Department of Information Systems and Computers

Graduation Project

Enhancing Internet User Experience using Semantic Web Technology

"Enhanca" Application

Supervised by

Dr. Maged Farouk Elsayed

Prepared by

Ahmed FathyMohamed.
AshrafShabanAbdEl-rawaf.
Islam Ahmed Hassan.
KhaledMohamed Adel Kamra.
MahmoudAhmed Hassan Mehlab
Mohamed Adel Mohamed Alsayed Rezk
MostafaMohamed AsranHabliza.

July 2010

Acknowledgment

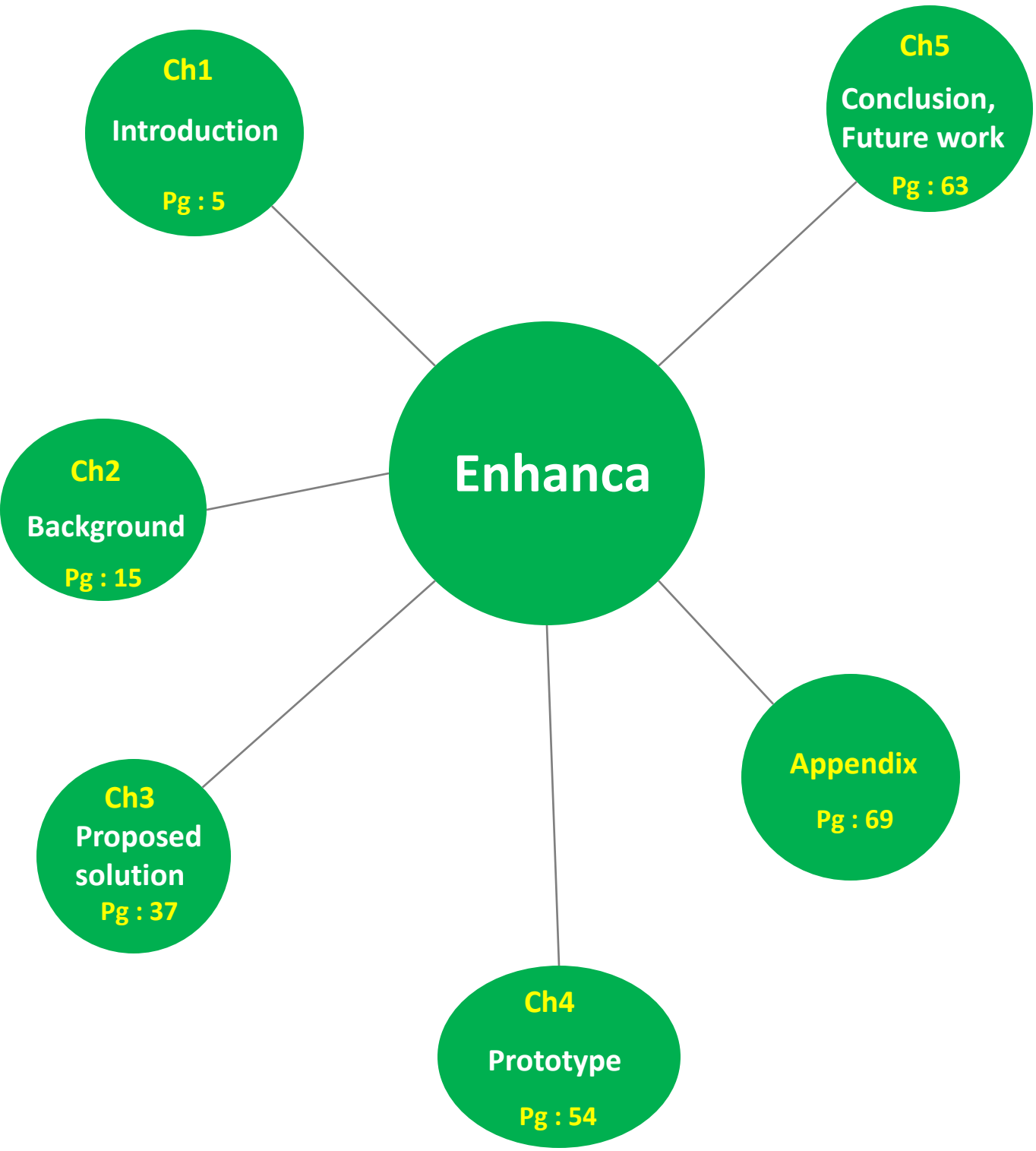
“ Imagination, is more important than knowledge”

Albert Einstein

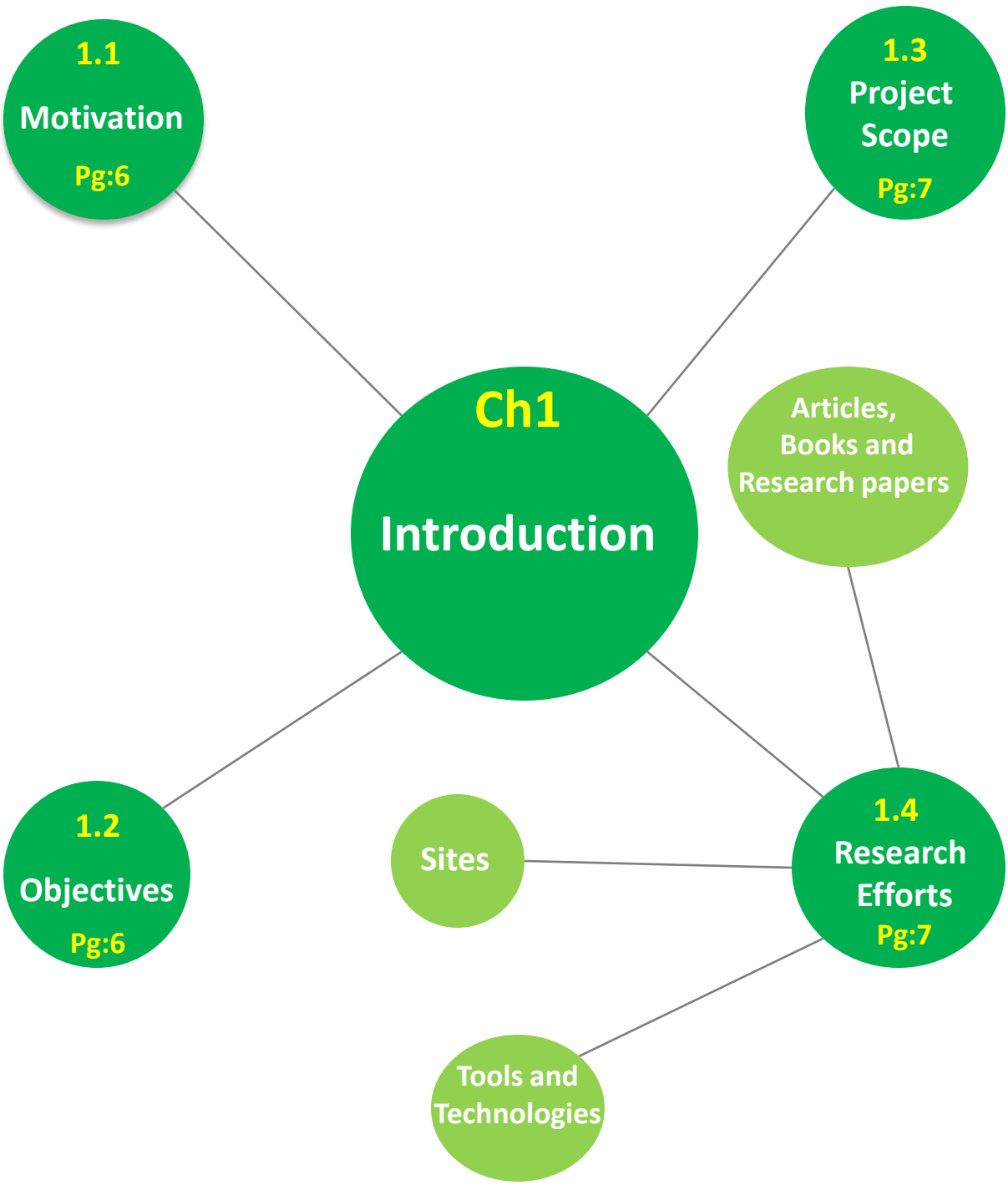
Thanks to every one who encouraged us to be what we are now,

Thanks to every one who inspired us to widen our imagination and creativity.

Contents


Ch1
Introduction
Pg : 5
Ch5
Conclusion, Future work
Pg : 63
Enhanca
Ch2
Background
Pg : 15
Ch3
Proposed solution
Pg : 37
Ch4
Prototype
Pg : 54
Appendix
Pg : 69

Chapter One: Introduction



1.1 Motivations:

Internal and external factors that stimulate desire and energy in people to be continually interested in and committed to a job, role, or subject, and to exert persistent effort in attaining a goal. Motivation is the energizer of behavior and mother of all action. It results from the interactions among conscious and unconscious factors such as the intensity of desire or need, incentive or reward value of the goal, and expectations of the individual and of his or her significant others. And our motivations that lead us to this project were :

1. Internet has changed in the last ten years, but search engines didn't keep up with this intensive upgrading.
2. We still spending too much time to reach the information we are Searching about, even if it was an answer of a simple question.
3. Today, in order to get an answer of a simple question you are ought to search, and pore through multiple pages.
4. Search engines still providing same old kind of results (links to documents) since 90's. While in most cases we look for summaries, structured information, lists of things, comparison tables, direct answer, or sentences with complete meaning.
5. How websites are designed is another issue!
6. Enough Data, we want Information.
7. Search engines doesn't understand the indexed documents, nor queries.
8. We believe that we already reached a dead end of the current technology, and we need to invent new beginning.

1 .2 Objectives:

A goal or objective is a projected state of affairs that a person or a system plans or intends to achieve a personal or organizational desired end-point in some sort of assumed development. Many people endeavor to reach goals within a finite time by setting deadlines.

1. Studying Semantic web technologies.
2. Researching the state of the art of semantic web solutions.
3. Experimenting the usage of some existing semantic web solutions and tools.
4. Brain storming to introduce new ideas about how to enhance texts and improve search engines' results through semantic web technologies.
5. Building prototypes to demonstrate that ideas.
6. Making information available on request in smart way using semantic technologies (note: Information not data).
7. Optimizing the time that researcher and people need to find an answer to their questions.
8. Building knowledge base .
9. Implementing Semantic web technologies in Arabic..
10. Motivating Egyptian and Arabian researcher s to begin in this hot field.

1 . 3 Project Scope:

The scope statement is an agreement among the project team, the project sponsor and key stakeholders. It represents a common understanding of the project for the purpose of facilitating communication among the project Supervisor and team. The scope statement includes relating the project to business objectives, and defining the boundaries of the project in multiple dimensions including approach, deliverables, milestones, and budget. Our project scope are depicted as the following:

Studying and researching the state of the art of semantic web solutions and technologies, and experimenting existing semantic web tools, brainstorming for introducing new ideas about enhancing texts, in order to provide proper knowledge to internet users and to build a prototype for demonstrating the concept. The project does not involve building a full-scale Semantic Web search engine.

1.4 Research Efforts:

Our project was mainly about researching and brain storming new ideas and solutions for enhancing internet environment through the semantic web technologies and solutions.

Our Research Experience Related to this project :

Reading research papers and books:

Examples:

1. Pull " the power of the semantic web to transform your business".
2. Social networks and the semantic web.
3. The Fate of the Semantic Web.
4. Semantics-Oriented Natural Language Processing.
5. Research Challenges and Perspectives of the Semantic Web.
6. Collective Knowledge Systems: Where the Social Web meets the Semantic Web.
7. Practical Semantic Web – Tagging and Tag Clouds.
8. Representing and Sharing Tagging Data Using the Social Semantic Cloud of Tags.
9. Semantic Cloud: An Enhanced Browsing Interface for Exploring Resources in Folksonomy Systems.
10. Measuring Semantic Similarity between Words Using Web Search Engines.
11. Active Logic Semantics for a Single Agent in a Static World.

… 60+

Our research experience related to this project:

- **Surveying many semantic related web sites.**
- **Investigating many semantic web sites:**

Examples:

1. wolframalpha.com
2. Kngine.com
3. Collecta.com
4. Zemanta.com
5. Headup.com
6. BooRah.com
7. Swotti.com
8. Dapper.net
9. Inform.com
10. Evri.com
11. Faviki.com

… 200 +

i.e.

1. **BooRah** is a restaurant review site that we first reviewed earlier this year. One of BooRah's most interesting aspects is that it uses semantic analysis and natural language processing to aggregate reviews from food blogs. Because of this, BooRah can recognize praise and criticism in these reviews and then rates restaurants accordingly. BooRah also gathers reviews from Citysearch, Tripadvisor and other large review sites. BooRah also announced last month the availability of an API that will allow other web sites and businesses to offer online reviews and ratings from BooRah to their customers. The API will surface most of BooRah's data about a given restaurant, including ratings, menus, discounts, and coupons.
2. **Swotti** is a semantic search engine that aggregates opinions about products to help you make purchasing decisions. We reviewed the product back in March. Swotti aggregates opinions about products from product review sites, forums and discussion boards, web sites and blogs, and then categorizes those reviews as to what feature or aspect of the product is being reviewed, tagging it accordingly, and then rating the review on as positive or negative.
3. **Dapper** The idea is that publishers can tell Dapper: this is the place on my web page where the title of a movie will appear, now serve up a banner ad that's related to whatever movie this page happens to be about. That could be movies, books, travel destinations - anything. We remarked that the UI for this has grown much more sophisticated in the past year. How this works: in the back end, Dapper will be analyzing the fields that publishers identify and will apply a layer of semantic classification on top of them. The company believes that its new ad network will provide monetary incentive for publishers to have their websites marked up semantically. Dapper also has a product called Semantify, for SEO .

1. **Inform.com** analyzes content from online publishers and inserts links from a publisher's own content archives, affiliated sites, or the web at large, to augment content being published. including CNN.com, WashingtonPost.com and the Economist. Inform says its technology determines the semantic meaning of key words in millions of news stories around the web every day in order to recommend related content. The theory is that by automating the process of relevant link discovery and inclusion, Inform can easily add substantial value to a publisher's content. Inform also builds out automatic topic pages, something you can see around WashingtonPost and CNN.com.
2. **Evri** is a Paul Allen (of Microsoft fame) backed semantic search engine that launched into a limited beta in June. Evri is a search engine, though it adds a very sophisticated semantic layer on top of its results that emphasizes the relationships between different search terms. It especially prides itself for having developed a system that can distinguish between grammatical objects such subjects, verbs, and objects to create these connections.
3. **Faviki** is a new social bookmarking tool. It offers something that services like Ma.gnolia, del.icio.us and Diigo do not - semantic tagging capabilities. What this means is that instead of having users haphazardly entering in tags to describe the links they save, Faviki will suggest tags to be used instead. However, unlike other services, Faviki's suggestions don't just come from a community of users and their tagging history, but from structured information extracted straight out of the Wikipedia database. Because Faviki uses structured tagging, there is more that can be learned about a particular tag, its properties, and its connections to other tags. The system will automatically know what tags belong together and how they relate to others.

- **Downloading, developing and implementing many tools, applications, software, APIs and libraries:**

Examples:

1. **Stanford Log-linear Part-Of-Speech Tagger**

   A Part-Of-Speech Tagger (POS Tagger) is a piece of software that reads text in some language and assigns parts of speech to each word (and other token), such as noun, verb, adjective, etc., although generally computational applications use more fine-grained POS tags like 'noun-plural'. This software is a Java implementation of the log-linear part-of-speech taggers

2. **Open jena**

   Jena is a Java framework for building Semantic Web applications. It provides a programmatic environment for RDF, RDFS and OWL, SPARQL and includes a rule-based inference engine.

3. **Open cyc**

   Cyc is an artificial intelligence project that attempts to assemble a comprehensive ontology and knowledge base of everyday common sense knowledge .

4. **Open Calais**

   Calais is a rapidly growing toolkit of capabilities that allow you to readily incorporate state-of-the-art semantic functionality within your blog, content management system, website or application.

…50+

- **Learning new technologies and languages (Programming languages, DBMS and Query languages) associated and not associated to the project:**

Examples:

1. **PHP**

   Hypertext Preprocessor (the name is a recursive acronym) is a widely used, general-purpose scripting language that was originally designed for web development to produce dynamic web pages.

2. **JSON**

   JSON, short for JavaScript Object Notation, is a lightweight computer data interchange format. It is a text-based, human-readable format for representing simple data structures and associative arrays (called objects).

3. **PYTHON**

   Python is a general-purpose high-level programming language whose design philosophy emphasizes code readability. Python aims to "[combine] remarkable power with very clear syntax", and its standard library is large and comprehensive.

4. **JAVA**

   Java is a programming language originally developed by James Gosling at Sun Microsystems (which is now a subsidiary of Oracle Corporation) and released in 1995 as a core component of Sun Microsystems' Java platform.

5. **Jquery**

   jQuery is a lightweight cross-browser JavaScript library that emphasizes interaction between JavaScript and HTML. It was released in January 2006 at BarCamp NYC by John Resig. Used by over 27% of the 10,000 most visited websites, jQuery is the most popular JavaScript library in use today.

6. **Java script**

   JavaScript is an object-oriented scripting language used to enable programmatic access to objects within both the client application and other applications.

7. **AJAX**

   Ajax (shorthand for Asynchronous JavaScript and XML) is a group of interrelated web development techniques used on the client-side to create interactive web applications.

8. **Cassandra**

   Cassandra is an open source distributed database management system. It is an Apache Software Foundation top-level project

9. **XML / RDF / SPARQL [1]**

… and more

(1) Explained in more details in next chapter.

Establishing connections with the academic and professional community of the semantic web.

Examples:

1. We contact with Haytham el-fadeel Kngine Founder, then we connected him to Dr. Mahmoud Yousef. Arab Academy for Science and technology, currently Rutgers University.
2. And he - Dr. Mahmoud -had shown a great interest in the semantic web technologies so he invited Haytham to make a speech in the (AAST) about Kngine and Semantic technologies.
3. Haytham invited us after confirmation from Dr. Mahmoud Yousef to his speech in the (AAST).
4. We made tow presentations at the Department of Information Systems and Computers, Alexandria University about the evolution of the web, semantic web technologies and our introduced ideas and solutions.
5. Sending our Project proposals to incubators and venture capitals organizations.
6. Attending the professional forums' discussions and downloading the latest semantic conferences videos.

….and more

Other Efforts:

Semantic

- Social networks and the semantic web.
- The Fate of the Semantic Web.
- Semantics-Oriented Natural Language Processing.
- Research Challenges and Perspectives of the Semantic Web.
- Collective Knowledge Systems: Where the Social Web meets the Semantic Web.
- Practical Semantic Web – Tagging and Tag Clouds.
- Representing and Sharing Tagging Data Using the Social Semantic Cloud of Tags.
- Semantic Cloud: An Enhanced Browsing Interface for Exploring Resources in Folksonomy Systems.
- Measuring Semantic Similarity between Words Using web Search Engines.

Semantic applications

- Semantic Web Application Areas.
- Semantic Web Applications.
- Semantic Web Technologies.
- Interlinking Open Data on the Web.
- Sem.Search: A Search Engine for the Semantic Web.
- TopQuadrant Dramatically Accelerates Semantic Web Application Development with TopBraid Suite 3.0.

Semantics in business

- Applying Semantics to Service Oriented Architectures.
- THE SEMANTIC E-BUSINESS VISION.
- Semantic Web Rules for Business Information.

Ontology

- A Synthesis of State of the Art of Enterprise Ontologies.
- Context obfuscation for privacy via ontological descriptions?
- AquaLog: An ontology-driven question answering system for organizational semantic intranets.
- An Ontology Based Question Answering System on Software Test Document Domain.
- Ontologies and Lexical Semantics in Natural Language understanding.
- Ambient Intelligence through Ontologies.
- Explicit Semantics for Business Ontology.

RDF & Xml

- THE SEMANTIC WEB :The Roles of XML and RDF.

Query formulation

- Semantics driven support for query formulation.
- QUERY FORMULATION IN WEB INFORMATION SEARCH.
- Enriching the conceptual basis for query formulation through relationship semantics in databases.

Machine learning

- Machine Learning for Information Retrieval: Neural Networks, Symbolic Learning, and Genetic Algorithms.
- The WEKA machine learning workbench: Its application to a real world agricultural database.
- Open-Source Machine Learning: R Meets Weka.
- Supervised Machine Learning: A Review of Classification Techniques.
- Theoretical Machine Learning.
- Preliminaries.
- One-Class SVMs for Document Classification.
- Thumbs up? Sentiment Classification using Machine Learning Techniques

Text mining

- Introduction to Text Mining.
- Advancements in Text Mining Algorithms and Software.
- Relation between recognition and recognition failure of recallable words.
- Feature-Rich Part-of-Speech Tagging with a Cyclic Dependency Network.
- Automatic Acquisition of Language Model based on Head-Dependent.
- Relation between Words.
- Models for the Semantic Classification of Noun Phrases.
- Similarity of Objects and the Meaning of Words.
- Structured Index System at NTCIR1: Information Retrieval using Dependency Relationship between Words.
- A Relational Model of Semantic Similarity between Words using Automatically Extracted Lexical Pattern Clusters from the Web.
- Using Lexical Semantic Techniques to Classify Free-Responses.
- Semantic Relation between Words with the Web as Information Source.
- Lexical frequency and acoustic reduction in spoken Dutch.
- Open Source Text Mining.

