we were caught by surprise by the results… we weren't really expecting any."

4.What benefits turned up?

"Within just a couple of months, we began to see an increase in our organic search results. Before long, it had increased by 30% over historical rates. We also saw an increase in our click-through rate. Yahoo did a study a while back and found that people that had rich snippets on the results pages were seeing around a 15% increase in CTR, which has proven to be the case for us. And of course, it makes our web site "smarter" and more open to machines, which ultimately benefits customers."

5.What led you to choose RDFa over alternative formats, such as microformats or microdata?

"I found that RDFa was a much more stable concept – based on the use of long established vocabularies (also known as ontologies) that have existed for years. My first foray into giving objects better definition on the web was revitalizing the hProduct microformat. Working through that implementation I found limitations in microformats that left me searching for a "beefier" solution. At the time, microdata hadn't hit the scene yet, so RDFa was the natural choice."

6.What architecture strategy did you employ, and what were the specific reasons for that choice?

"Initially, we deployed RDFa markup through our local stores' WordPress blogs simply by weaving the rich markup and attributes into the WordPress themes. An advantage to "front-end semantics" is that a developer or team doesn't have to use a particular platform or employ a particular arch strategy to populate their sites with rich data. A savvy developer could hand code the stuff into their

work just as easily as adopting a semantic-specific platform "

7 How hard was it for Best Buy to implement RDFa? What was the methodology?

"Initially, implementing RDFa was a personal challenge, as a developer has to shift their development mindset and consider not only what the visual output of the code is, but what the machine output is as well Outside of a couple of extra validation steps to confirm the RDFa output could be successfully distilled by machines and software, the coding methodology really didn't change that much Over the past two years I have worked hard to automatically build semantics into my project methodologies – making it a standard rather than a separate implementation step "

8 What are the greatest benefits you've seen from implementing RDFa?

"The rise in our organic search traffic, and the adoption by major companies like Facebook and Google "

9 How fast did you begin to see benefits, and what were they, intially?

"Within 3 months of our initial deployment we saw a surge in our organic search engine traffic That eventually reached 30% and held it "

10 Did you see a rankings increase on Google?

"This is hard to tell – I've had a difficult time separating search traffic from scraper applications utilizing traditional web analytics To my knowledge, there are a handful of interested individuals running distilling and parsing software against our RDFa pages "

11 You said you saw about a 15% increase in CTR Did you notice any appreciable change in your bounce rate?

"Not really Some stores, of course, may be less active on their blog, and they see a higher bounce rate because of that, but that's something we want to watch closer "

(This started out as an experiment, so the results caught them by surprise They hadn't really been monitoring to the extent they might have, had they expected such results)

12 What are your plans now?

"We're continuing to implement RDFa on product pages across our various web properties and have our attention focused on the perceived impact that rich RDFa markup could have on sales numbers "

13 Do you foresee any major impact to the acceptance and implementation of RDFa by the release of Drupal 7?

"Kudos to the Drupal folks for implementing semantics into their CMS I believe that the Drupal 7 release will make semantics even more accessible than ever before to everyday developers (and many people with no dev experience working on the web), and allow them to easily open up and release useful data through RDFa "

14 Any specific comments or recommendations you'd like to offer those that are either on the fence about implementing RDFa for their site, or have decided it doesn't offer them sufficient ROI to make it worth the effort?

"There's definitely a benefit in being an early adopter, that's one thing Additionally, with RDFa, we're getting closer to the Semantic Web Tim Berners Lee is the one asking for our data, and I think it makes sense to do that RDFa is one of the what I call a gateway drug to the Semantic Web

You don't have to be an expert, you can be an every day developer or somebody that uses a tool like Drupal or another CMS to put out rich data in RDFa So I think it's not as hard as people make it out to be Again, it was more of an experiment for us, it wasn't a hard-core effort, and the hope is that we can make this part of the everyday development, just like coding HTML it should be just that easy

We're seeing a lot of adopters, including some big names, like Google with its rich snippets and also Facebook Facebook has something called the Open Graph, that also uses RDFa and rich data So there's a benefit there, and I see more adoption by the larger and more important firms It just makes sense We're living in a world where there's a huge amount of data on the web, and it's only going to continue to get larger And I don't believe that current or traditonal SEO practices, page sculpting, things like that I think that can only go so far But traditional SEO efforts, in combination with RDFa and some more development-centered efforts, can really go a long way for people So that would be my words of encouragement

We're starting, as the Semantic Web community, to engage the SEO community a little bit more, to try and get everybody on the same page So I'm going to be speaking in Austin on March 1st, at the Semantic Web Meetup there What I'm trying to do is tailor this message not only to a technical audience, but to the SEO audience, too There is an effort now to try to tie development and SEO together The more rich data we have on the web, the better off we're going to be "