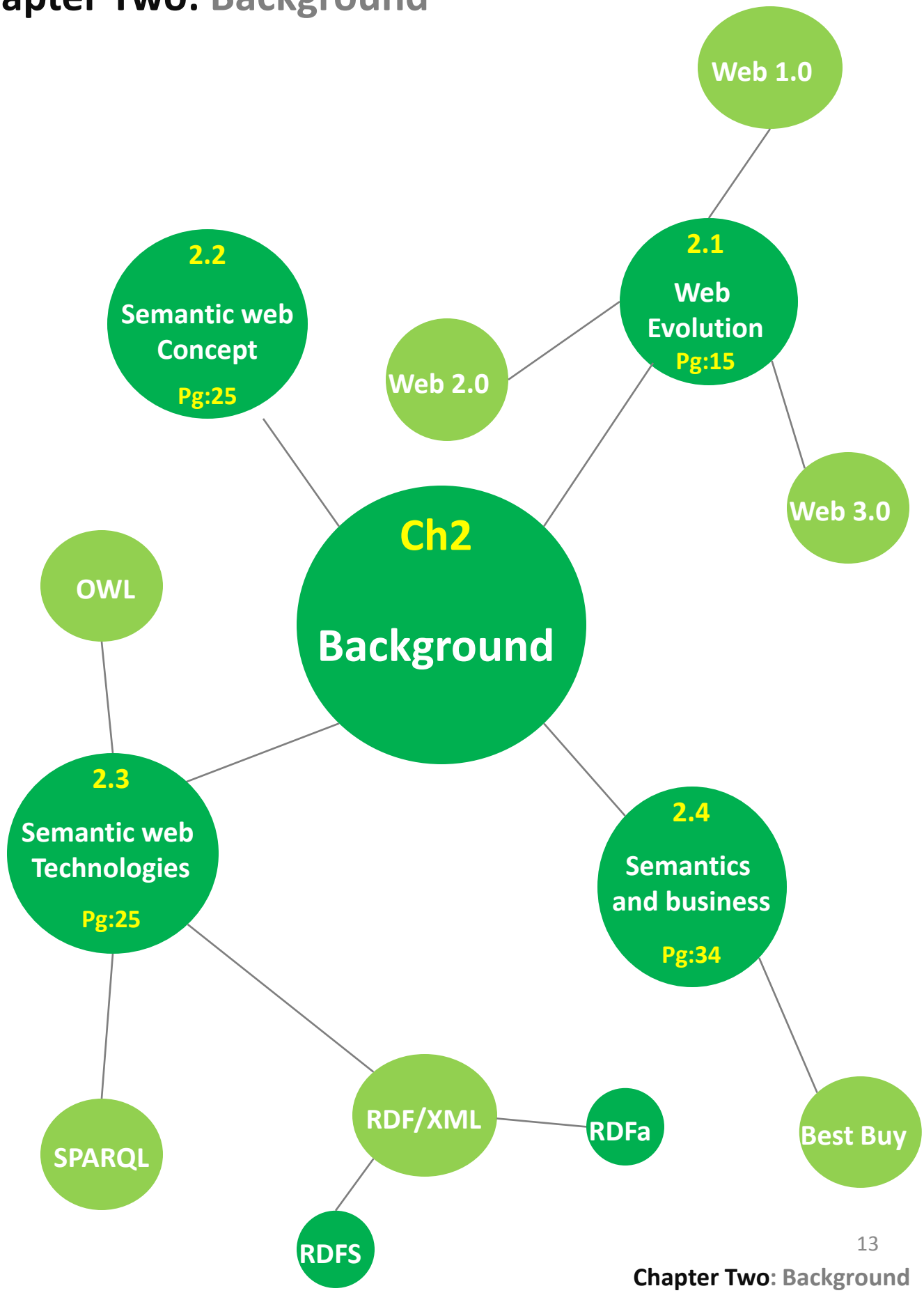
Stemming

- Light Stemming for Arabic Information Retrieval.
- A rule-Based extensible stemmer for information retrieval with application to Arabic.
- Improving Stemming for Arabic Information Retrieval: Light Stemming and Co-occurrence Analysis.
- An implementation of a word stemming algorithm for English.
- Vocabulary size and term distribution: tokenization, text normalization and stemming.

Others

- Example of Business Process Improvement.
- Genetic algorithms ,noise , and the sizing of populations.
- An Automated Approach for Webpage Classification.
- Faceted Wikipedia Search.
- Wikipedia-based Semantic Interpretation for Natural Language Processing.
- Artificial Intelligence: An Overview.
- Signed Approach for Mining Web Content Outliers.

Chapter Two: Background



2.1. Web Evolution:

2.1.1. Web 1.0 (1991-2003):

Concept:

- Is a retronym which refers to the state of the World Wide Web, and any website design style used before the advent of the Web 2.0 phenomenon. Web 1.0 began with the release of the WWW to the public in 1991, and is the general term that has been created to describe the Web before the "bursting of the Dot-com bubble“ [1] in 2001, which is seen by many as a turning point for the internet.

Characteristics and Features:

- Static pages instead of dynamic user-generated content.
- The use of framesets.
- Online guestbooks.
- GIF buttons, typically 88x31 pixels in size promoting web browsers and other products.
- HTML forms sent via email. A user would fill in a form, and upon clicking submit their email client would attempt to send an email containing the form's details.

Web 1.0 examples:

- Yahoo.com
- msn.com
- Altavista.com
- Icq.com
- Masrawy.com
- Sun.org

(1)The "dot-com bubble" (or sometimes "IT bubble“ or "TMT bubble") was a speculative bubble covering roughly 1995–2000 (with a climax on March 10, 2000 with the NASDAQ peaking at 5132.52) during which stock markets in industrialized nations saw their equity value rise rapidly from growth in the more recent Internet sector and related fields. While the latter part was a boom and bust cycle, the Internet boom sometimes is meant to refer to the steady commercial growth of the Internet with the advent of the world wide web as exemplified by the first release of the Mosaic web browser in 1993 and continuing through the 1990s.

(Different web eras and their technologies) This diagram shows that every era has its own technologies which make it possible and reachable.

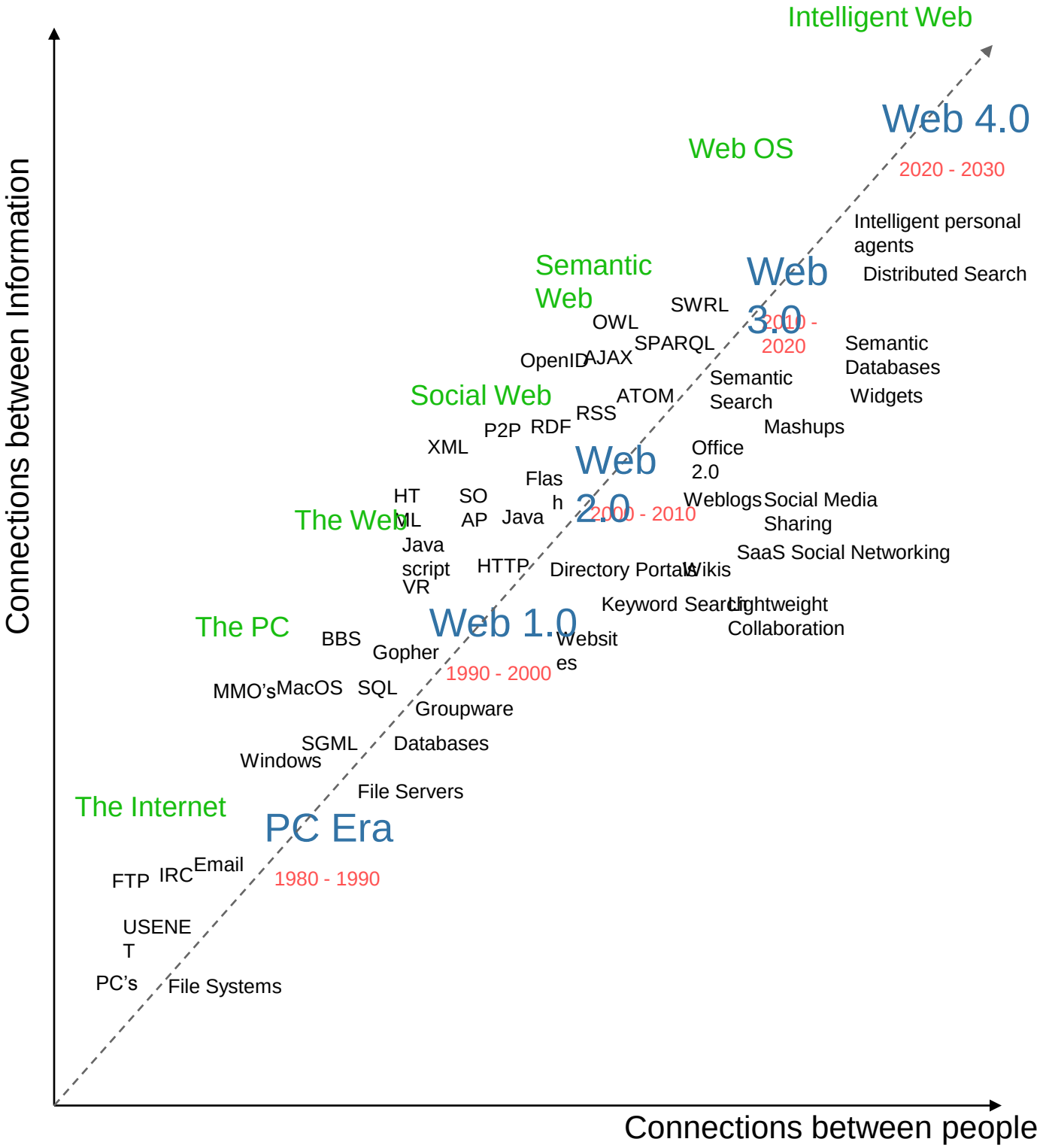


Figure 2.1

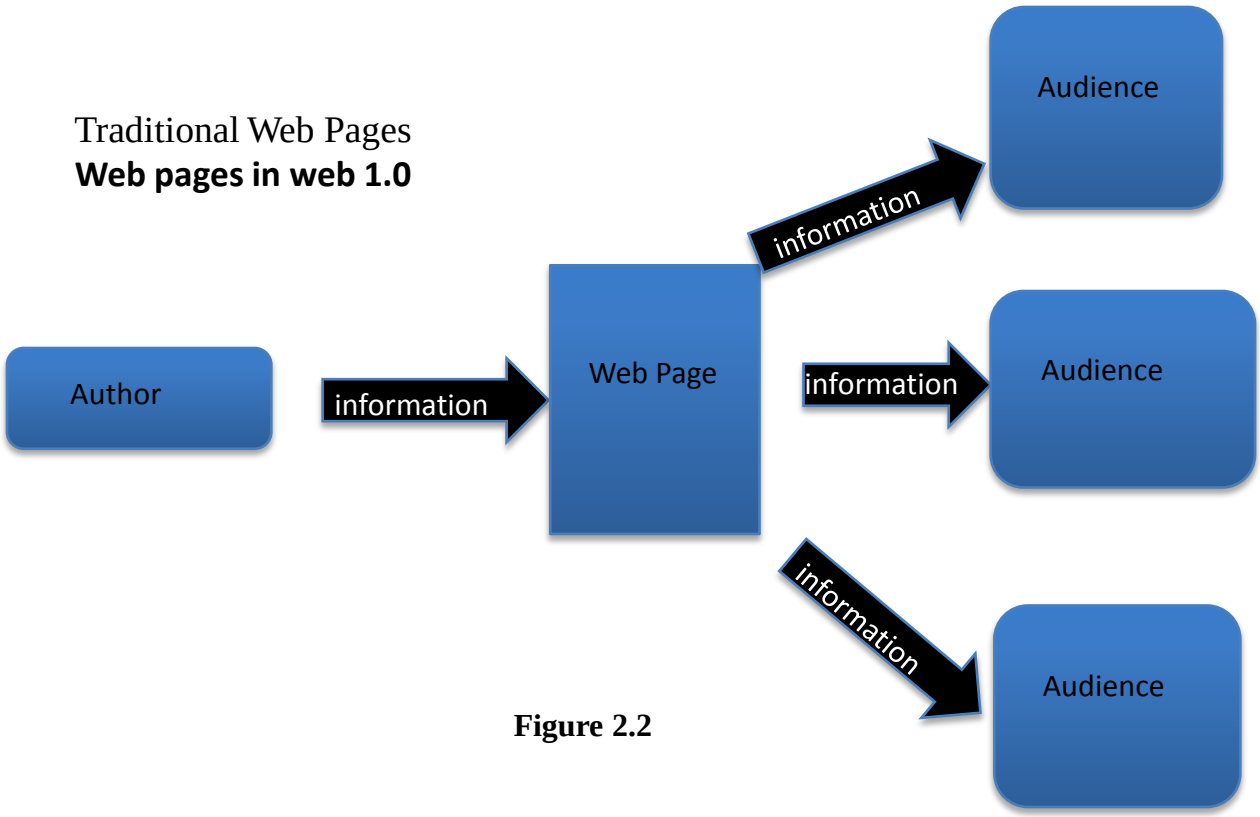


Figure 2.2

2.1.2. Web 2.0:

Concept:

- Is commonly associated with web applications that facilitate interactive information sharing, interoperability, user-centered design, and collaboration on the World Wide Web. A Web 2.0 site allows its users to interact with each other as contributors to the website's content, in contrast to websites where users are limited to the passive viewing of information that is provided to them. Examples of Web 2.0 include web-based communities, hosted services, web applications, social-networking sites, video-sharing sites, wikis, blogs, mashups, and folksonomies.
- The term is closely associated with Tim O'Reilly[1] because of the O'Reilly Media Web 2.0 conference in 2004 . Although the term suggests a new version of the World Wide Web, it does not refer to an update to any technical specifications, but rather to cumulative changes in the ways software developers and end-users use the Web. Whether Web 2.0 is qualitatively different from prior web technologies has been challenged by World Wide Web inventor Tim Berners-Lee, who called the term a "piece of jargon"— precisely because he specifically intended the Web to embody these values in the first place.

(1) Tim O'Reilly (Irish: *Tadhg Ó Raghallaigh*) (born June 6, 1954) is the founder of O'Reilly Media (formerly O'Reilly & Associates) and a supporter of the free software and open source movements.

History:

- The term "Web 2.0" was coined in 1999 by Darcy DiNucci. In her article, "Fragmented Future," DiNucci writes:
- The Web we know now, which loads into a browser window in essentially static screenfulls, is only an embryo of the Web to come. The first glimmerings of Web 2.0 are beginning to appear, and we are just starting to see how that embryo might develop. The Web will be understood not as screenfulls of text and graphics but as a transport mechanism, the ether through which interactivity happens. It will [...] appear on your computer screen, [...] on your TV set [...] your car dashboard [...] your cell phone [...] hand-held game machines [...] maybe even your microwave oven.
- Her use of the term deals mainly with Web design and aesthetics; she argues that the Web is "fragmenting" due to the widespread use of portable Web-ready devices. Her article is aimed at designers, reminding them to code for an ever-increasing variety of hardware. As such, her use of the term hints at – but does not directly relate to – the current uses of the term.
- The term did not resurface until 2003. These authors focus on the concepts currently associated with the term where, as Scott Dietzen puts it, "the Web becomes a universal, standards-based integration platform".
- In 2004, the term began its rise in popularity when O'Reilly Media and MediaLive hosted the first Web 2.0 conference. In their opening remarks, John Battelle and Tim O'Reilly outlined their definition of the "Web as Platform", where software applications are built upon the Web as opposed to upon the desktop. The unique aspect of this migration, they argued, is that "customers are building your business for you". They argued that the activities of users generating content (in the form of ideas, text, videos, or pictures) could be "harnessed" to create value.
- O'Reilly et al. contrasted Web 2.0 with what they called "Web 1.0". They associated Web 1.0 with the business models of Netscape and the Encyclopedia Britannica Online. For example,
- Netscape framed "the web as platform" in terms of the old software paradigm: their flagship product was the web browser, a desktop application, and their strategy was to use their dominance in the browser market to establish a market for high-priced server products. Control over standards for displaying content and applications in the browser would, in theory, give Netscape the kind of market power enjoyed by Microsoft in the PC market. Much like the "horseless carriage" framed the automobile as an extension of the familiar, Netscape promoted a "webtop" to replace the desktop, and planned to populate that webtop with information updates and applets pushed to the webtop by information providers who would purchase Netscape servers.

History:

- In short, Netscape focused on creating software, updating it on occasion, and distributing it to the end users. O'Reilly contrasted this with Google, a company which did not at the time focus on producing software, such as a browser, but instead focused on providing a service based on data. The data being the links Web page authors make between sites. Google exploits this user-generated content to offer Web search based on reputation through its "page rank" algorithm. Unlike software, which undergoes scheduled releases, such services are constantly updated, a process called "the perpetual beta".
- A similar difference can be seen between the Encyclopedia Britannica Online and Wikipedia: while the Britannica relies upon experts to create articles and releases them periodically in publications, Wikipedia relies on trust in anonymous users to constantly and quickly build content. Wikipedia is not based on expertise but rather an adaptation of the open source software adage "given enough eyeballs, all bugs are shallow", and it produces and updates articles constantly.
- O'Reilly's Web 2.0 conferences have been held every year since 2004, attracting entrepreneurs, large companies, and technology reporters. In terms of the lay public, the term Web 2.0 was largely championed by bloggers and by technology journalists, culminating in the 2006 TIME magazine Person of The Year – "You". That is, TIME selected the masses of users who were participating in content creation on social networks, blogs, wikis, and media sharing sites. The cover story author Lev Grossman explains:
- It's a story about community and collaboration on a scale never seen before. It's about the cosmic compendium of knowledge Wikipedia and the million-channel people's network YouTube and the online metropolis MySpace. It's about the many wresting power from the few and helping one another for nothing and how that will not only change the world, but also change the way the world changes.
- Since that time, Web 2.0 has found a place in the lexicon; in 2009 Global Language Monitor declared it to be the one-millionth English word

Characteristics:

- Flickr, a Web 2.0 web site that allows its users to upload and share photos
- Web 2.0 websites allow users to do more than just retrieve information. They can build on the interactive facilities of "Web 1.0" to provide "Network as platform" computing, allowing users to run software-applications entirely through a browser. Users can own the data on a Web 2.0 site and exercise control over that data These sites may have an "Architecture of participation" that encourages users to add value to the application as they use it .
- The concept of Web-as-participation-platform captures many of these characteristics. Bart Decrem, a founder and former CEO of Flock, calls Web 2.0 the "participatory Web“ and regards the Web-as-information-source as Web 1.0.

Characteristics:

- The impossibility of excluding group-members who don't contribute to the provision of goods from sharing profits gives rise to the possibility that rational members will prefer to withhold their contribution of effort and free-ride on the contribution of others. This requires what is sometimes called Radical Trust by the management of the website. According to Best, the characteristics of Web 2.0 are: rich user experience, user participation, dynamic content, metadata, web standards and scalability. Further characteristics, such as openness, freedom and collective intelligence by way of user participation, can also be viewed as essential attributes of Web 2.0.
- **Search :**

  Finding information through keyword search. Links Connects information together into a meaningful information ecosystem using the model of the Web, and provides low-barrier social tools.
- **Authoring :**

  The ability to create and update content leads to the collaborative work of many rather than just a few web authors. In wikis, users may extend, undo and redo each other's work. In blogs, posts and the comments of individuals build up over time.
- **Tags :**

  Categorization of content by users adding "tags" - short, usually one-word descriptions - to facilitate searching, without dependence on pre-made categories. Collections of tags created by many users within a single system may be referred to as "folksonomies" (i.e., folk taxonomies).
- **Extensions :**

  Software that makes the Web an application platform as well as a document server.
- **Signals :**

  The use of syndication technology such as RSS to notify users of content changes.

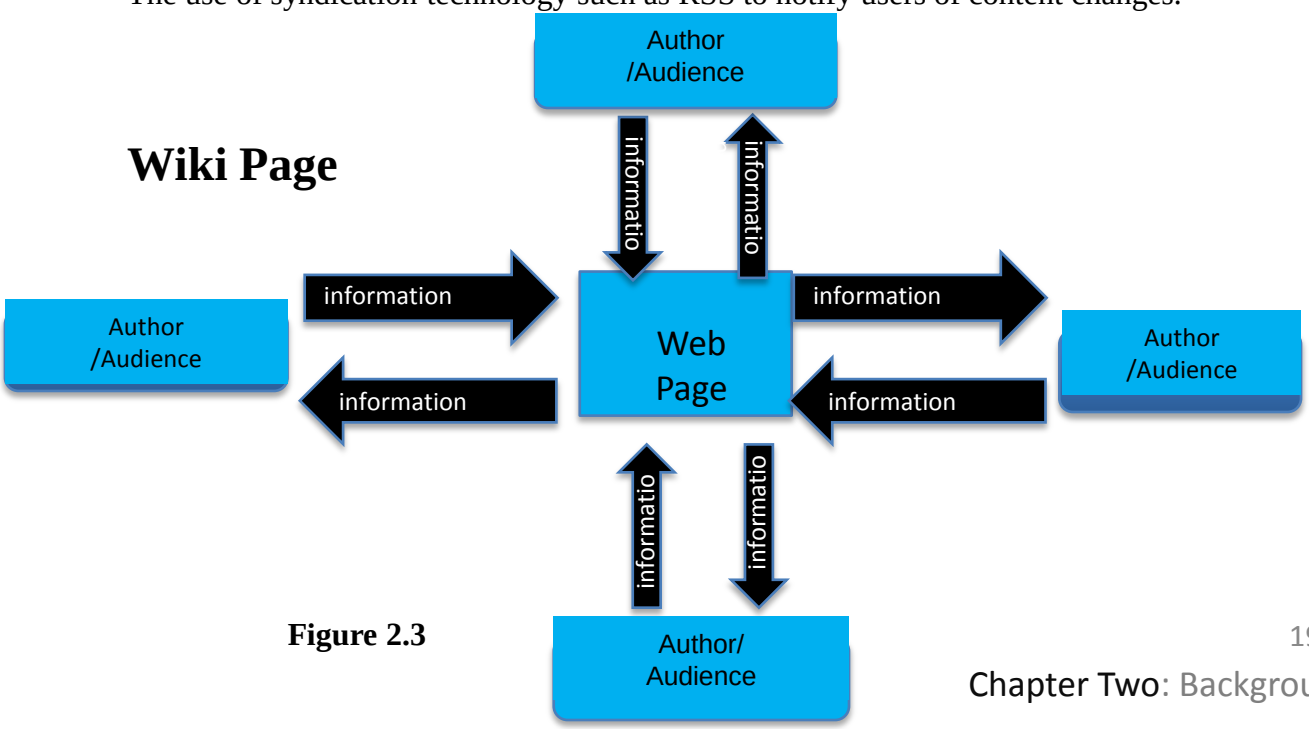


Figure 2.3

Web 2.0 Examples:

1. Blog
2. Wiki
3. Tagging
4. Social Bookmarking Sites
5. Social Networking
6. GEO World
7. Cloud Computing

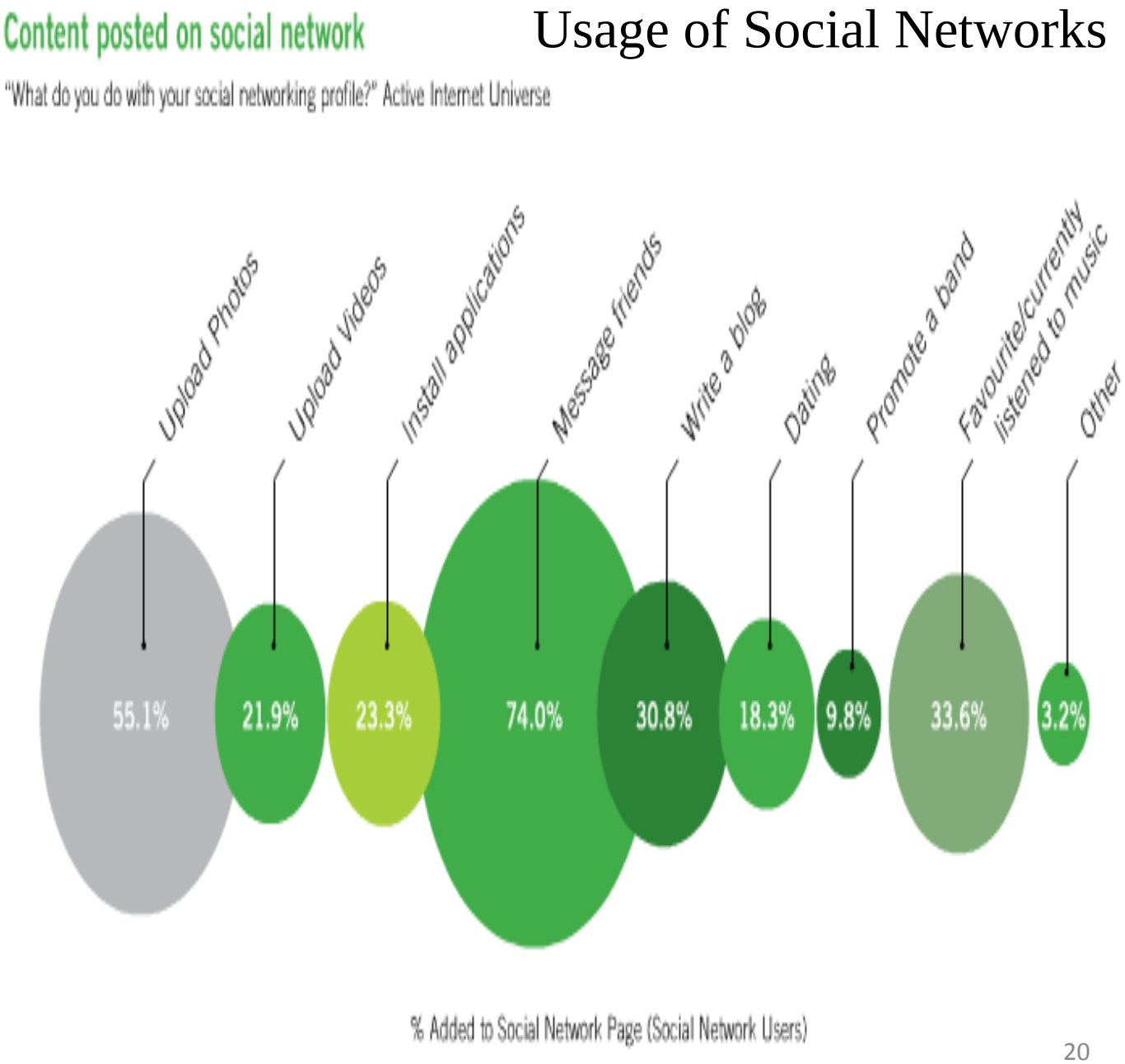


Figure 2.4

2.1.3. Web 3.0:

Overview:

- Not much time passed before "Web 3.0" was coined. Definitions of Web 3.0 vary greatly. Amit Agarwal states that Web 3.0 is, among other things, about the Semantic Web and personalization .Andrew Keen, author of The Cult of the Amateur, considers the Semantic Web an "unrealisable abstraction" and sees Web 3.0 as the return of experts and authorities to the Web. For example, he points to Bertelsman's deal with the German Wikipedia to produce an edited print version of that encyclopedia. CNN Money's Jessi Hempel expects Web 3.0 to emerge from new and innovative Web 2.0 services with a profitable business model.Conrad Wolfram has argued that Web 3.0 is where "the computer is generating new information", rather than humans.Others still such as Manoj Sharma, an organization strategist, in the keynote "A Brave New World Of Web 3.0" proposes that Web 3.0 will be a "Totally Integrated World" - cradle-to-grave experience of being always plugged onto the net.
- Rajnish Sharma (Systems Officer, UPTEC Computer Consultancy, Lucknow) believes that, "The next generation of the Web is Web 3.0, will make tasks like your search faster and easier. Instead of multiple searches, you might type a complex sentence or two in your Web 3.0 browser, and the Web will do the rest. The browser will analyze your response, search the Internet for all possible answers, and then organize the results for you. The Web 3.0 browser will act like a personal assistant. As you search the Web, the browser learns what you are interested in. The more you use the Web, the more your browser learns about you and the less specific you'll need to be with your questions.
- Eventually you might be able to ask your browser open questions like 'where should I go for lunch?' Your browser would consult its records of what you like and dislike, take into account your current location and then suggest a list of restaurants".
- Futurist John Smart, lead author of the Metaverse Roadmap echoes Sharma's perspective, defining Web 3.0 as the first-generation (convergence of the virtual and physical world), a web development layer that includes TV-quality open video, 3D simulations, augmented reality, human-constructed semantic standards, and pervasive broadband, wireless, and sensors. Web 3.0's early geosocial (Foursquare, etc.) and augmented reality (Layar, etc.) webs are an extension of Web 2.0's participatory technologies and social networks (Facebook, etc.) into 3D space. Of all its metaverse-like developments, Smart suggests Web 3.0's most defining characteristic will be the mass diffusion of NTSC-or-better quality open video to TVs, laptops, tablets, and mobile devices, a time when "the internet swallows the television."Smart considers Web 4.0 to be the Semantic Web and in particular, the rise of statistical, machine-constructed semantic tags and algorithms, driven by broad collective use of conversational interfaces, perhaps circa 2020.
- David Siegel's perspective in Pull: The Power of the Semantic Web, 2009, is consonant with this, proposing that the growth of human-constructed semantic standards and data will be a slow, industry-specific incremental process for years to come, perhaps unlikely to tip into broad social utility until after 2020.

Definition:

- „The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in co-operation.“ [Berners-Lee et al., 2001]
- The idea of the Semantic Web is to apply advanced knowledge technologies in order to fill the knowledge gap between human and machine.

Advantages of semantic web:

- Enrich the structure of the Web.
- Improve the quality of search.
- Enables applications to become more integrated and intelligent.
- Map the various data onto an abstract data representation.
  - make the data independent of its internal representation.
- Merge the resulting representations.
- Start making queries on the whole!
  - queries that could not have been done on the individual data sets.

The semantic graph (web 3.0) connects everything…

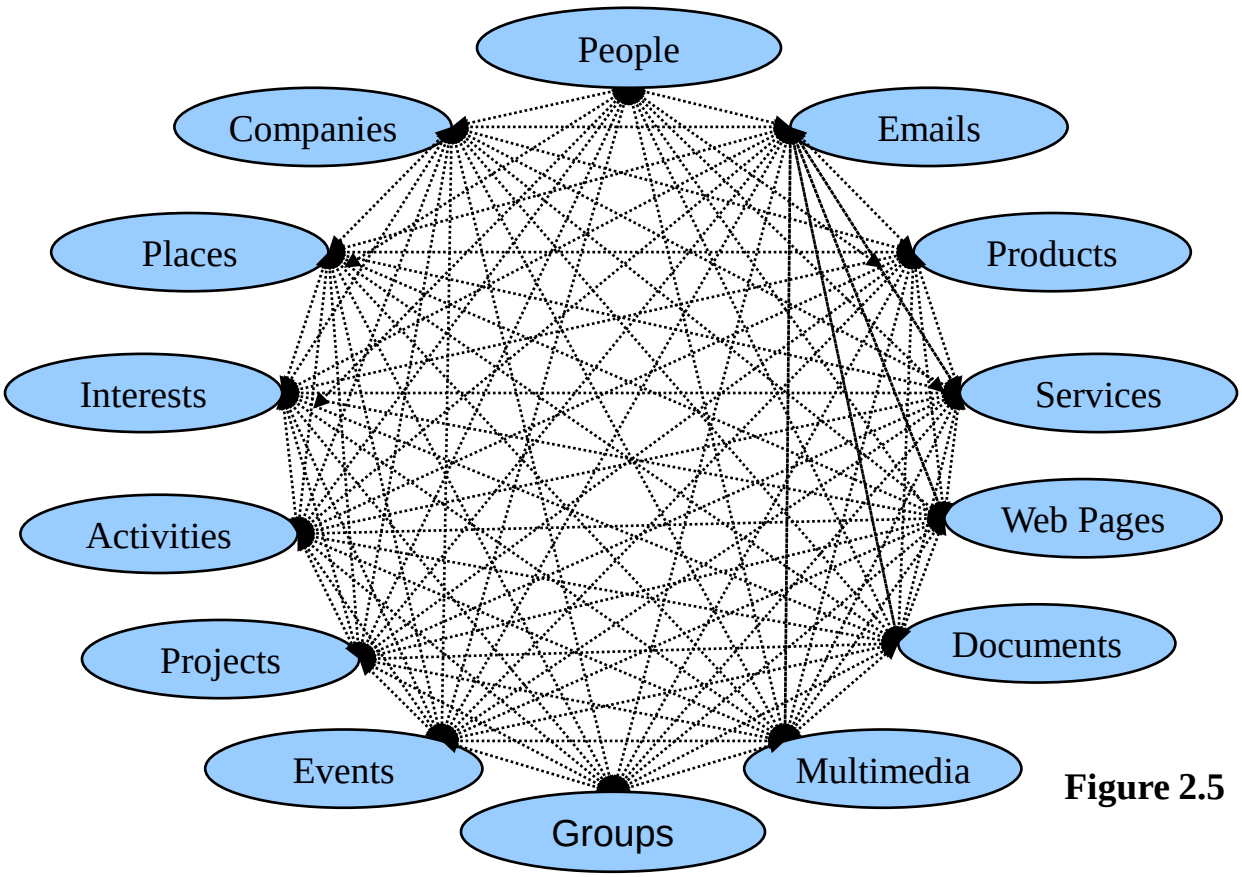


Figure 2.5

- Better search
- More targeted ads
- Smarter collaboration
- Deeper integration
- Richer content
- Better personalization

Web 1.0 vs. Web 2.0 vs. Web 3.0:

Web 1.0	Web 2.0	Semantic Web 3.0
Personal Websites	Blogs	Semantic Blogs
Content Management Systems	Wikis	Semantic Wikis
Altavista, Google	Google Personalised, DumbFind	Semantic Search Ex : Kngine
CiteSeer, Project Gutenberg	Google Scholar, Book Search	Social Semantic Digital Libraries
Message Boards	Community Portals	Semantic Forums and Community Portals
Buddy Lists, Address Books	Online Social Networks Ex : Facebook	Semantic Social Networks Ex : Facebook
-	-	Semantic Social Information Spaces

Figure 2.6

2.2. Semantic web concept:

- Semantic Web is a term coined by World Wide Web Consortium (W3C) director Sir Tim Berners-Lee [1]. It describes methods and technologies to allow machines to understand the meaning - or "semantics" - of information on the World Wide Web .
- According to the original vision, the availability of machine-readable metadata would enable automated agents and other software to access the Web more intelligently. The agents would be able to perform tasks automatically and locate related information on behalf of the user.
- While the term "Semantic Web" is not formally defined it is mainly used to describe the model and technologies proposed by the W3C [2]. These technologies include the Resource Description Framework (RDF), a variety of data interchange formats (e.g. RDF/XML, N3, Turtle, N-Triples), and notations such as RDF Schema (RDFS) and the Web Ontology Language (OWL), all of which are intended to provide a formal description of concepts, terms, and relationships within a given knowledge domain.
- Many of the technologies proposed by the W3C already exist and are used in various projects. The Semantic Web as a global vision, however, has remained largely unrealized and its critics have questioned the feasibility of the approach.

2.3. Semantic web technologies:

2.3.1. Web Ontology Language (OWL):

Definition:

- “Ontology is about the exact description of things and their relationships.” World Wide Web Consortium (W3C)
- An ontology is a way of declaring what types of things exist, and what types of relationships they have with each other.
- Provides common meaning to use in machine-understandable content.
- It is constituted by a specific vocabulary used to describe a certain reality (domain), plus
- a set of explicit assumptions regarding the intended meaning of the vocabulary.
- Thus, an ontology describes a formal specification of a certain domain
- Shared understanding of a domain of interest
- Formal and machine manipulable model of a domain of interest (telecoms systems, gene structures, public services)
- Ontology provide a shared and common understanding of a domain

(1) Sir Timothy John "Tim" Berners-Lee, OM, KBE, FRS, FREng, FRSA (born 8 June 1955, also known as "TBL"), is a British engineer and computer scientist and MIT professor credited with inventing the World Wide Web, making the first proposal for it in March 1989.On 25 December 1990, with the help of Robert Cailliau and a young student at CERN, he implemented the first successful communication between an HTTP client and server via the Internet.
(2) Berners-Lee is the director of the World Wide Web Consortium (W3C)

Domains and sub-domains:

In the beginning they designed a domains and sub-domains to have the same things in the same domain some how to understand not the exact meaning but the whole topic meaning that give us some of the semantics beside the old syntax ways …but those ways wasn't enough because

it don't give so correct answers .

it work on a specific domains .

it can't help in decision making or in recommendations .

Taxonomy :

Then and on the technology of domains they designed the 'Taxonomy.

the taxonomy was the way to give some semantics out of the syntax

but the taxonomy has no relations to help in recommendation.

Every child has only one parent then if we had a door then it's parent must be one we can't say that it may be a car's door or room's door .

they can't enable cross-application aggregation and value-added services .

they are not for sharing, not for finding because the taxonomy are designed for a program to have an information about it but we can't use it for other.

Taxonomy structure

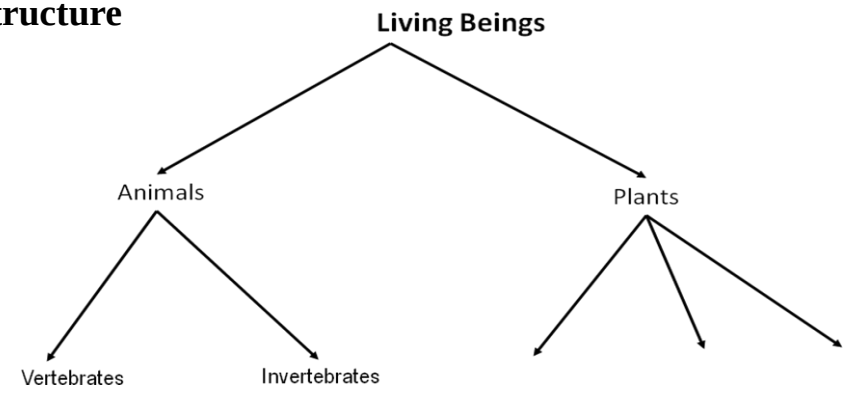


Figure 2.7

OWL uses:

- To define web resources more precisely and make them more amenable to machine processing
- To make domain assumptions explicit
  - Easier to change domain assumptions
  - Easier to understand and update legacy data
- To separate domain knowledge from operational knowledge
  - Re-use domain and operational knowledge separately
- A community reference for applications
- To share a consistent understanding of what information means

OWL structure:

Ontology structure

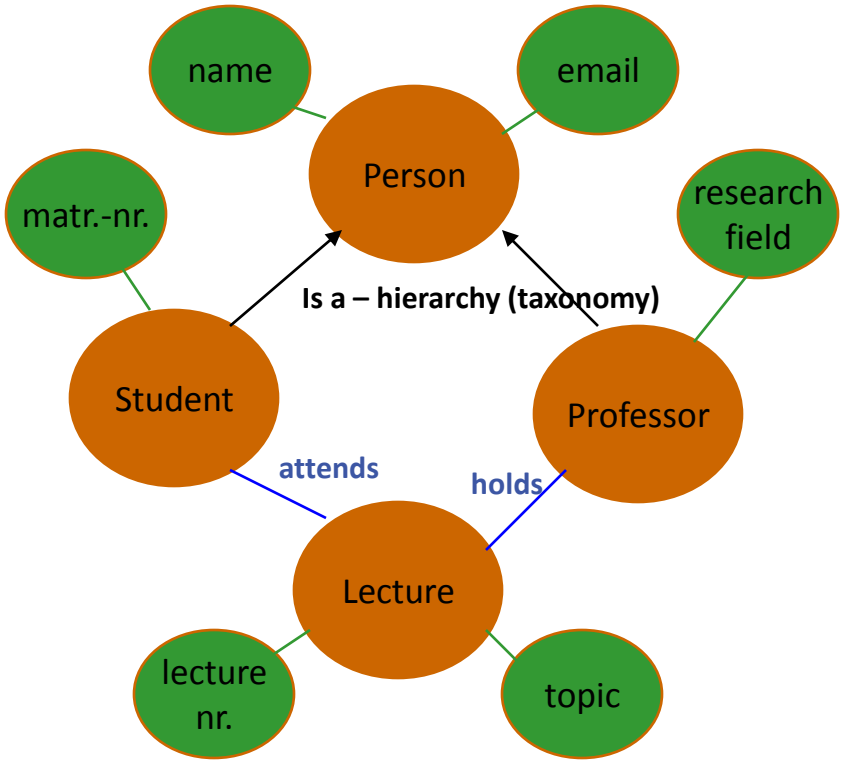


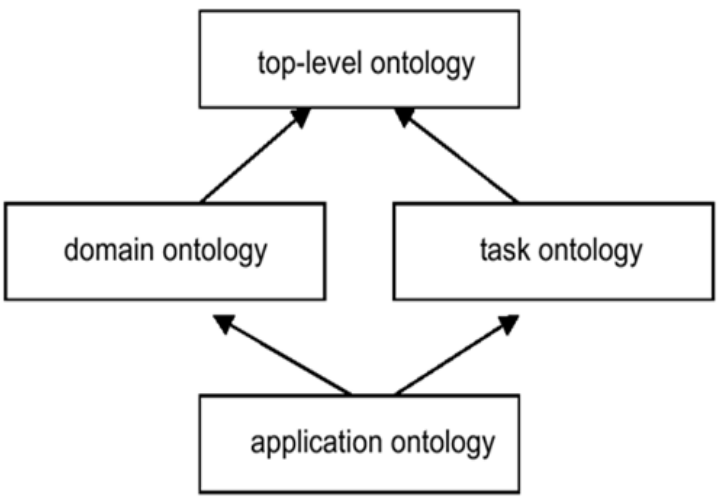
Figure 2.8

Types of Ontologies:

[Guarino, 98]

Describe **very general concepts** like space, time, event, which are independent of a particular problem or domain. It seems reasonable to have unified top-level ontologies for large communities of users.

Describe the vocabulary related to a **generic domain** by specializing the concepts introduced in the top-level ontology.

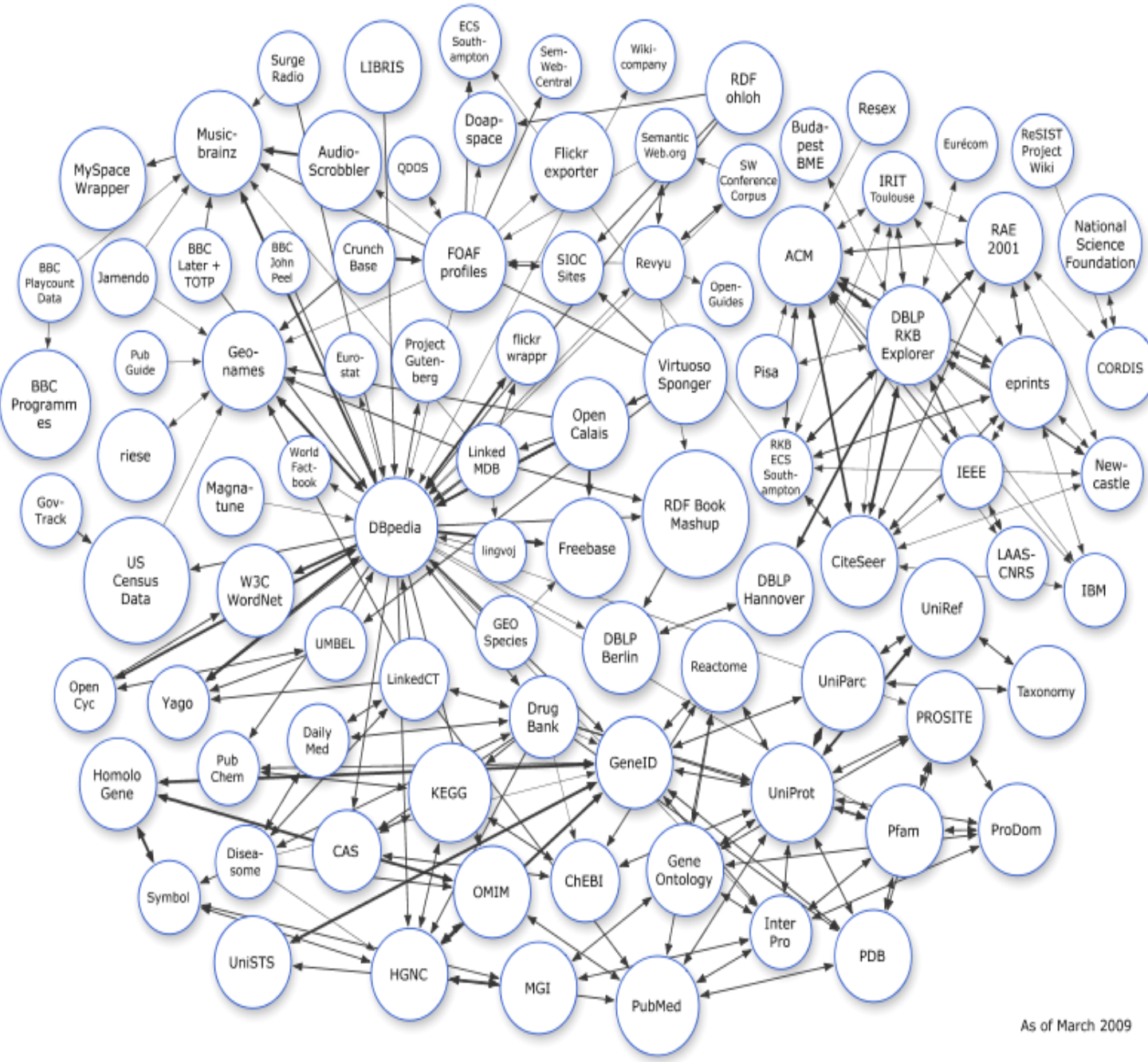


Describe the vocabulary related to a **generic task or activity** by specializing the top-level ontologies.

These are the most specific ontologies. Concepts in application ontologies often correspond to **roles played by domain entities while performing a certain activity**.

Figure 2.9

Available OWLs :



Linked data project [1] status as of march 2009

Figure 2.10

(1) The goal of the W3C Semantic Web Education and Outreach group's Linking Open Data community project is to extend the Web with a data commons by publishing various open datasets as RDF on the Web and by setting RDF links between data items from different data sources. In October 2007, datasets consisted of over two billion RDF triples, which were interlinked by over two million RDF links. By May 2009 this had grown to 4.2 billion RDF triples, interlinked by around 142 million RDF links. There is also an interactive visualization of the linked data sets to browse through the cloud.