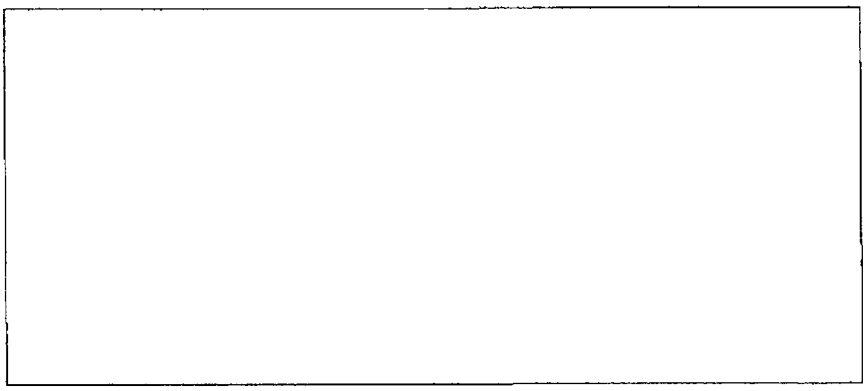
10.2 Appendix B

Questionaire

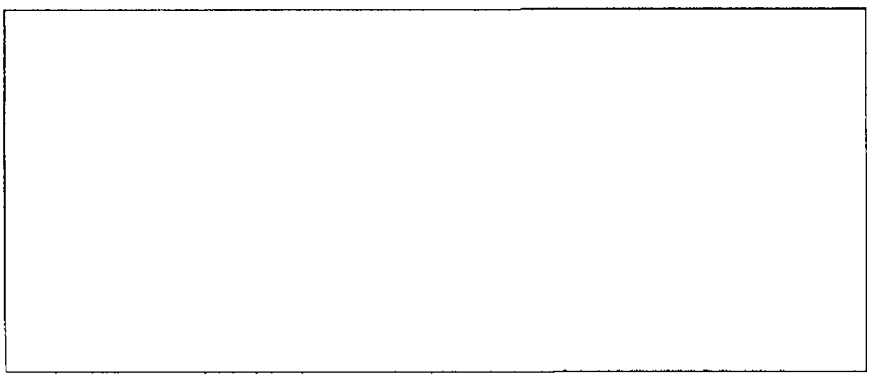
Is Web Development your line of work?

What type of languages do you use to develop your websites?

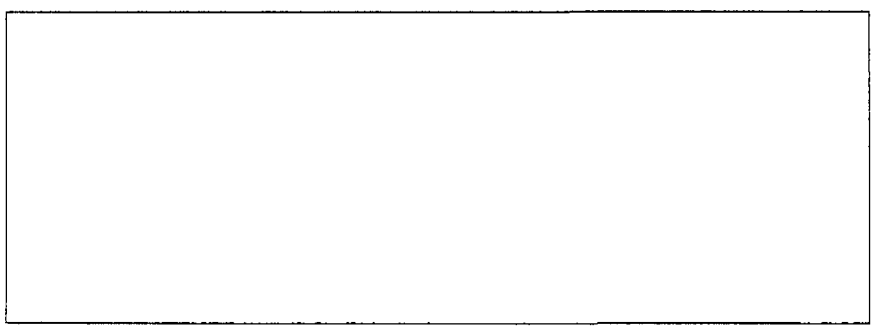
What SEO techniques do you incorporate?



Have you heard of RDFa?



Do you use RDFa?



If you use RDFa

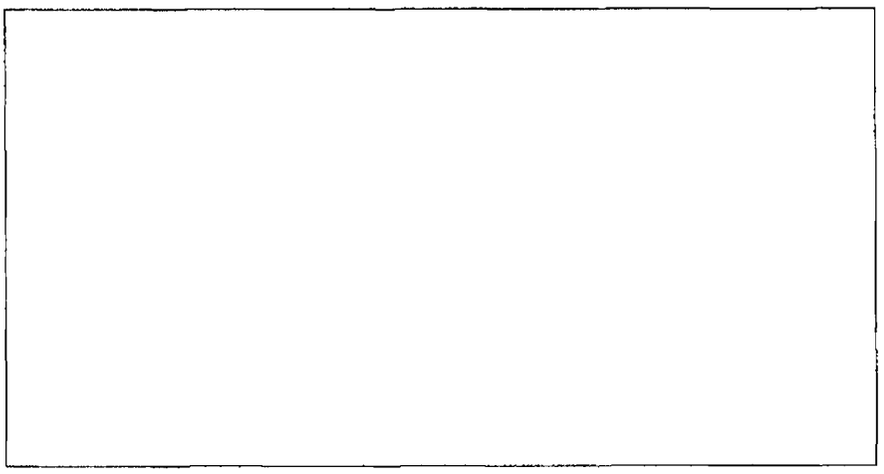
Why did you choose to use it?