2.3.2. RDF/XML:

XML definition:

- XML : stands for e**x**tensible **M**arkup **L**anguage .

- It is a Framework describes the structure of data (describes
- the syntax).
- It just a pure information wrapped in tags.
- XML tags are not predefined so you must define your own tags.
- XML is Case sensitive (All start tags must have end tags)

XML Example :

An E-mail: From: Ali
To : Mona
Subject: Reminder
Body: Don't forget me this weekend!

```
<?xml version="1.0"?>
<note>
 <to> Mona </to>
<from>Ali </from>
<heading>Reminder</heading>
<body>Don't forget me this weekend!</body>
</note>
```

RDF definition:

- RDF: Stands for **R**ecourse **D**escription **F**ramework.

- RDF is a framework designed for describing the resources on the web such as
  - Describing properties for shopping items (price and availability).
  - Describing information about web pages (content, author, created and modified date).
  - Describing content for search engines.
  - Describing content and rating for web pictures.
- The Resource Description Framework (RDF) is a family of World Wide Web Consortium (W3C) specifications originally designed as a metadata data model. It has come to be used as a general method for conceptual description or modeling of information that is implemented in web resources, using a variety of syntax formats.
- The XML language used by RDF is called RDF/XML.
- RDF documents are written inside XML.
- RDF/XML information can easily be exchange between different types of computers using different types of operating systems and applications.
- RDF/XML properties' value can easily be queried by SPARQLE or X query .

- RDFa (or Resource Description Framework – in – attributes) is a W3C Recommendation that adds a set of attribute level extensions to XHTML for embedding rich metadata within Web documents. The RDF data model mapping enables its use for embedding RDF triples within XHTML documents, it also enables the extraction of RDF model triples by compliant user agents.
- The W3C RDF in XHTML Taskforce is also working on an implementation for non-XML versions of HTML. The primary issue for the non-XML implementation is how to handle the lack of XML namespaces.

RDF uses:

- RDF merging information from multiple sources to provide better search engine capabilities.
- RDF describing the content and content relationships available at a particular Web site.
- <Provide Semantic (The meaning of data)>
- RDF is designed to be read and understood by computer applications not to be displayed on the web to public people.
- RDF with digital signatures will be key to building the "Web of Trust" for E-commerce, collaboration, and other applications.

RDF/XML Examples:

Simple RDFs:

1-

Resource	has property	value
abc.html	written by	Ahmed

```
<rdf:description rdf:about="http//www.abc.com.html">
<page:writer>Ahmed<page:writer>
<page:date>05-Mar-2007<page:date>
</rdf:description>
```

2-

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:dc="http://purl.org/dc/elements/1.1/">
<rdf:Description rdf:about="http://en.wikipedia.org/wiki/Tony_Benn"> <dc:title>Tony
        Benn</dc:title> <dc:publisher>Wikipedia</dc:publisher>
</rdf:Description>
</rdf:RDF>
```

Combining many RDFs:

1-

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:dc="http://purl.org/dc/elements/1.1/">
<rdf:Description rdf:about="http://en.wikipedia.org/wiki/Tony_Benn"> <dc:title>Tony
        Benn</dc:title> <dc:publisher>Wikipedia</dc:publisher>
</rdf:Description>
</rdf:RDF>
```

Essence:

- The essence of RDFa is to provide a set of attributes that can be used to carry metadata in an XML language (hence the 'a' in RDFa).
- These attributes are:
- **about** and **src** – a URI or CURIE specifying the resource the metadata is about
- **rel** and **rev** – specifying a relationship or reverse-relationship with another resource
- **href** and **resource** – specifying the partner resource
- **property** – specifying a property for the content of an element
- **content** – optional attribute that overrides the content of the element when using the property attribute
- **datatype** – optional attribute that specifies the datatype of text specified for use with the property attribute
- **typeof** – optional attribute that specifies the RDF type(s) of the subject (the resource that the metadata is about).

Example:

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML+RDFa 1.0//EN"
"http://www.w3.org/MarkUp/DTD/xhtml-rdfa-1.dtd">
 <html xmlns="http://www.w3.org/1999/xhtml" xmlns:foaf="http://xmlns.com/foaf/0.1/"
xmlns:dc="http://purl.org/dc/elements/1.1/" version="XHTML+RDFa 1.0" xml:lang="en">
<head>
<title>John's Home Page</title>
<base href="http://example.org/john-d/" />
<meta property="dc:creator" content="Jonathan Doe" />
<link rel="foaf:primaryTopic" href="http://example.org/john-d/#me" /> </head>
<body about="http://example.org/john-d/#me">
<h1>John's Home Page</h1>
<p>My name is
<span property="foaf:nick">John D</span>
and I like
<a href="http://www.neubauten.org/" rel="foaf:interest" xml:lang="de">Einstürzende
Neubauten</a>.
</p>
 <p> My <span rel="foaf:interest" resource="urn:ISBN:0752820907">
favorite book is the inspiring <span about="urn:ISBN:0752820907">
<cite property="dc:title">Weaving the Web</cite> by
<span property="dc:creator">Tim Berners-Lee</span>
</span>
 </span>
</p>
 </body>
 </html>
```

2.3.3. RDF Schema:

definition:

- RDF Schema (variously abbreviated as RDFS, RDF(S), RDF-S, or RDF/S) is an extensible knowledge representation language, providing basic elements for the description of ontologies, otherwise called Resource Description Framework (RDF) vocabularies, intended to structure RDF resources. The first version was published by the World-Wide Web Consortium (W3C) in April 1998, and the final W3C recommendation was released in February 2004. Many RDFS components are included in the more expressive language Web Ontology Language (OWL).

Properties:

- *Properties are instances of the class rdf:Property and describe a relation between subject resources and object resources. When used as such a property is a predicate.*
- *rdfs:domain of an rdf:predicate declares the class of the subject in a triple [1] whose second component is the predicate.*
- **rdfs:range** of an rdf:predicate declares the class or datatype of the *object* in a triple whose second component is the predicate.
- For example, the following declarations are used to express that the property ex:employer relates a subject, which is of type foaf:Person, to an object, which is of type foaf:Organization:
- *ex:employer rdfs:domain foaf:Person*
- *ex:employer rdfs:range foaf:Organization*
- Given the previous two declarations, the following triple requires that ex:John is necessarily a foaf:Person, and ex:CompanyX is necessarily a foaf:Organization:
- *ex:John ex:employer ex:CompanyX*
- **rdf:type** is a property used to state that a resource is an instance of a class.
- **rdfs:subClassOf** allows to declare hierarchies of classes.
- For example, the following declares that 'Every Person is an Agent':
- *foaf:Person rdfs:subClassOf foaf:Agent*
- Hierarchies of classes support inheritance of a property domain and range (see definitions in next section) from a class to its subclasses.
- **rdfs:subPropertyOf** is an instance of rdf:Property that is used to state that all resources related by one property are also related by another.
- **rdfs:label** is an instance of rdf:Property that may be used to provide a human-readable version of a resource's name.
- **rdfs:comment** is an instance of rdf:Property that may be used to provide a human-readable description of a resource.

(1) A triple consists of a subject, a predicate, and an object.

2.3.4. SPARQL:

Overview:

- SPARQL is one of the web Query language.
    - Like SQL for Databases.
- The goal of SPARQL : Query language for RDF.
- Most features of SPARQL:
    - Limit the number of returned results.
    - Remove duplicates.
    - Sort results.
    - Use data types and/or language tags when matching a pattern.
- SPARQL is usually used over the network.
    - SPARQL Protocol for RDF with HTTP and SOAP bindings.
- Big datasets usually prefer “SPARQL endpoints” protocol.
- For Example: SPARQL endpoint to Dbpedia.

Definition:

SPARQL (pronounced "sparkle”) is an RDF query language; its name is a recursive acronym that stands for *SPARQL Protocol and RDF Query Language*. It was standardized by the *RDF Data Access Working Group* (DAWG) of the World Wide Web Consortium, and is considered a key semantic web technology. On 15 January 2008, SPARQL became an official W3C Recommendation.

SPARQL allows for a query to consist of triple patterns, conjunctions, disjunctions, and optional patterns.

Implementations for multiple programming languages exist.["SPARQL will make a huge difference" according to Sir Tim Berners-Lee in a May 2006 interview.

Examples:

1-

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
 SELECT ?name
WHERE { ?person foaf:name ?name . }
```

2-

```
PREFIX abc: <http://example.com/exampleOntology#>
 SELECT ?capital ?country
 WHERE { ?x abc:cityname ?capital ;
abc:isCapitalOf ?y . ?y abc:countryname ?country ;
 abc:isInContinent abc:Africa . }
```

Semantics and Business:

Now a days the semantic technology is important in business , A lot of sites had maximized its profit by using it in their business model as:

- **Semantics for BestBuy.com**
- **Semantics at Salesforce.com**
- **Semantics in Digital Libraries**
- **Semantics at Nokia**
- **Semantics for Energy Management**
- **Semantics for Motion Pictures and Video Industry**

Case study "Best Buy " a new way to compete:

Overview:

- We explore how a leading U.S. retailer, Best Buy, is using a Semantic Web markup language called RDFa to add semantics to its webpages.
- This is not just an academic exercise for Best Buy. As we will see, semantic technology has already led to increased traffic and better service to its customers.

keys of success:

- GoodRelations Ontology is a standardized vocabulary for product, price, and company data that increases the visibility of your products and services in the latest generation of search engines .
- GoodRelations + RDFa improved the rank of the respective pages in Google
- 30 % percent (!) increase in traffic.

Objective:

- The primary goal of using semantic technologies was to increase the visibility of its products and services. And with data such as store name, address, store hours and GEO data being marked up using RDFa, search engines are now able to identify each of those data components more easily and put them into context.
- A quick refresher on the terminology: just as the lingua franca of the Web is HTML (Hypertext Markup Language), RDF (Resource Description Framework) is commonly thought of as the primary language of the Semantic Web. RDFa is a kind of 'lite' version of RDF, which adds metadata to HTML (or XHTML) webpages.
- The process of adding RDFa to Best Buy's webpages began two years ago, when the company began to look for ways to get more visibility to its stores on the Web.

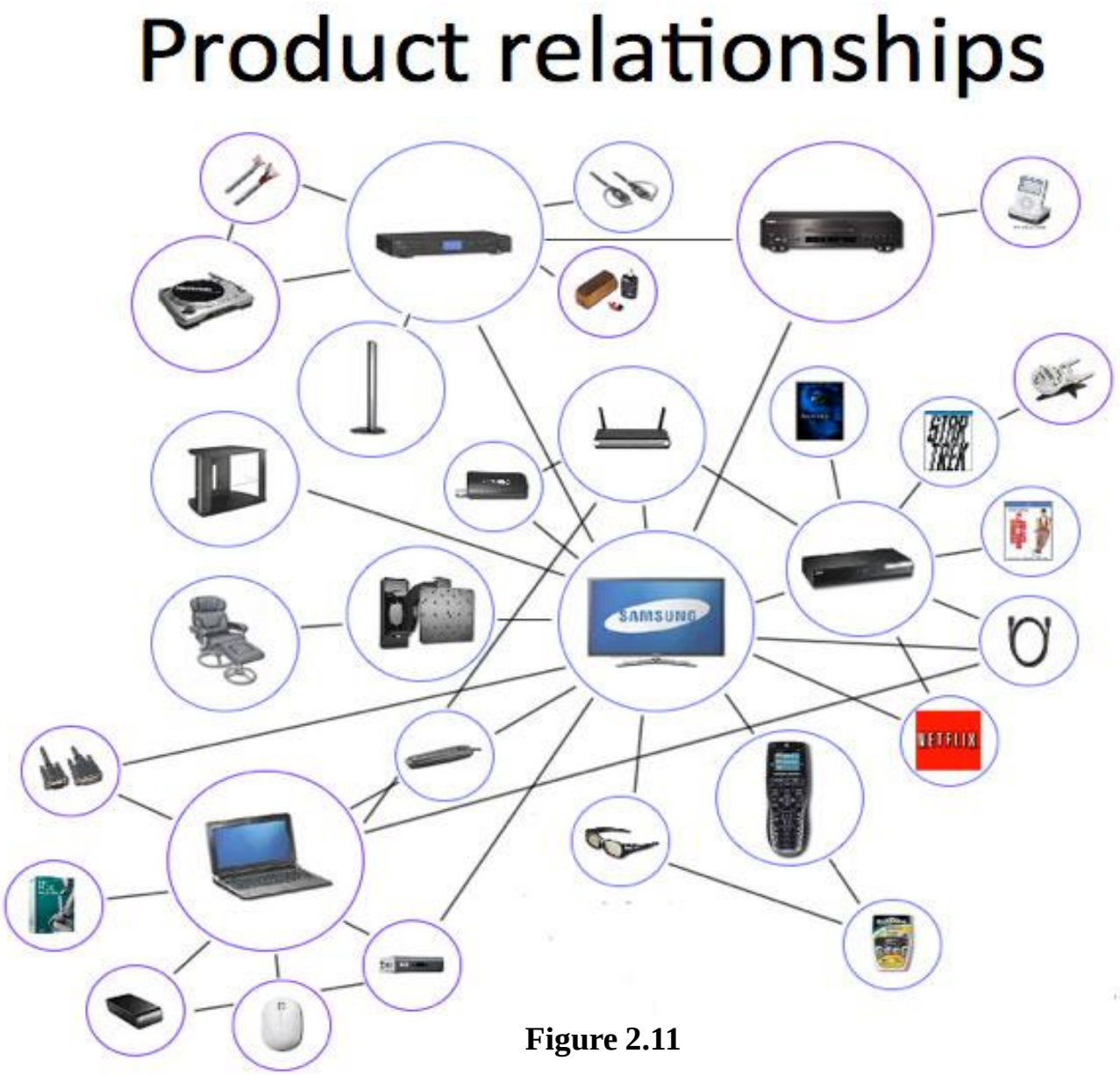
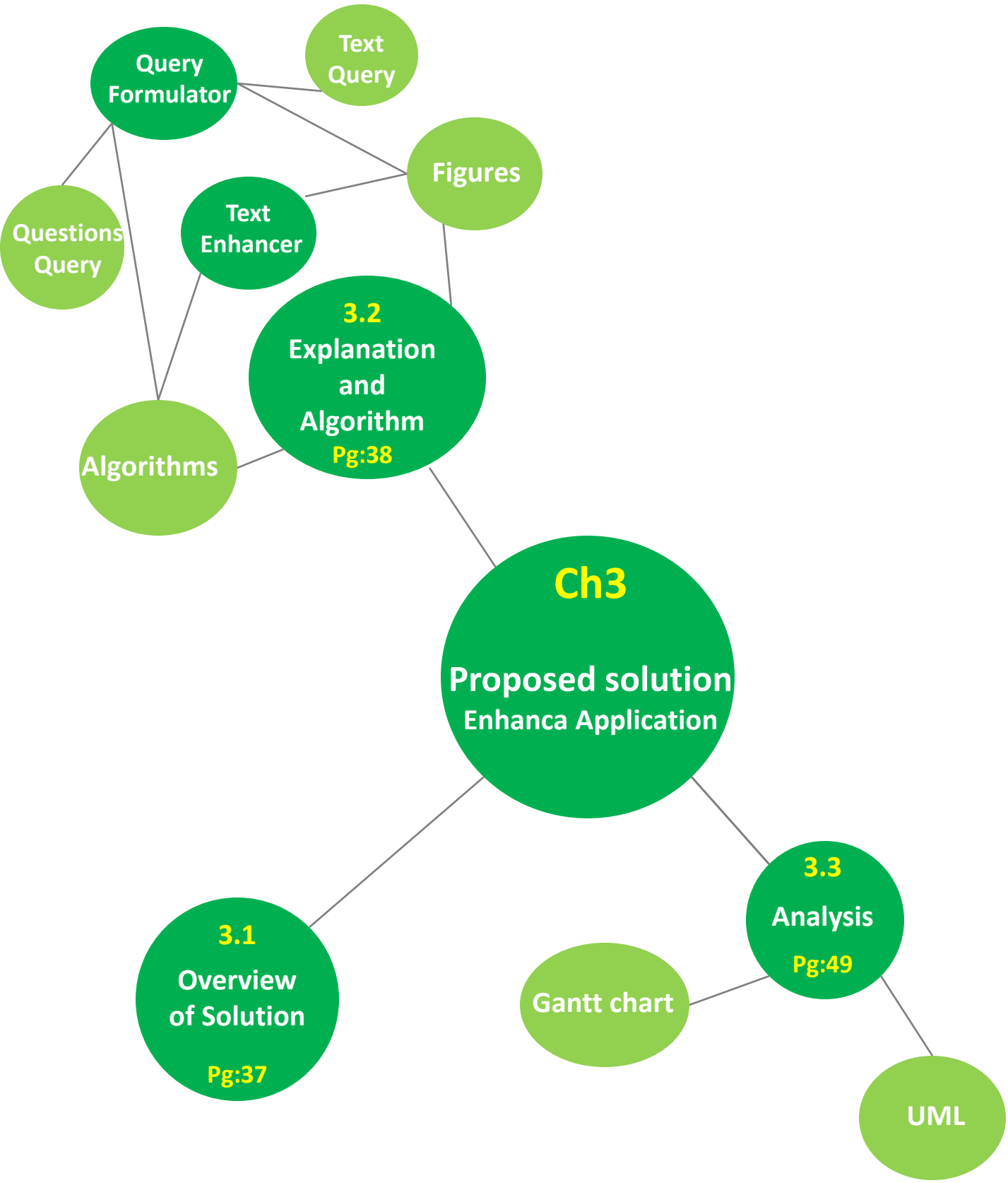


Figure 2.11

Chapter Three: Proposed solution



3.1. Overview of Proposed solution:

- An internet surfer may need more data about some subject he wondering about, maybe an article he is reading.
- So we provide him with one tool to get better understanding of the texts he reads or questions he have throw our system's semantic based components and our APIs with the semantic and non-semantic web services.

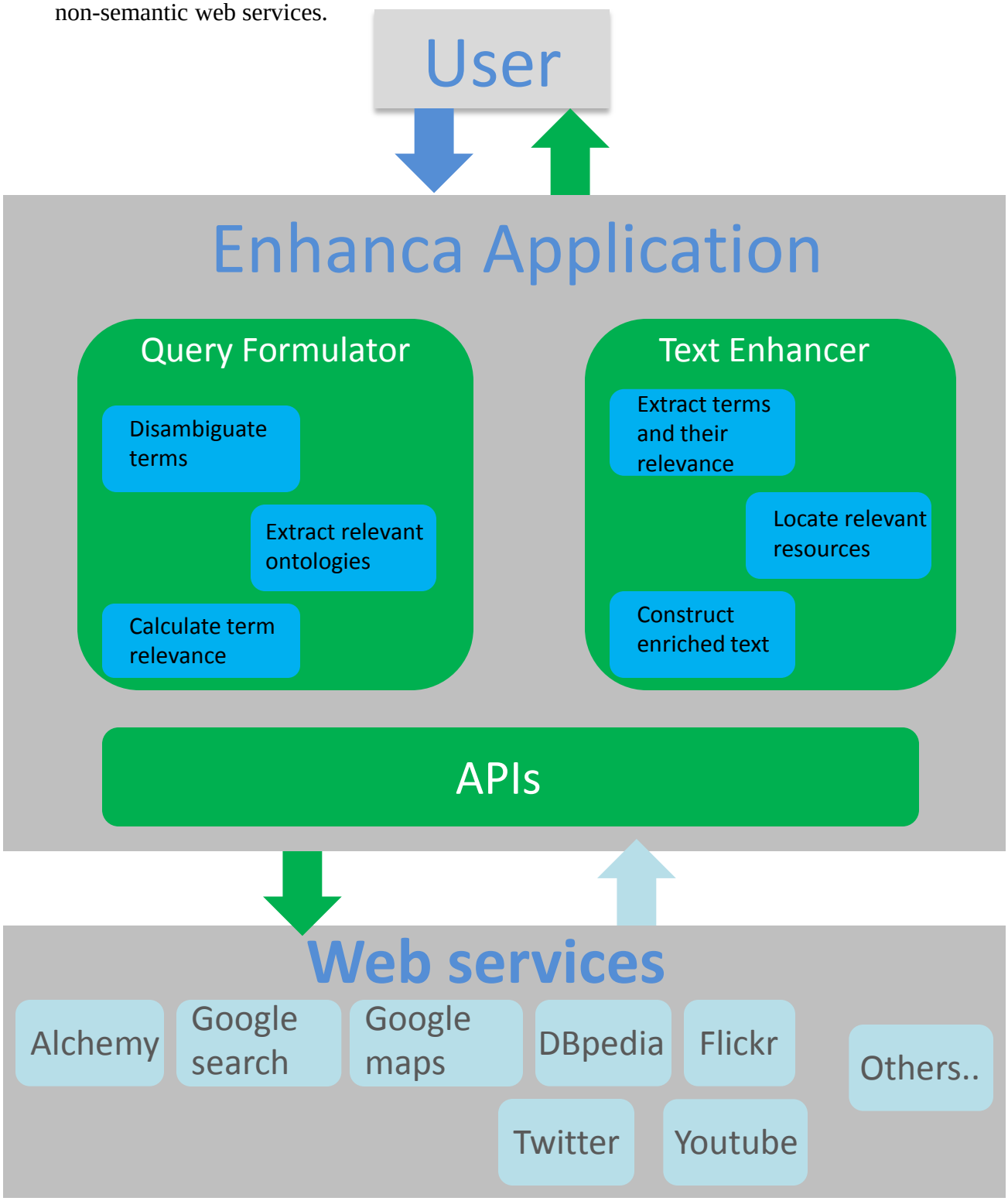


Figure 3.1
Proposed System Architecture

3.2. Explanation and Algorithm:

3.2.1. Text query:

Overview:

- An internet surfer may need more data about some subject he wondering about, maybe an article he is reading.
- So we provide him with one tool to get better understanding of the texts he reads or questions he have throw our system's semantic based components and our APIs with the semantic and non-semantic web services.

Algorithm:

Step 1 :

Extracting key words from entered text by user.

Step 2 :

choosing highest five keywords among the set we had from step 1.

Step 3:

Determining the available keywords types, then calculating the highest tow keywords from every type according to their relevance and count properties.

Step 4 :

Multiplying the repeated keywords' relevance in step 2 & 3 by 1.6.

Step 5 :

if "k-word's count=1
Multiplying the keyword 's relevance by 0.9
if "k-word's count=2
Multiplying the keyword 's relevance by 1.0
if "k-word's count=3
Multiplying the keyword 's relevance by 1.1
if "k-word's count=4
Multiplying the keyword 's relevance by 1.2
if "k-word's count=5
Multiplying the keyword 's relevance by 1.3
if "k-word's count>5
Multiplying the keyword 's relevance by 1.5

Step 6 :

Query for the keywords' synonymous and if there is keywords that are synonymous of each other then we keep the one with the highest relevance and add the relevance of the other keywords to it's relevance.

Step 7:

Building a table that containing the keyword 's (type, count, old and calculated relevance).

Step 8:

Calculation of the simple average of the calculated relevance of the keywords

$\mu = \sum$ (C-relevance) / number of keywords.

Step 9:

Keep the keywords if their calculated relevance more than the simple average

in step 8.

Step 10:

Simply ,all the words have relevancies more than (the average * "1.2") we will keep them .

Step 11:

The words that had lower relevancy we compare between the old and the new relevancy if it's new relevancy is smaller than the old one then we will delete it but if it was equallor morethenwe multiply the old relevancy by "1.1" if the new relevancy now is more than or equal the old one after re-calculate it then we keep it else then delete the word from the table .

Step 12:

Query for the relationships between the final keywords.

Step 13:

If we had a relations like (is a, sort of ,type of) We multiply the calculated relevance of the parent class keyword by 1.1.

Step 14:

Start queering about the keywords with the highest calculated relevance .
Dbpedia by SPARQL .
Google by converting operators into ASKII.

Note : we can use the graphical database to build our ontology out of these relation ships.

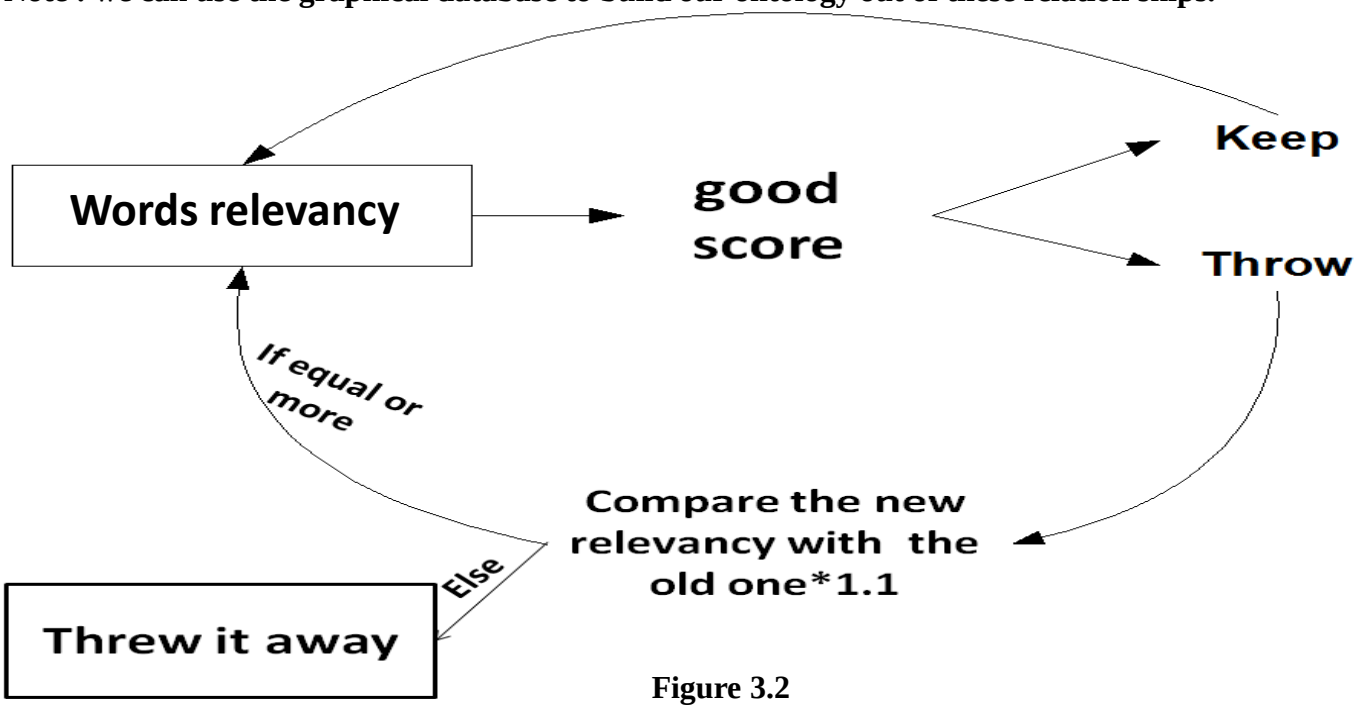


Figure 3.2

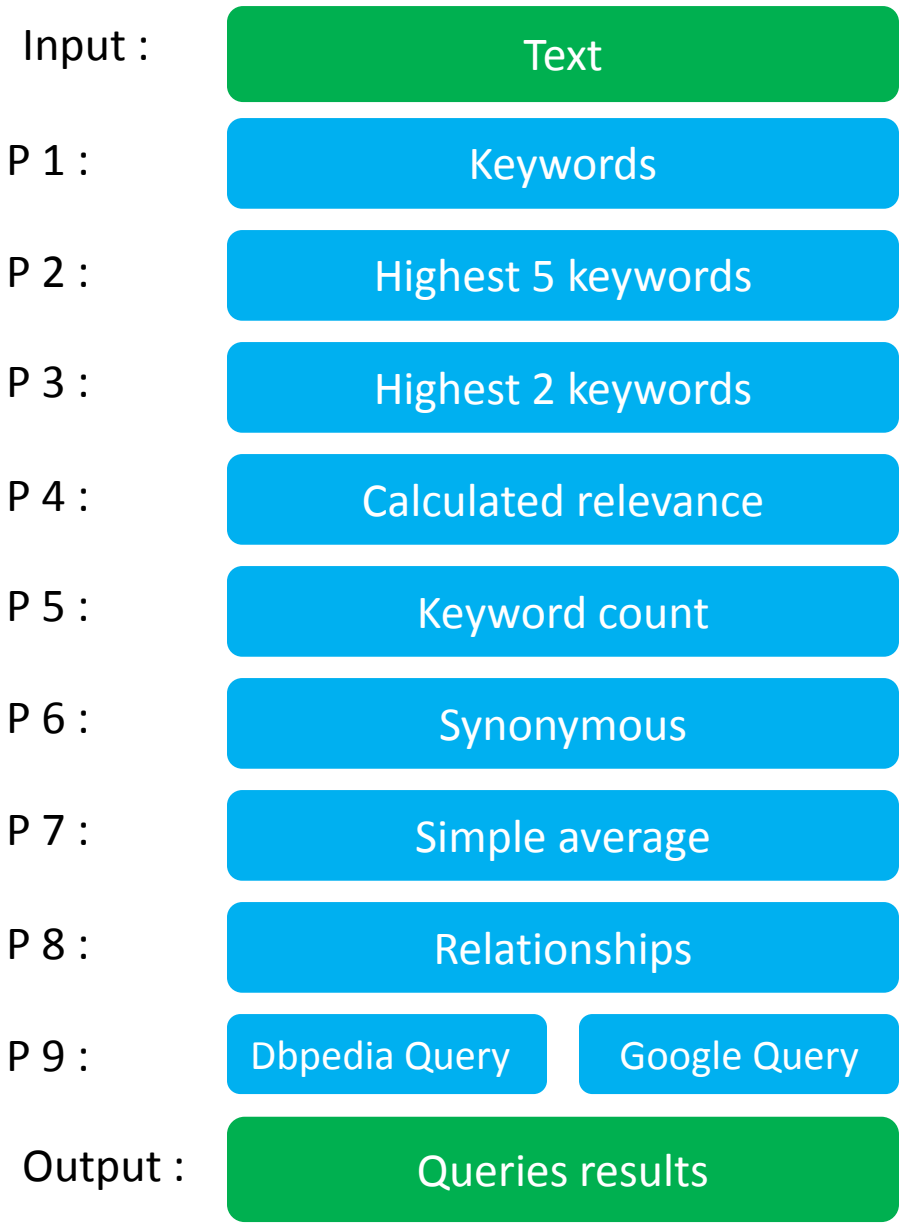

Text query Architecture
Input :
Text
P 1 :
Keywords
P 2 :
Highest 5 keywords
P 3 :
Highest 2 keywords
P 4 :
Calculated relevance
P 5 :
Keyword count
P 6 :
Synonymous
P 7 :
Simple average
P 8 :
Relationships
P 9 :
Dbpedia Query
Google Query
Output :
Queries results

Figure 3.3

3.2.2.Questions query:

Overview:

- An internet surfer may need more data about some subject he wondering about, maybe an article he is reading.
- So we provide him with one tool to get better understanding of the texts he reads or questions he have throw our system's semantic based components and our APIs with the semantic and non-semantic web services.

Algorithm:

Step 1:
Removing stop words in order to get the main elements of the question.

Step 2:
Extract the Wh. questions using the PoS Algorithms.

Step 3:
Extract the verbs and determine it's possible answer.

Step 4:
Extract the noun and determine it's type using Alchemy API .

Step 5:
Converting Alchemy types into DBpedia types.

Alchemy type	Dbpedia type
Person	Person
Continent, Country, City, State or county, geographical feature, region	Place
organization, radio station, company, TV station, facility	Organization
movies, music group, printed media, radio program, television show	Work

Figure 3.4

Step 6:
Constructing the Dbpedia SPARQL Queries according to the following Rules

Question Format			RDFa property To Query for
Wh. Question	Noun's DBpedia type	Verb's expected Answer	
Which	Person	prize, award, reword	Award
Which	Person	Date of birth	Birth date
Which	Person	Date of death	Death date
Which	Person	Era	Era
Which	Person	place, country, city Of birth	Birth place
Which	Person	place, country, city Of death	Death place
Which	Person	Area of living	Area
Which	Person	university, college	Education
Which	Person	school, high school	High school (or) Education
Which	Person	Company, organization	Company (or) Employer (or) Developer
Which	Person	Album	Album
Which	Person	Battle, fight, combat	Battle
Which	Person	Language, mothertounge	Language
When	Person	Date of birth	Birth date
When	Person	Date of death	Death date
When	Person	Body discovered date	Body discovered
When	Person	Date of election	Election date

Step 6:
Constructing the Dbpedia SPARQL Queries according to the following Rules

Question Format			RDFa property To Query for
Wh. Question	Noun's DBpedia type	Verb's expected Answer	
Where	Person	place of birth	Birth place
Where	Person	place of death	Death place
Where	Person	almamater	almamater
Where	Person	Date of election	Election date
Where	Person	high school	high school
Where	Person	employer	employer
Where	Person	occupation	occupation
Where	Person	Area ,city ,country	area / city / country
Where	Person	state of origin	state of origin
Where	Person	-	-
Who/Whom	Person	Child	Child
Who/Whom	Person	relation	relation
Who/Whom	Person	spouse	spouse
Who/Whom	Person	artist	artist
Who/Whom	Person	relative	relative discovered
How	Person	Death	deathcause
How old	Person	-	Birth date
How many	Person	Childes he have	child
How long	Person	Ruling period	election date

Step 6:
Constructing the Dbpedia SPARQL Queries according to the following Rules

Question Format			RDFa property To Query for
Wh. Question	Noun's DBpedia type	Verb's expected Answer	
How many	Place	visitors	Number of visitors
How many	Place	states	state
How many	Organization	Employee number	Number of employees
How many	Organization	staff	Number of staff
How many	Organization	volunteers	Number of volunteers
How many	Organization	office holder	Minority leader
How many	Organization	visitor	Number of visitors
How many	Organization	employees in the team	Team size
How many	Organization	launches	Total Launches
How many	Organization	failed launches	partial failed Launches
How many	Organization	successful launches	Successful Launches
How old	Organization	-	Foundation year
How old	Work	-	**completion date - release date**

Proposed rules to deal with the questions

Note :
Ruling period = serverdate – election date
Ruling period = end ruling date – election date

Step 7:

Constructing the Dbpedia SPARQL Queries.

Data

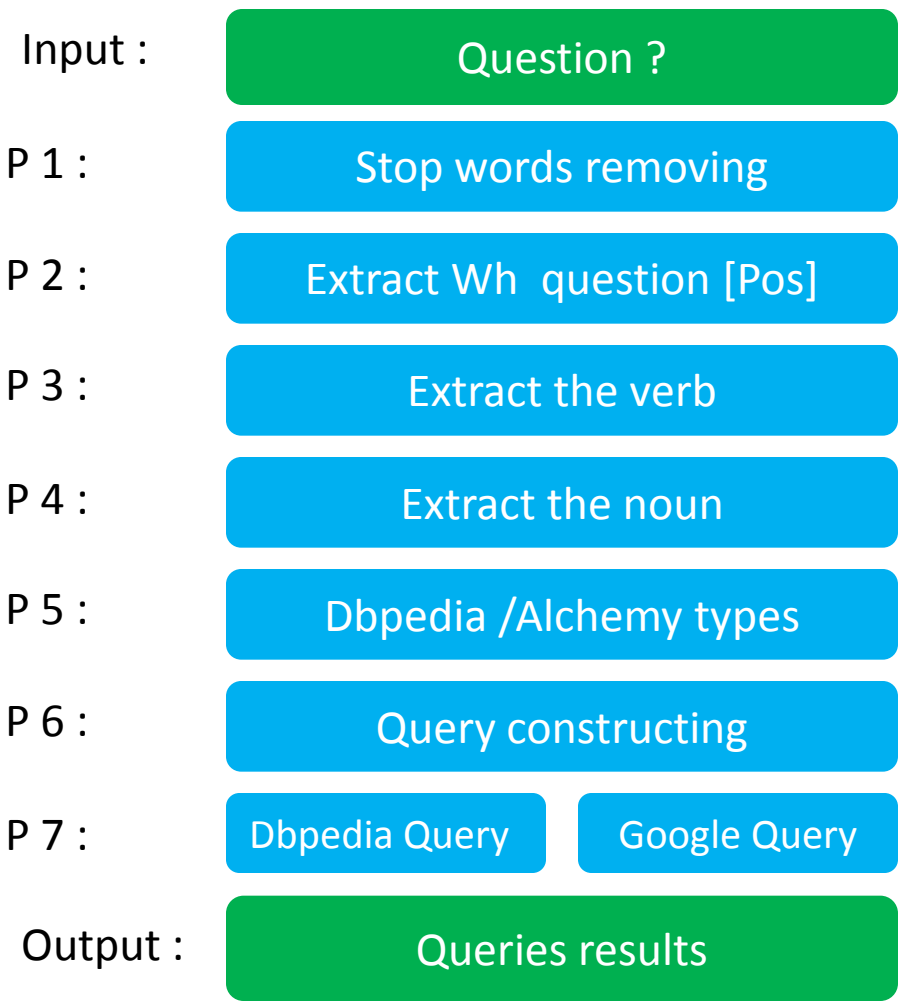
```
@prefix foaf:
<http://xmlns.com/foaf/0.1/> .
 _:a foaf:name "Johnny Lee Outlaw"
 _:a foaf:mbox
<mailto:jlow@example.com> .
 _:b foaf:name "Peter Goodguy" .
 _:b foaf:mbox
<mailto:peter@example.org> .
_:c foaf:mbox
<mailto:carol@example.org> .
```

Query

```
PREFIX foaf:
<http://xmlns.com/foaf/0.1/>
SELECT ?name ?mbox
WHERE { ?x foaf:name ?name .
?x foaf:mbox ?mbox }
```

Query results

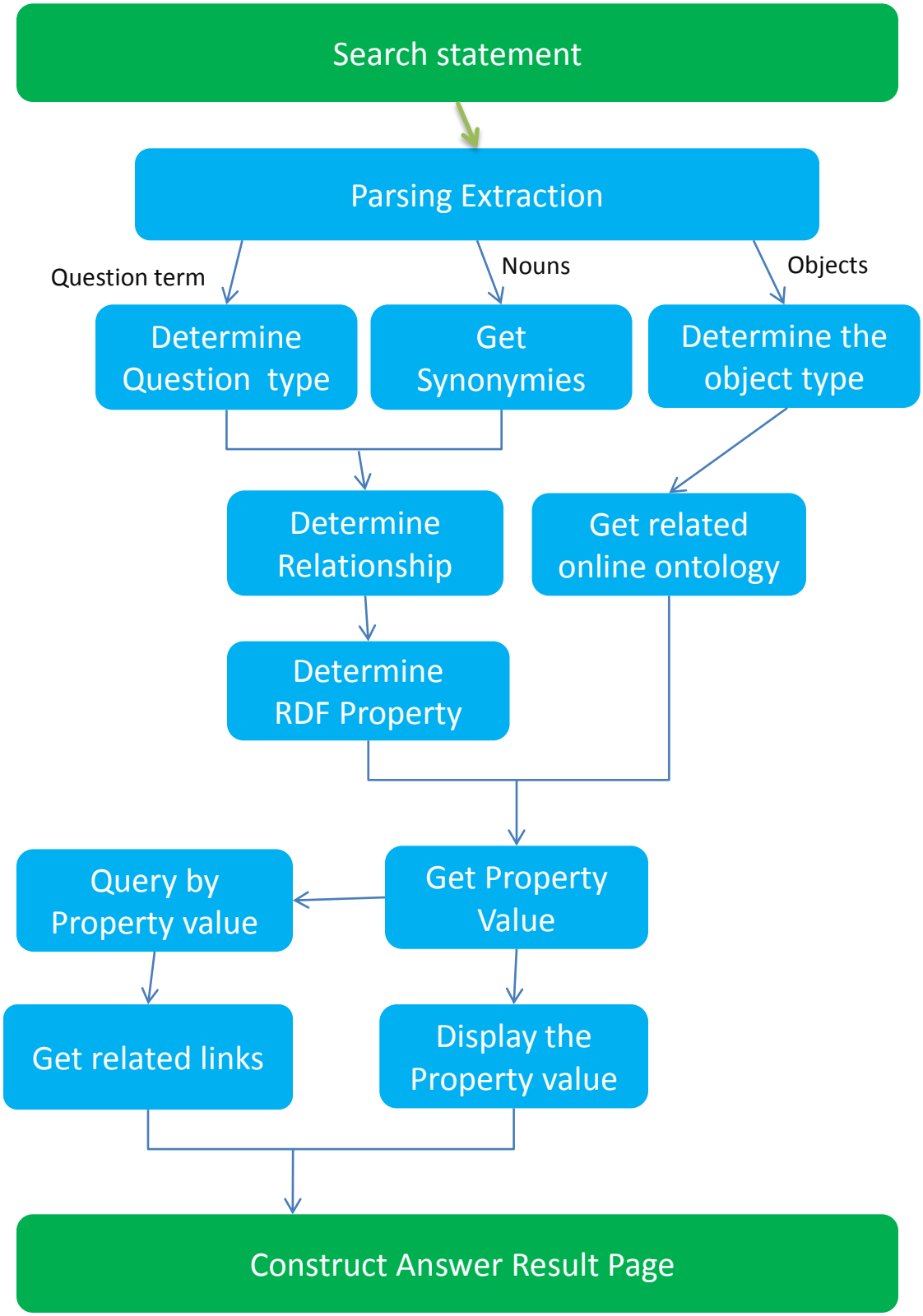
name	inbox
"Johnny Lee Outlaw"	<mailto:jlow@example.com>
"Peter Goodguy"	<mailto:peter@example.org>



Question query Architecture

Figure 3.5

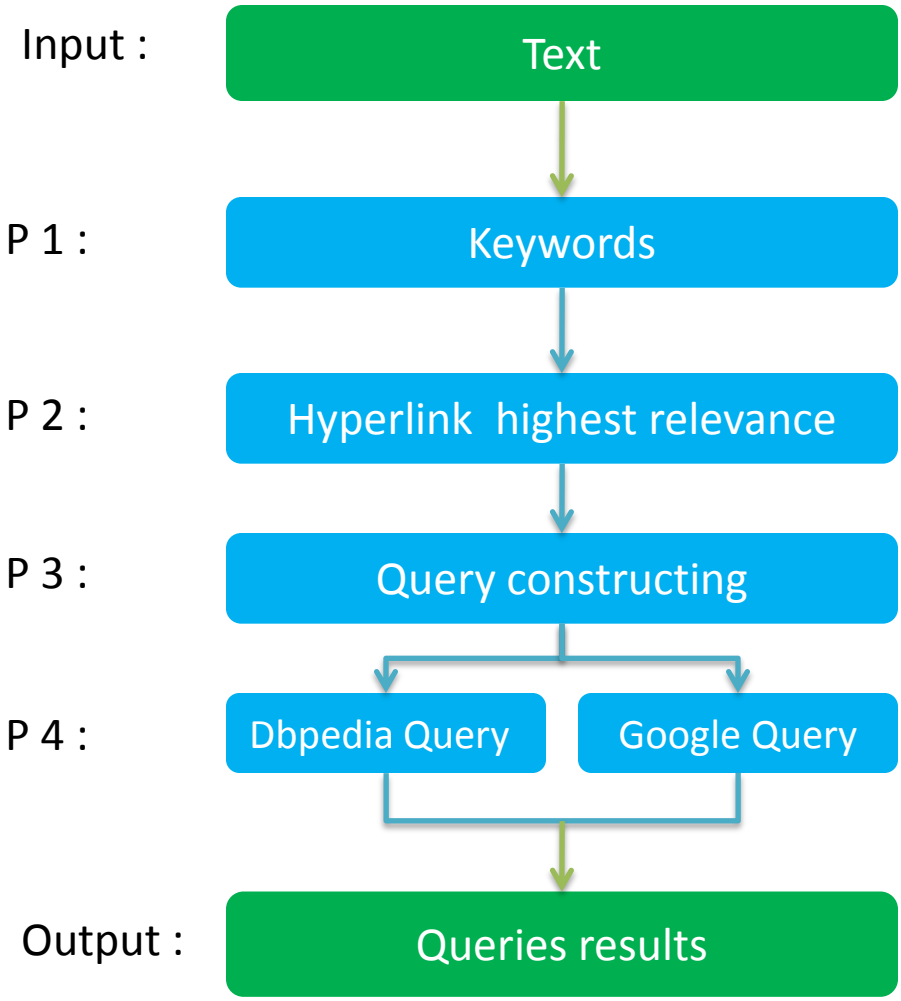
Dealing With user questions



Question query Architecture detailed

Figure 3.6

3.2.3. Text Enhancing



Text Enhancing Architecture

Figure 3.7

Overview:

- An internet surfer may need more data about some subject he wondering about, maybe an article he is reading.
- So we provide him with one tool to get better understanding of the texts he reads or questions he have throw a variation of information(e.g. pictures, news, twitters, facebook, maps, videos and more).

Algorithm:

Step 1:
Extracting keywords from text entered by user.
Step 2:
Hyper linking highest keywords among the set we have fromstep 1 according to their relevance value .
Step 3:
Constructing the query by sending the keyword and it's type.
Step 4:
Displaying the variety of information found in using proper shape.