What differences have you noticed?

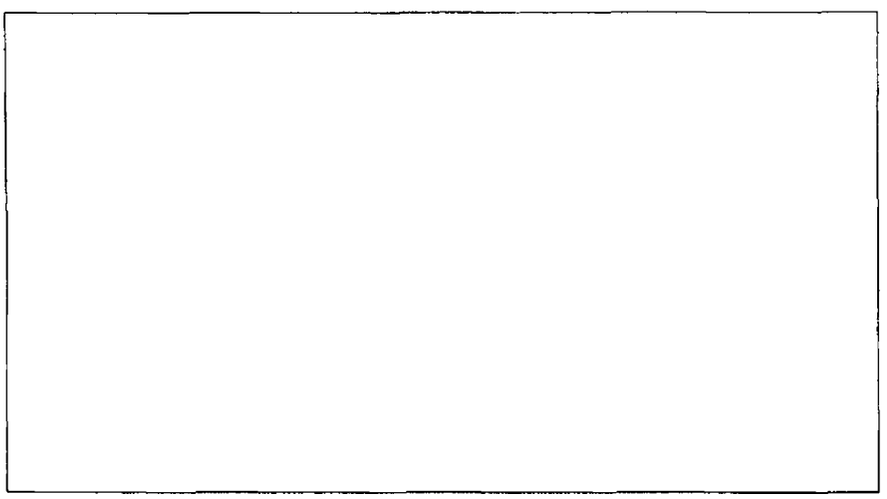
If you have heard of RDFa but don't implement it

- Why not?

- What do you use instead?



- Would you consider using it in the future?



10.3 Appendix C – Project Proposal

RDFa is Beneficial to Companies using the World Wide Web

Orla Butterly,
x06350682,
orlamb@hotmail com

MSc in Web Technologies
Date 26th September 2011

Objectives

The objective of this research is to show how RDFa can be beneficial to companies that use the World Wide Web It is based on my research statement "This technology is beneficial to companies but is not implemented due to lack of awareness in the area " The dissertation will include research on The Semantic Web, Web Marketing and Search Engine Optimization The system that will be developed alongside this will aim to prove or disprove the statement that companies benefit from RDFa

"Search Engine Optimization is an Internet marketing strategy and a process, widely used nowadays for improving the volume or quality of traffic to a website through search engines "

- Exploring SEO Techniques for Web 2 0 Websites, Najam Nazar

Background

The World Wide Web has been around for several decades now but it hasn't been until recent years that it began to evolve rapidly to what we see today There has been a large influx of companies big and small deciding to share their information on this platform and because of this there is a greater

amount of data available One company needs to stand out from the other Their marketing and advertising campaigns have all been altered to ensure they are still at a competitive advantage There are various ways for companies to market themselves and due to the growth in the social web some of the most favourable ways are by media such as images or video Search Engines are also another way for companies to try to get their information out there to the right target market

Technical Approach

The application will be a sample website that is built using different types of media that is seen in online marketing It will embed the RDFa into the web document so that the data becomes machine-readable

I will be using journals, articles and various other works in relation to marketing, Semantic Web and Search Engine Optimization to execute my research

Special resources required

No special resources are required

Technical Details

The main technologies that will be implemented to build this application are

- XHTML
- RDFa
- CSS

Project Plan

As I do not own a version Microsoft Project I will list the dates below

Due Date	Deliverable
03/10/11	Literature Review
N/A	Research Background
24/10/11	Research Position Paper
16/12/11	Architecture and

	Implementation
23/12/11	Evaluation
02/01/12	Conclusion and Summary
16/01/12	Master's Thesis
21/01/12	Viva

Testing and Evaluation

The application will be tested by comparing the data between the semantically marked up website and a "placebo" website It will monitor the search rankings, page traffic using Google Analytics I will also conduct various interviews regarding the area of research

Consultation 1

Ron Elliot

Ron provided me with a series of ideas regarding the area of web marketing He was a great help in figuring out from me the area I was most interested in and thought I should take a usability approach

Consultation 2

Paul Stynes

Paul's area of expertise is in semantic web His advice and insight in the area was very helpful He pointed me in the right direction regarding technologies

Proposed Supervisor

Paul Stynes / Ron Elliot

References

http //www optimum7 com/internet-marketing/internet-marketing-strategy/the-future-of-internet-marketing-and-social-media html - 28/9/11, 6 00pm

The New Rules of Marketing & PR (Second Edition) – Scott D M (2010)

http //rdfa info

The Semantic Web, Berners- Lee T (2001)

__Orla Butterly___

Signature of student and date