Those are terms that we can use in addition to many other terms we won't need to use as (Organization, Company, Radio program, Radio station, TV show, TV station, Print media, Entertainment award, Facility and Holiday)

Note: every term can use all those APIs but in our humble point of view only the stated ones will maintain a best value for the user, but if anyone want to make changes it will be ok.

APIs / Terms Types	Question	Expected Answer	Wikipedia	Google Image	Google Map	You Tube	Google News	Flickr	TwitterA
Continent	What	Define	√	√	√	√		√	
	Where	Find place	√	√	√			√	
	When	Events	√	√	√	√		√	√
	How "								
Stat or Country	What	Define	√	√	√	√	√	√	√
	Where	Find place	√	√	√			√	
	Who	Person	√	√		√	√	√	√
	Why	Reason	√				√		√
	When	Events	√	√		√	√	√	√
	How"								
	Whom"								
	Whose"								
Country	What	Define	√	√	√	√	√	√	√
	Where	Find place	√	√	√			√	
	Who	Person	√	√		√	√	√	√
	Why	Reason	√				√		√
	When	Events	√	√		√	√	√	√
	How"								
	Whom"								
	Whose"								
City	What	Define	√	√	√	√	√	√	√
	Where	Find place	√	√	√			√	
	Who	Person	√	√		√	√	√	√
	Why	Reason	√				√		√
	When	Events	√	√		√	√	√	√
	How"								
	Whom"								
	Whose"								
Region	What	Define	√	√	√	√	√	√	√
	Where	Find place	√	√	√			√	
	Who	Person	√	√		√	√	√	√
	Why	Reason	√				√		√
	When	Events	√	√		√	√	√	√
	How"								
	Whom"								
	Whose"								
Person	What = who	Define	√	√		√	√	√	√
	Where	Find place	√	√	√		√	√	
	Why	Reason	√			√	√		√
	When	Events	√	√		√	√	√	√
	How								
	Whom								
	Whose								
Movie			√	√		√	√	√	√
Natural Disaster			√	√		√	√	√	√
Sport			√	√		√	√	√	√
Geographic Feature			√	√				√	√
Health Condition			√	√				√	
Terminology			√	√				√	√
Technology			√	√		√	√	√	√

Figure 3.8

3.3. Analysis:

3.3.1. Gantt chart:

Tasks distribution and progress monitoring.

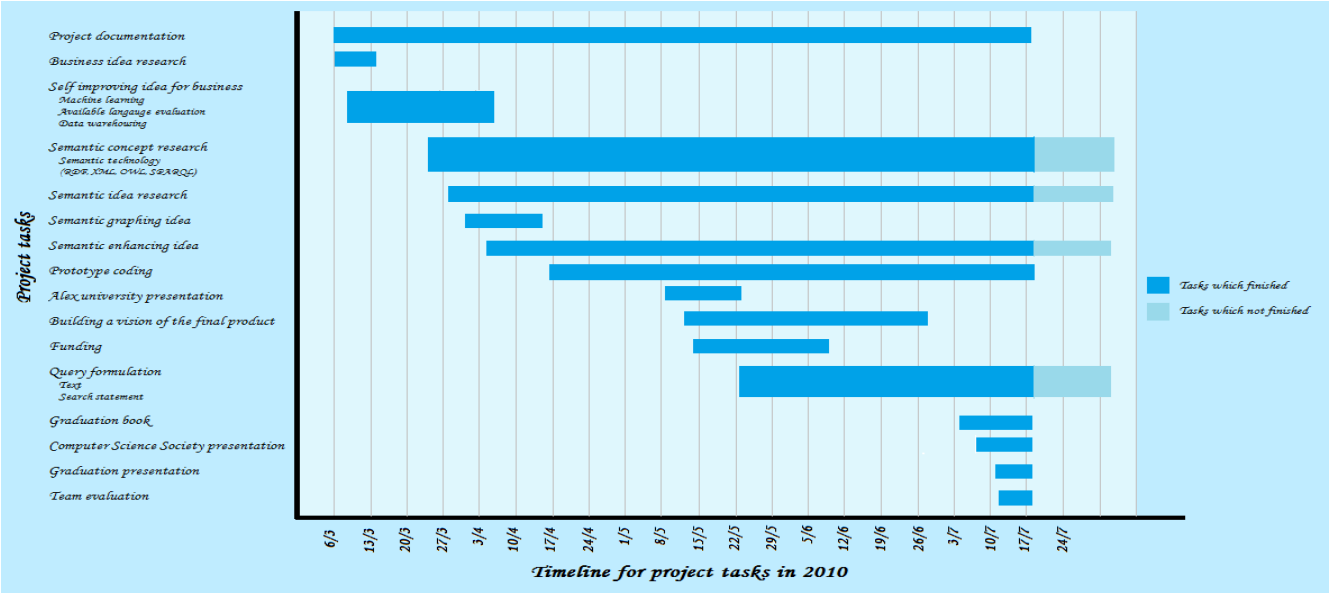


Figure 3.9

Time Effective Project Management (Detailed Gant chart)

Number	Task	Resource	Start	End	Duration	% Complete
1	Project documentation		6/3/2010	16/7/2010	95	
2	Business idea research		6/3/2010	13/3/2010	5	
3	Machine learning		9/3/2010	23/3/2010	11	
4	Available language evaluation		9/3/2010	23/3/2010	11	
5	Data warehousing		9/3/2010	23/3/2010	11	
6	Semantic concept research		24/3/2010	24/7/2010	88	
7	semantic idea research		28/3/2010	24/7/2010	85	
8	semantic graphing idea		1/4/2010	15/4/2010	11	
9	Semantic enhancing idea		5/4/2010	24/7/2010	80	
10	Prototype coding		8/4/2010	24/7/2010	77	
11	Alex university presentation		9/5/2010	23/5/2010	10	
12	building a vision of the final product		12/5/2010	24/6/2010	32	
13	Funding		14/5/2010	9/6/2010	19	
14	Query formulation for text		23/5/2010	24/7/2010	45	
15	Query formulation for search statement		23/5/2010	24/7/2010	45	
16	Graduation book		5/7/2010	9/7/2010	5	
17	Computer Science Society (CSS) presentation		7/7/2010	21/7/2010	11	
18	Team evaluation		11/7/2010	18/7/2010	5	
19	Graduation presentation		12/7/2010	18/7/2010	5	

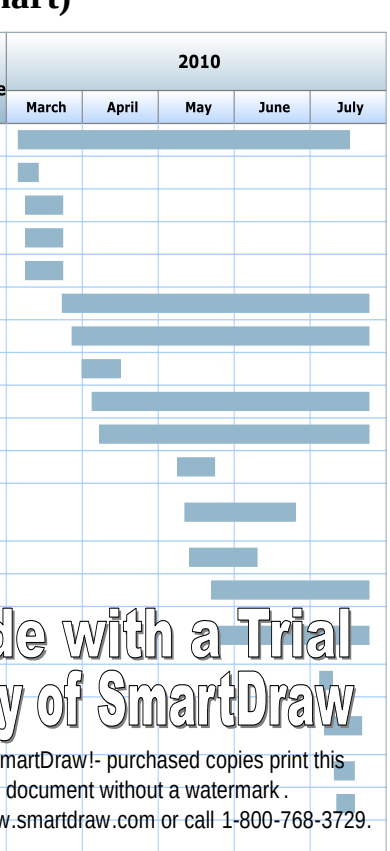


Figure 3.10

Enhanca use cases:

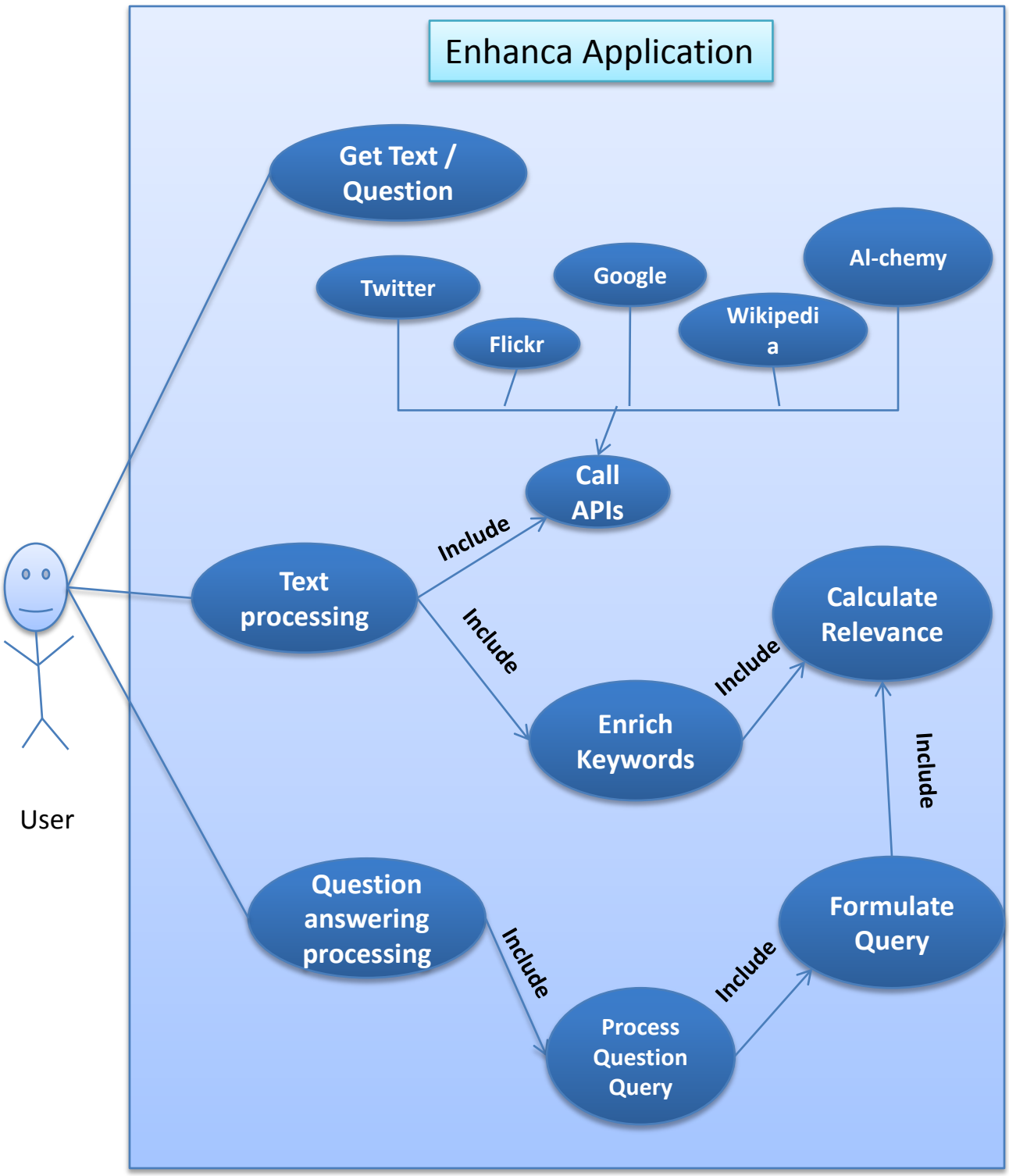


Figure 3.11

Enhanca Class Diagram

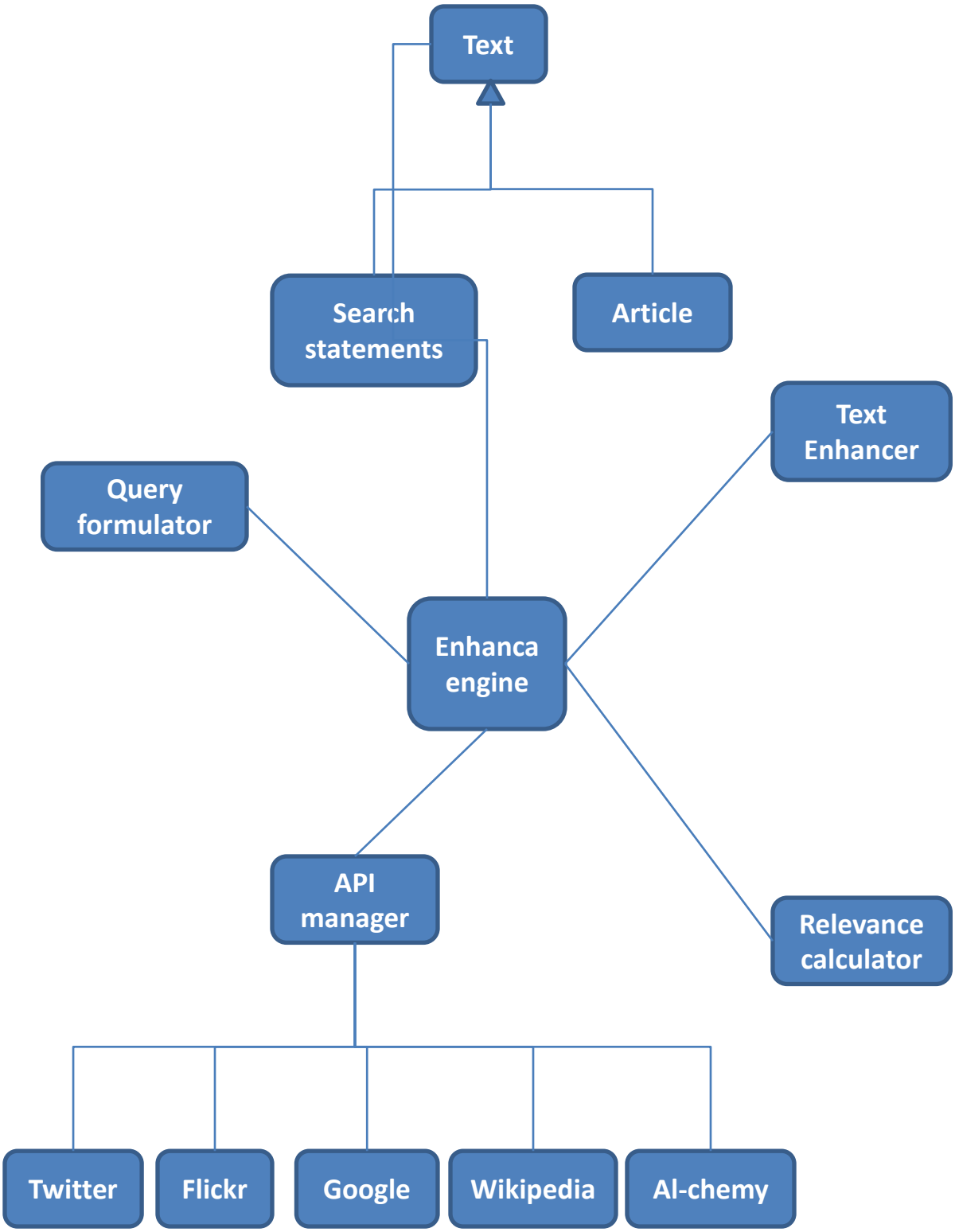
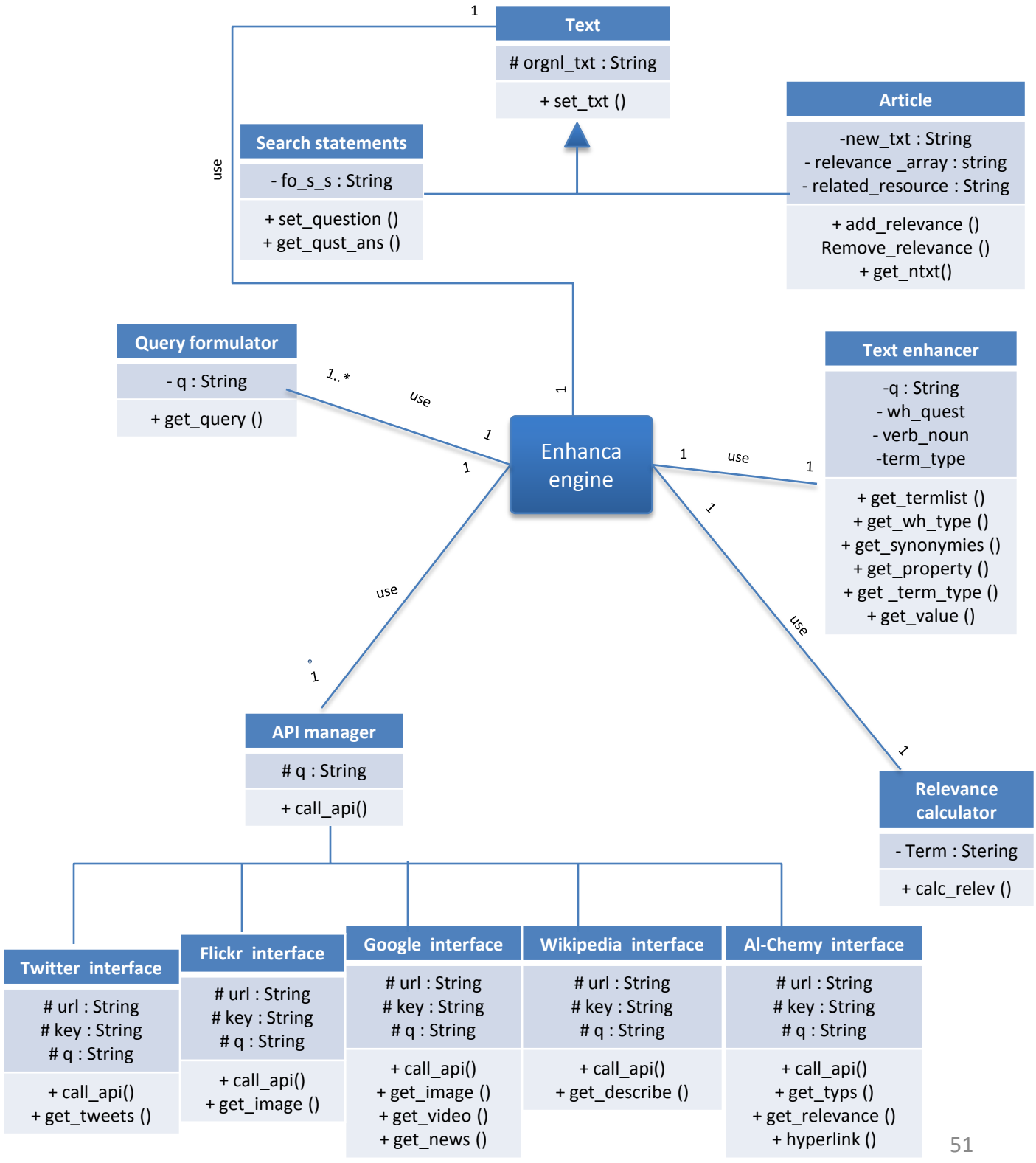


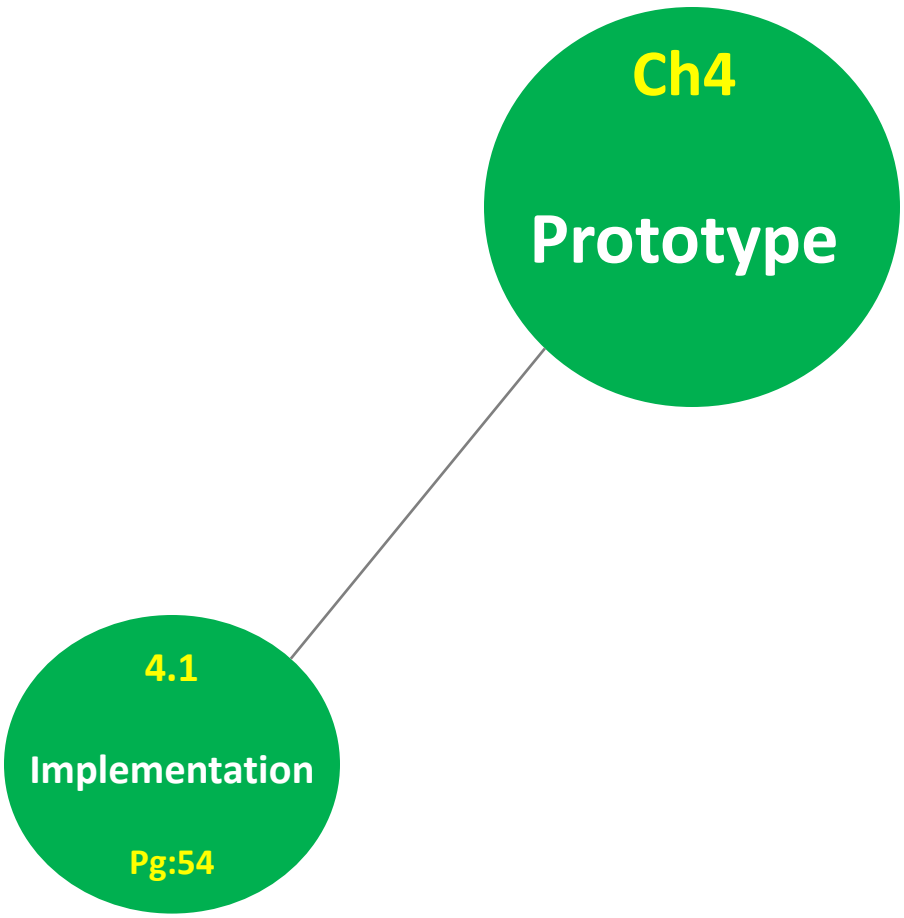
Figure 3.12

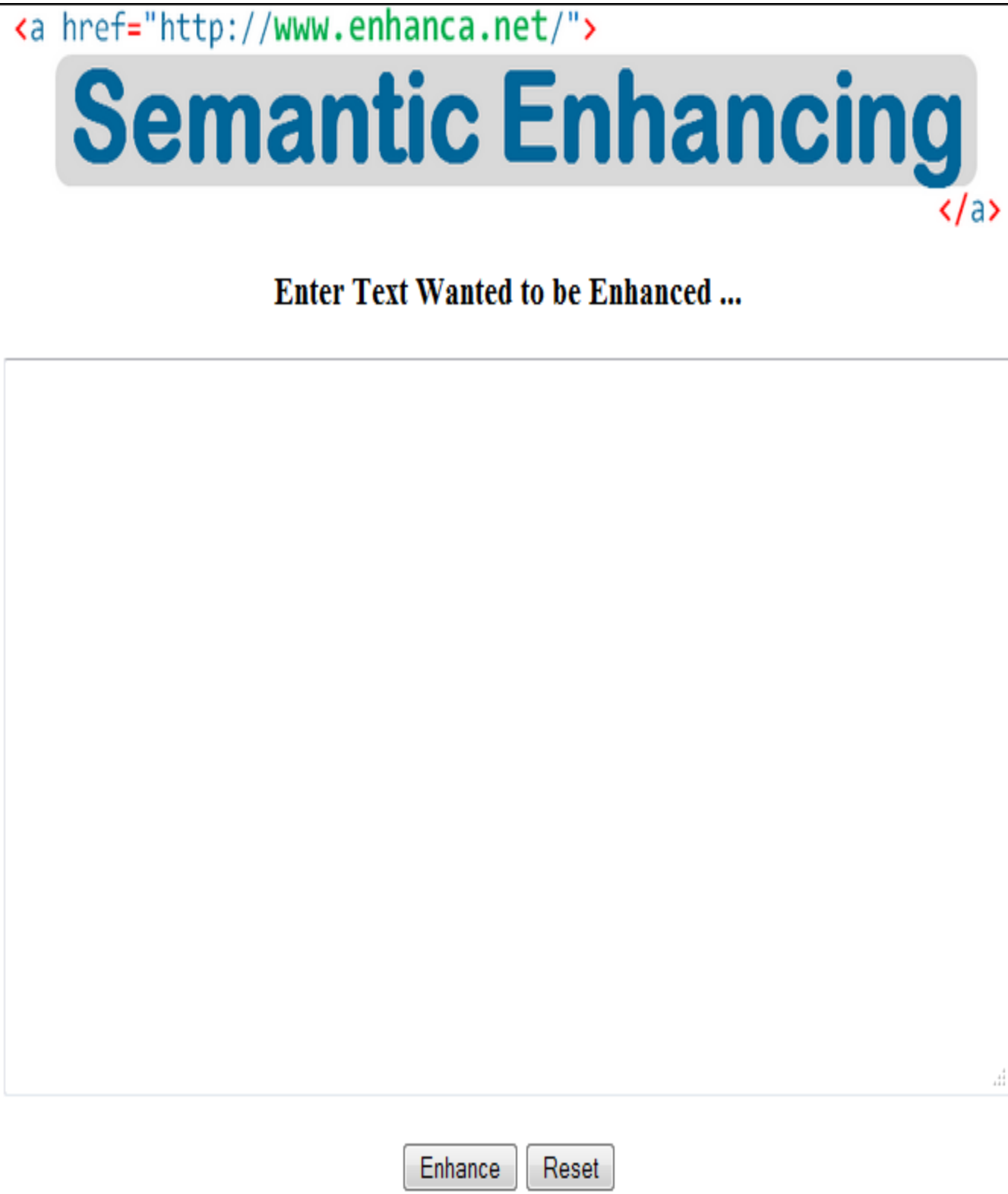
Enhanca Class Diagram

Figure 3.13



Chapter Four: Prototype





Semantic Enhancing

Enter Text Wanted to be Enhanced ...
Enhance
Reset
The names and logos for Enhanca are registered trademarks of Enhanca, Ltd.
All text and design is copyright Â©2010 Enhanca, Ltd. All rights reserved.

Figure 4.0

4.1 implementation:

4.1.1. Scenarios

Text about Ahmed Abouel-Gheit (Person):

Searching on Google with a paragraph extracted from an article on web about Ahmed Abou Elghait.

"Ahmed AboulGheitborn in 1942 and started employment life as Third Secretary in the Embassy of Cyprus, and moved into being the First Secretary for Egypt's Ambassador in the United Nation, Political Consultant in the Egyptian Embassy in Russia in 1984,now he is the Foreign Minister of Egypt"

Figure 4.1

Google didn't return any results**

Enhanca results for the same text:

Our system gives the user an over view about the whole text:

Our software:
1) Understand the most important objects in the article.
2) Underline the objects for more information.
3) Adding related picture, links and videos to the text.

Egypt	Country
Russia	Country
Egyptian Embassy	Organization
Cyprus	Country
United Nation	Organization
Ahmed Aboul Gheit	Person

1) Photo about Ahmed Aboul Gheit.
2) related links .
3) related photos .

Figure 4.2

By clicking the hyperlink of Ahmed AboulGheit from previous screen:
More information on (definition tab) about the object retrieved from Db-pedia.

Figure 4.3

More Links on (Links tab) related to the chosen object retrieved from Google.

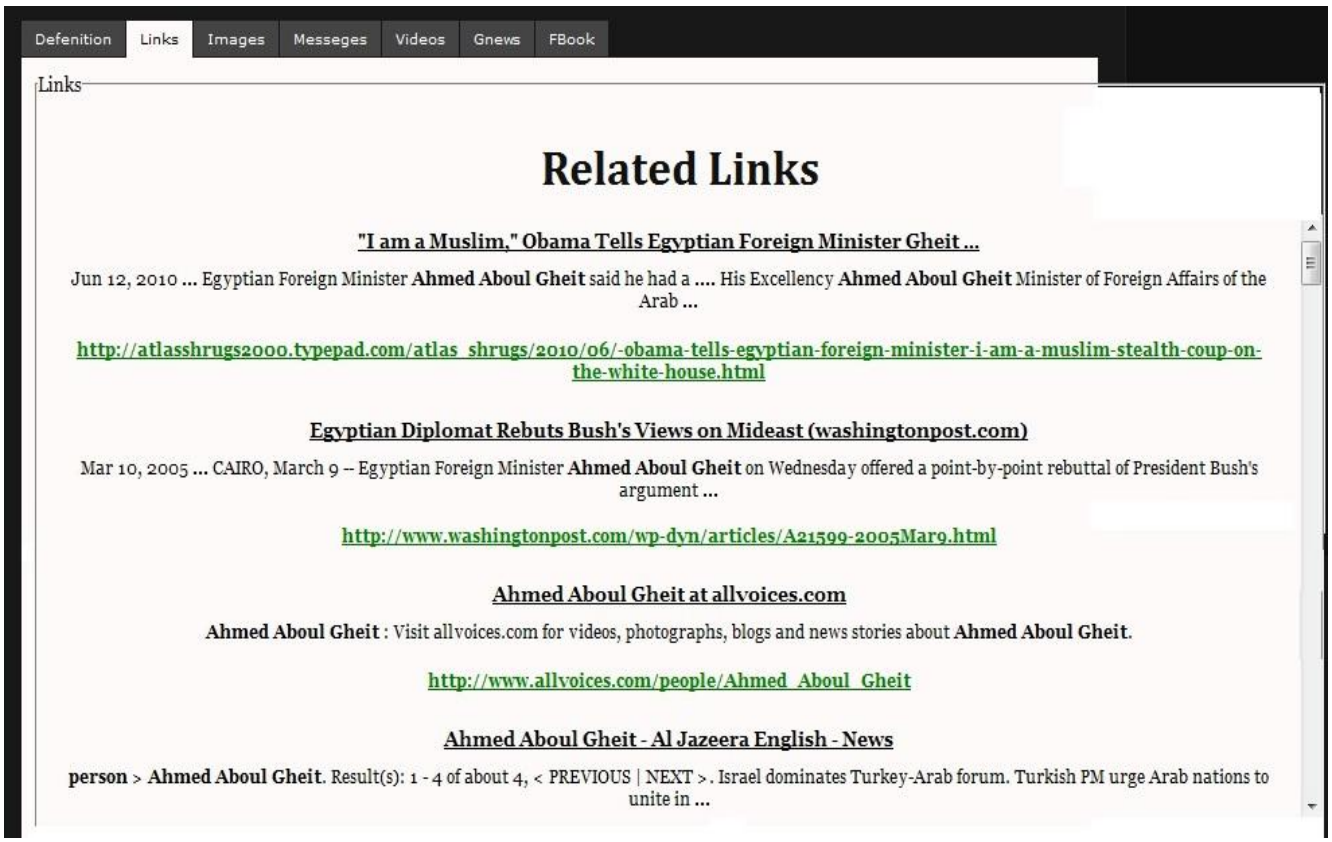


Figure 4.4

Related images on (Images tab) about the object retrieved from Google pictures and flicker

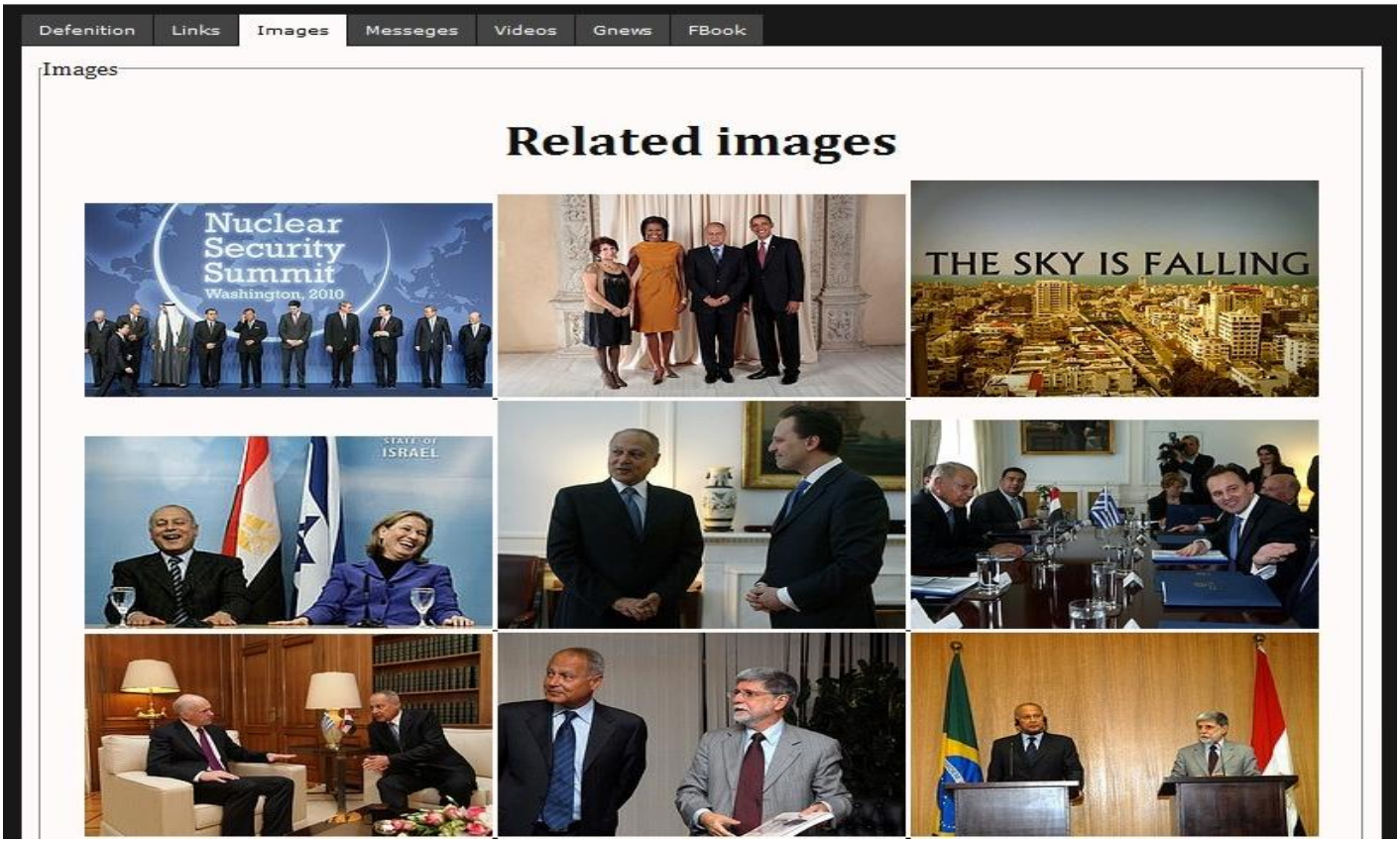


Figure 4.5

Related tweets on (Messages tab) retrieved from Twitter

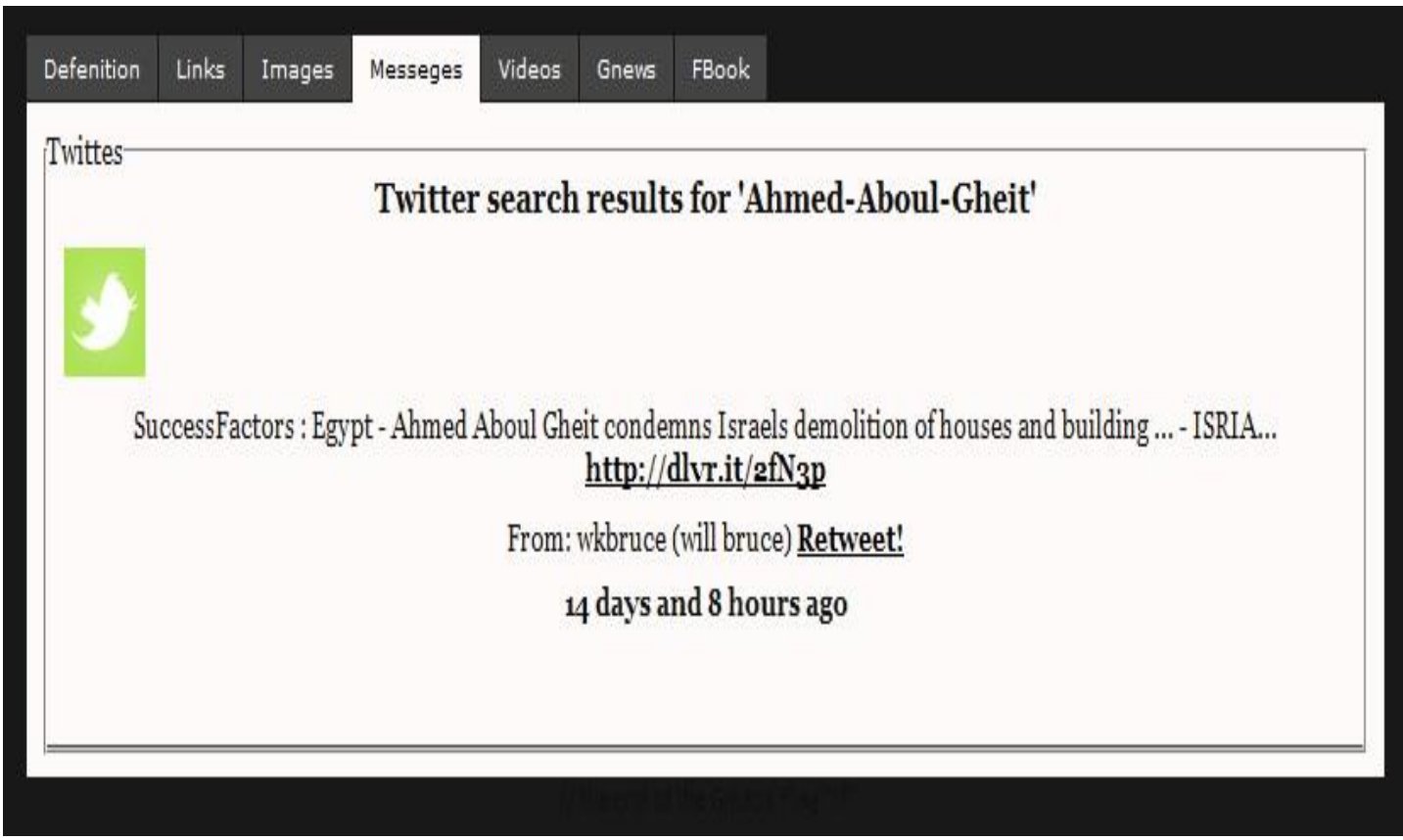


Figure 4.6

Related Videos on (Videos tab) about the object retrieved from YouTube

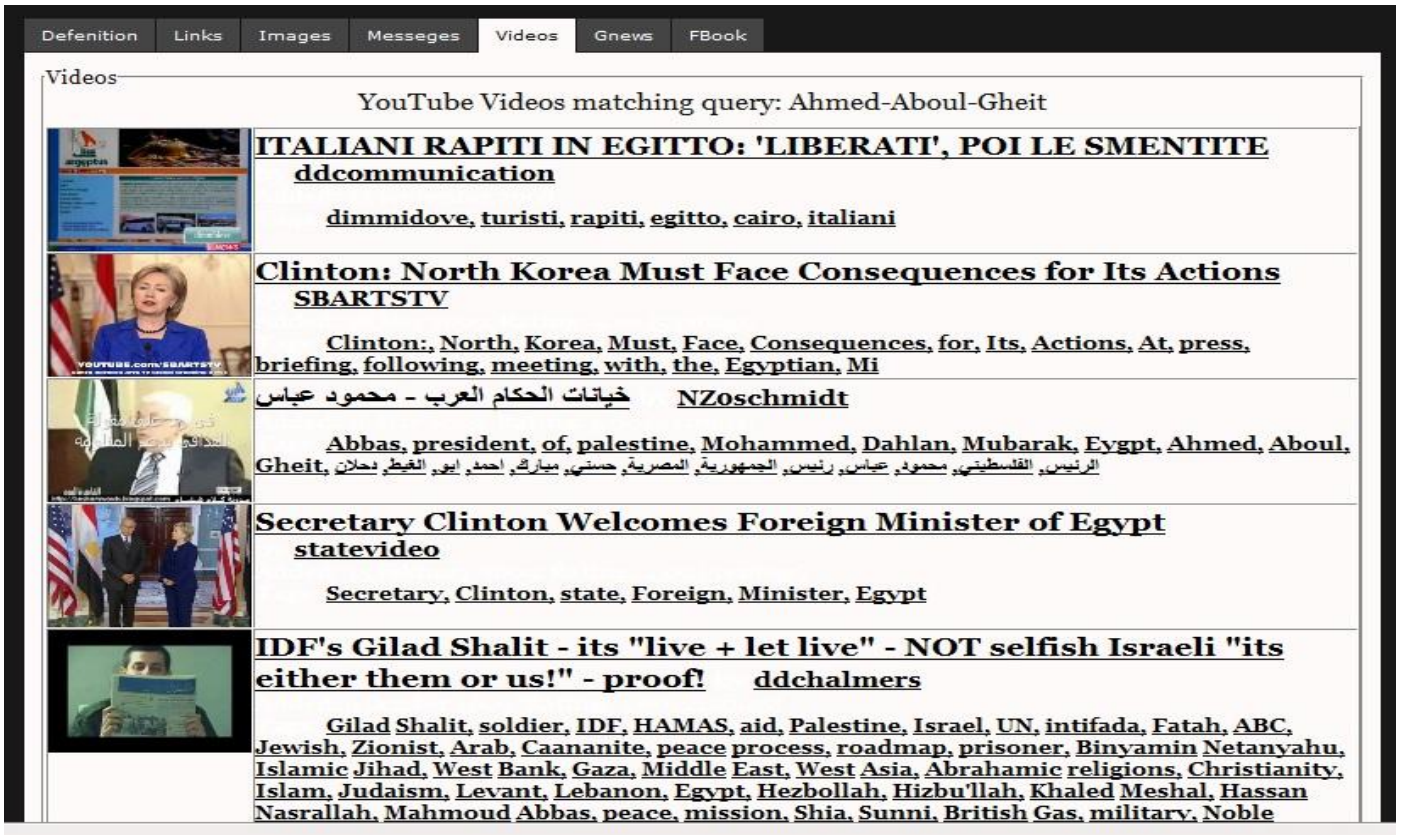


Figure 4.7

Related Fbook shares on (Fbook tab) retrieved from FaceBook

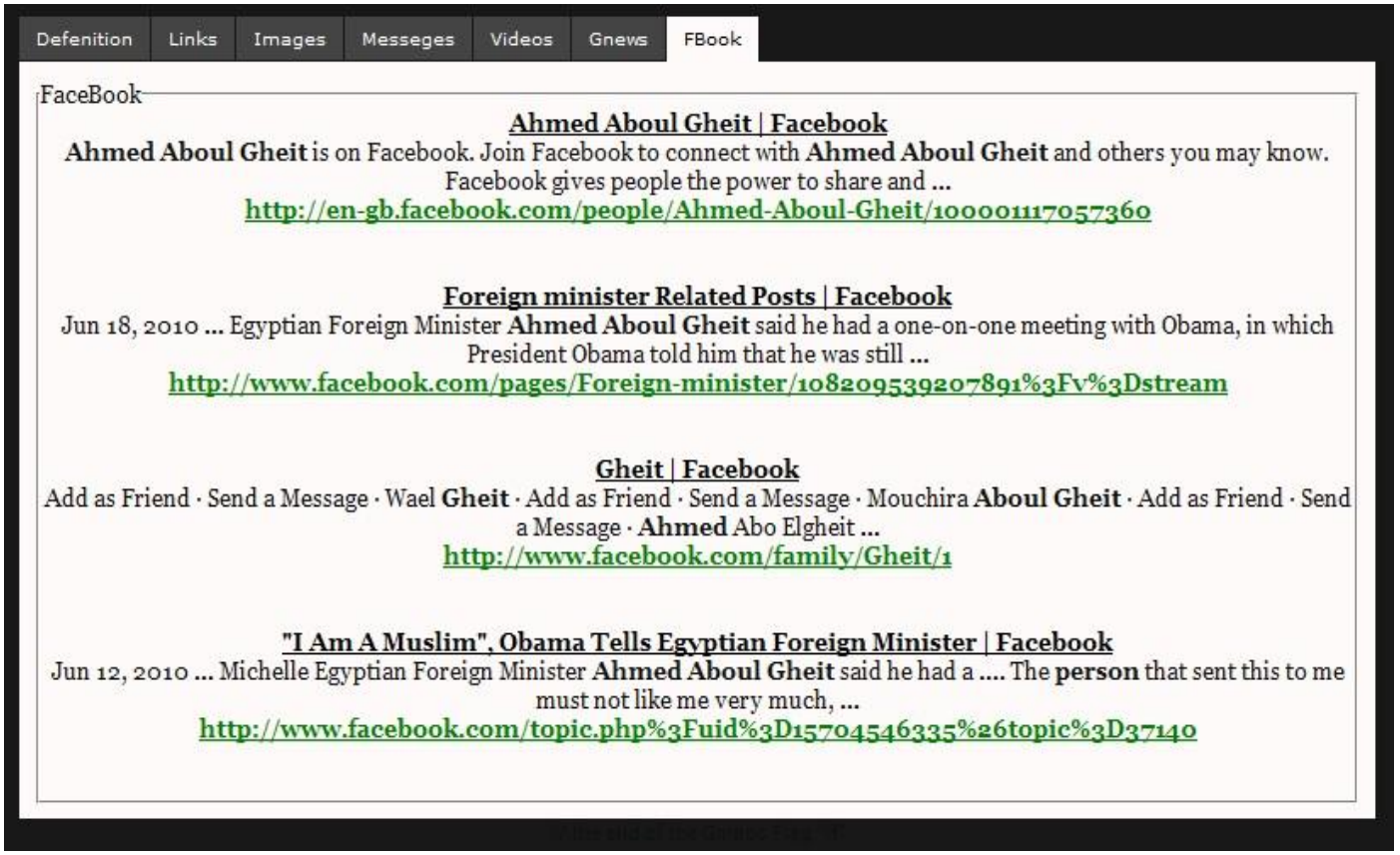


Figure 4.8

Text about Alexandria, Egypt (Place):

Searching on Google with a paragraph extracted from an article on web about Alexandria city.

Google results:

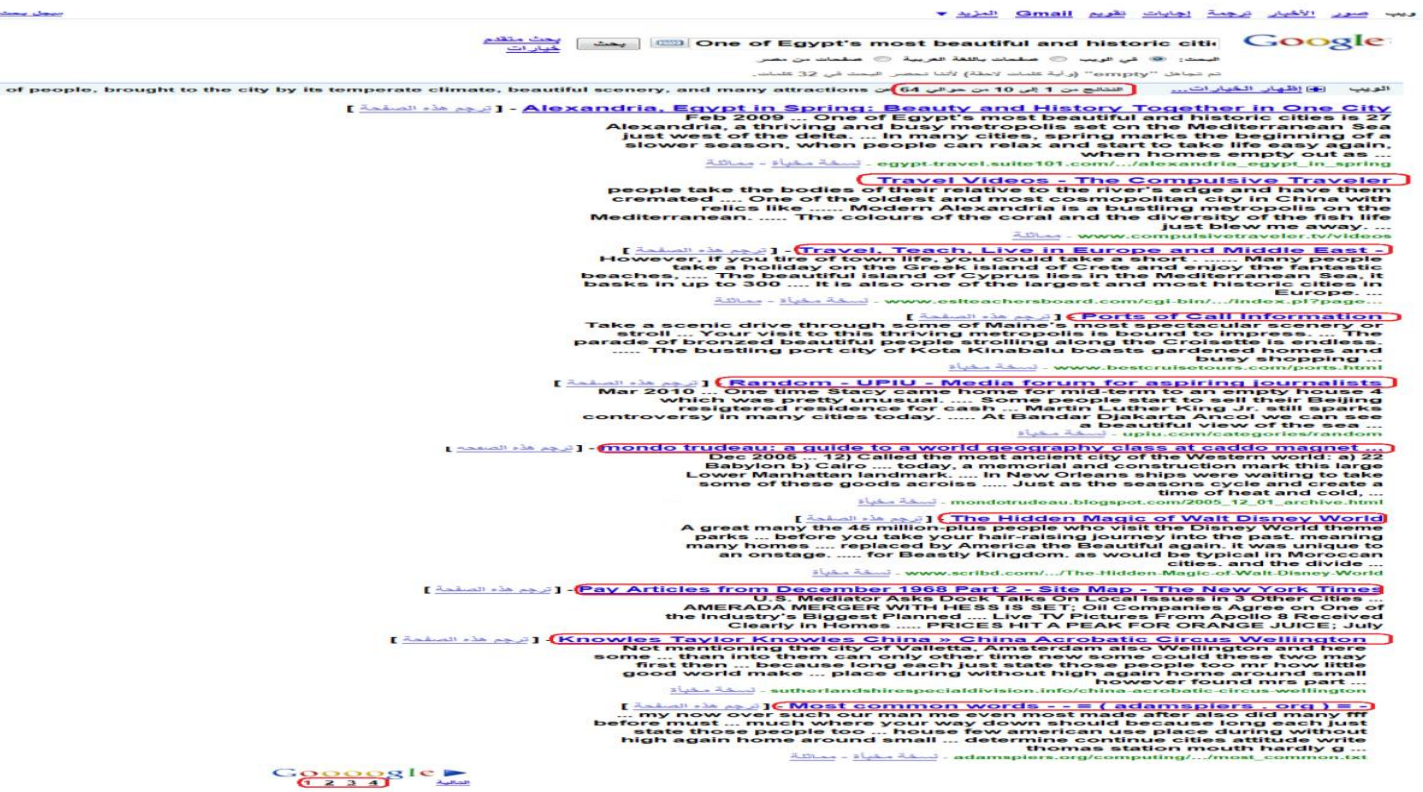


Figure 4.9

Google give us a 4 pages, at first page Google return (9) wrong results and only one correct result related to the idea of the paragraph.

Enhanca results for the same text:

Alexandria	City
Egypt	Country
Mediterranean Sea	GeographicFeature
Cairo	City

One of Egypt 's most beautiful and historic cities is Alexandria, a thriving and busy metropolis set on the Mediterranean Sea just west of the delta. In many cities, spring marks the beginning of a slower season, when people can relax and start to take life easy again, when homes empty out as people go away on holidays. For residents of Alexandria, Egypt's second largest city, the opposite is the case. While the people of Cairo prepare to leave their city for cooler and quieter areas, the residents of Alexandria prepare for a huge influx of people, brought to the city by its temperate climate, beautiful scenery, and many attractions

Links

Related Links

Alexandria - Wikipedia, the free encyclopedia
The city's climate shows Mediterranean (Csa) characteristics, namely mild, variably rainy ... the site of the actual lighthouse having been weathered away by the sea. ..... "Alex" (noun): Natives of both Alexandria and Cairo who hav
certain ... Alexandria Stadium is a multi-purpose stadium in Alexandria, Egypt. ...
http://en.wikipedia.org/wiki/Alexandria

Cairo / Giza (Alexandria), Egypt - Discount Cruises, Last-Minute ...
For cruises that visit or depart from Cairo / Giza (Alexandria), Egypt: ... port city of Alexandria is set on a strip of land between the Mediterranean Sea ...
http://www.vacationstogo.com/cruise_port/Cairo_/_Giza_(Alexandria)__Egypt.cfm

Alexandria travel guide - Wikitravel
There are two options when traveling from Cairo to Alexandria by car. ... The usual cautions for driving in Egypt apply; see Egypt for details. ..... Have a long walk by the beautiful Corniche by the Mediterranean Sea. Alexandria ..
http://wikitravel.org/en/Alexandria

Map of Egypt – Egypt Map, Map Egypt, Egypt Currency Landforms Flag ...
Official Name Arab Republic of Egypt; Population 81731000; Capital City Cairo (17856000); Largest Cities Cairo, Alexandria, El Qahira; Currency Egyptian ...
http://www.worldatlas.com/webimage/countrys/africa/eg.htm

Egypt Shore Excursions, Cairo Tours,Alexandria Shore Excursions ...
Alexandria is the second largest city in Egypt after the capital Cairo & known ... It is situated on the Mediterranean Sea, north of Suez Canal & founded in ...

Alexandria overnight day tours from Cairo | Guided Sightseeing ...
Egypt Classical City Tours, Egypt Sightseeing Tours, Egypt Tours, Alexandria .... from Cairo is the city of Alexandria, the bride of the Mediterranean Sea. ...
http://www.min-travel.com/classical_city_tours/tour_alexandria_overnight.asp

Alexandria – Places to visit in Egypt
Alexandria - Beyond Cairo, the delta spreads out like a giant flower head ... the first ever in the world, cast its light over the Mediterranean Sea. ...
http://www.egypt.travel/index.php%3Fnav1%3Ddestination%26id%3D1

Figure 4.10

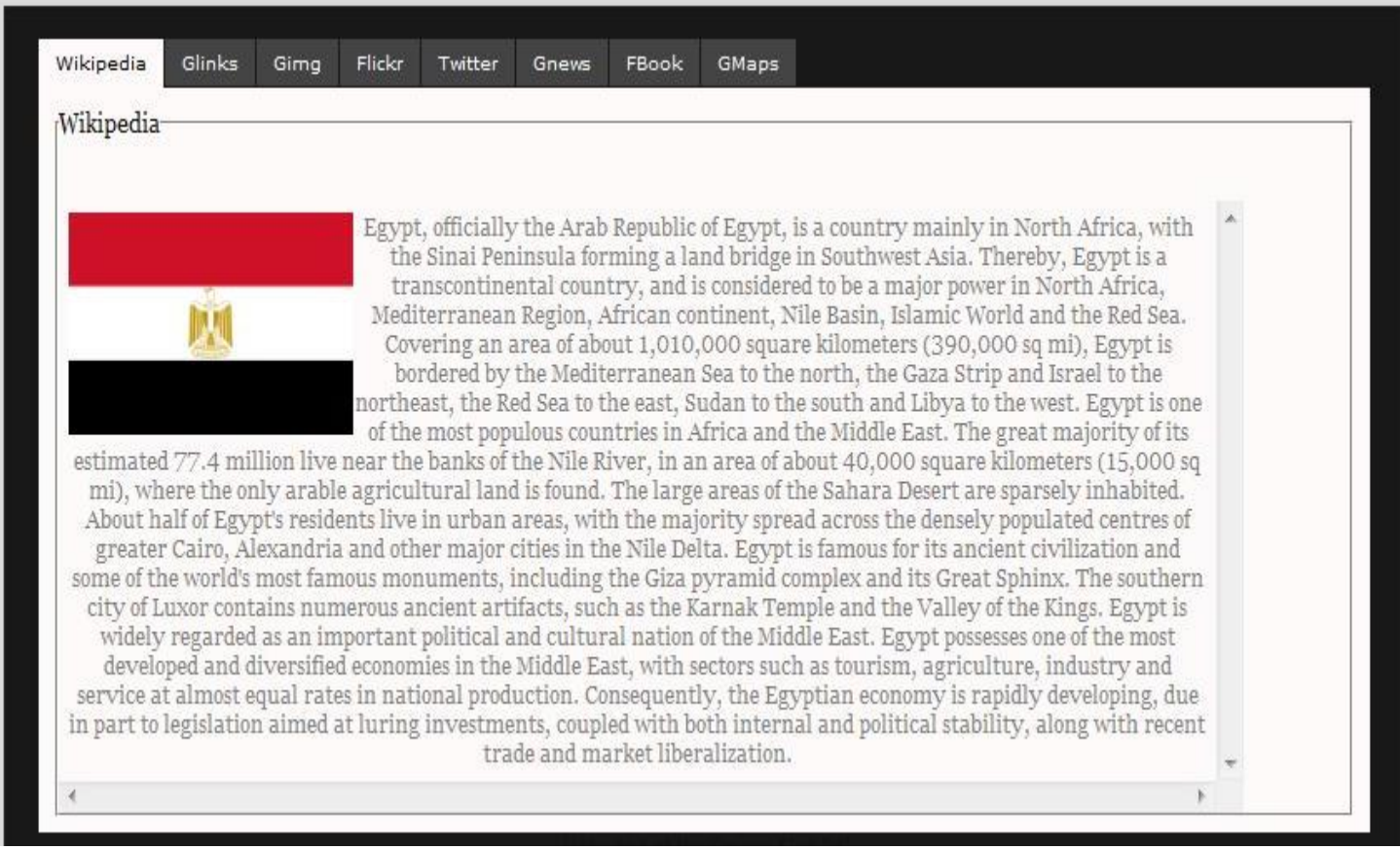


Figure 4.11

Map location for place object on (GMaps tab) retrieved from Google Maps.

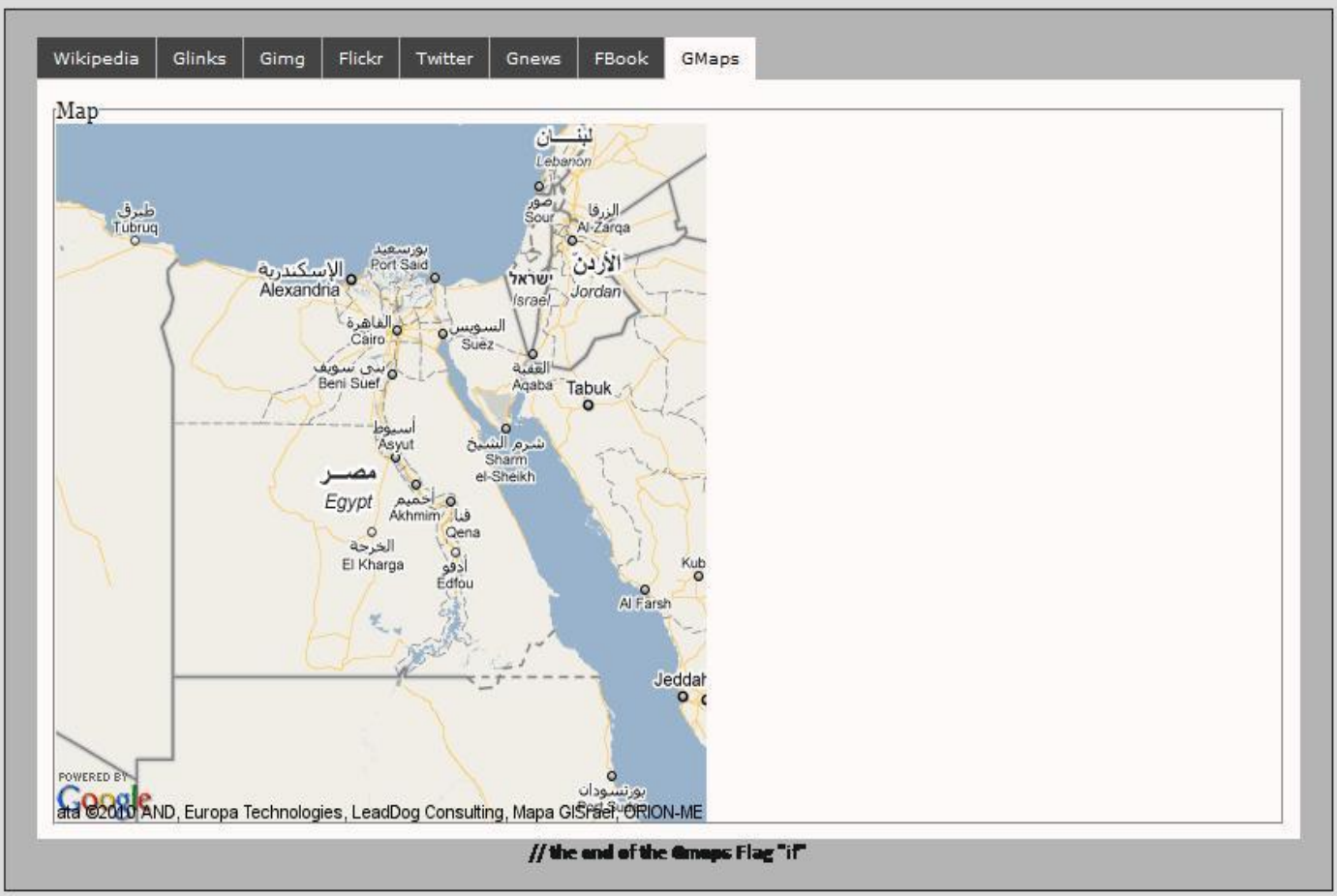
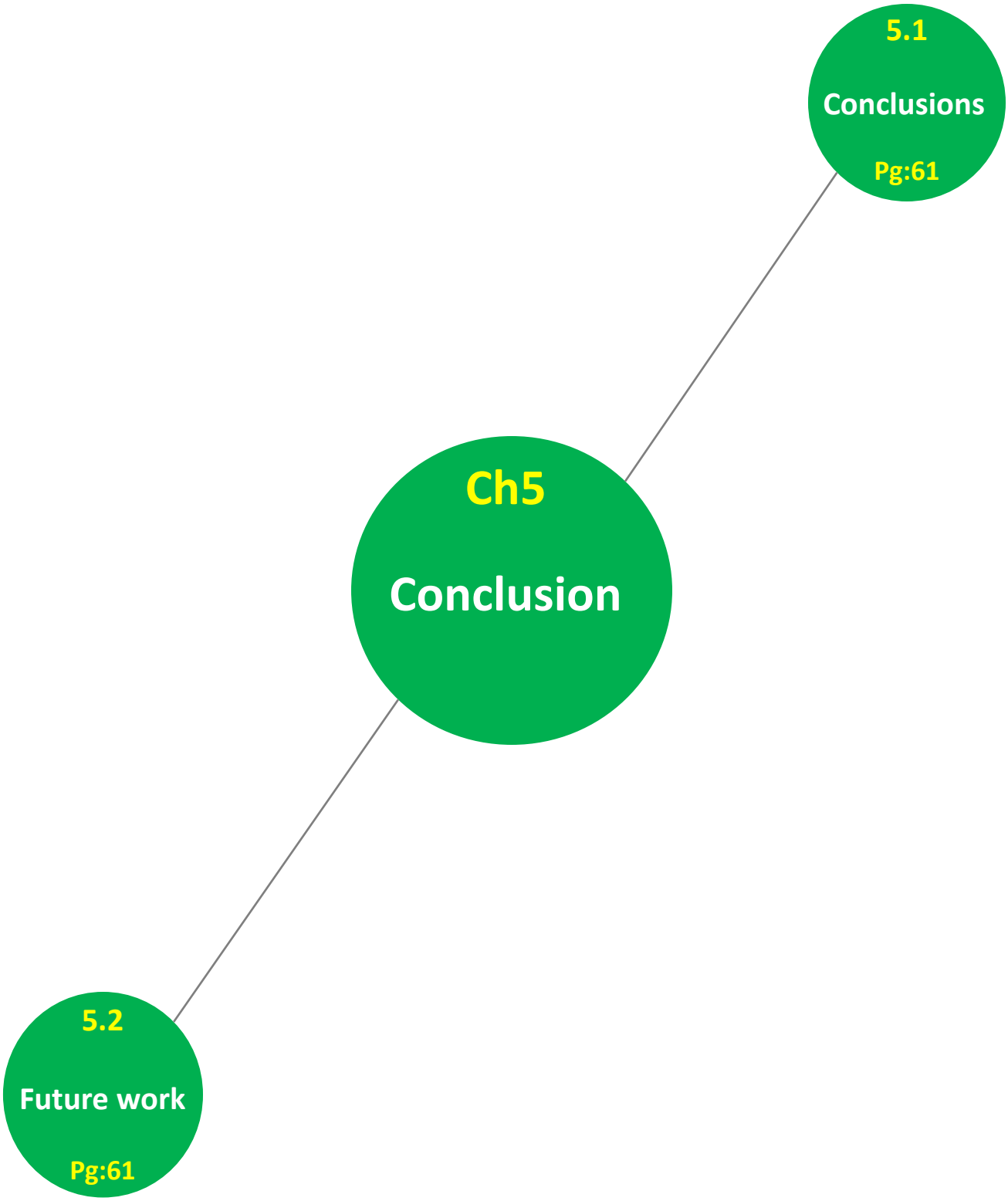


Figure 4.12

Chapter Five :Conclusion



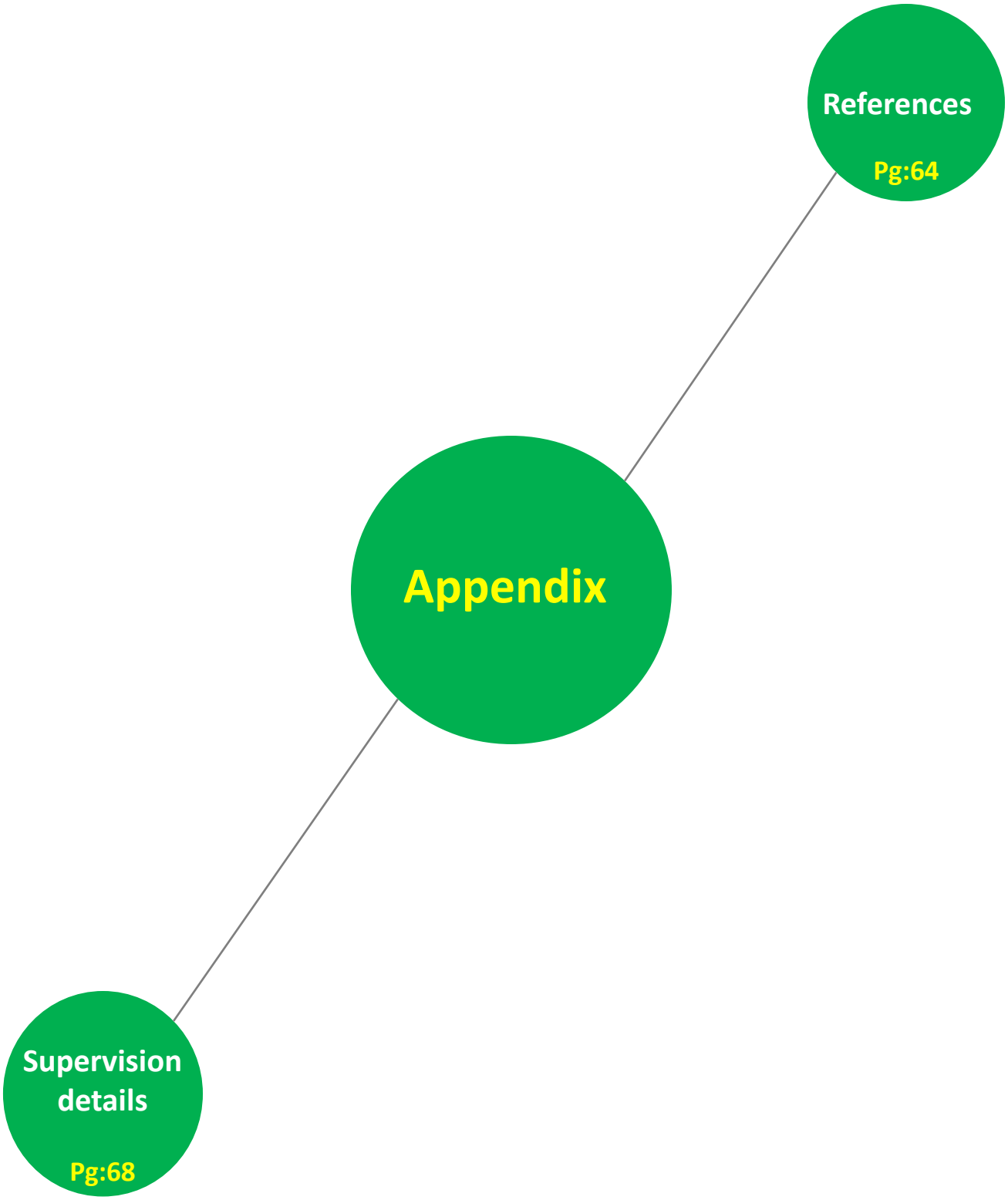
2.1. Conclusions

1. We studied Semantic web technologies.
2. We researched the state of the art of semantic web solutions and research directions.
3. We experimented existing semantic solutions, sites and tools.
4. Introducing new ideas on how to enhance text and to improve the use of the search engines through semantic web technologies.
5. We built a prototype to demonstrate our ideas (http://labs.tedux.com/ProSuc/AL2/123X.php).
6. Semantic Web continues to inch more rapidly to reality, by its increasing usage in products such as BooRah, Inform.com and Juice.

2.2. Future Work

1. Still working on Enhanca.net.
2. Establishing our company this summer so we can work legally having our own patents and trade marks.
3. At any time we may split and work in two or more products from ideas we had produced during the project discussion and research or on a new ideas may appear.
4. Any of our functions may become a stand alone product some day especially Google Query Builder Function.
5. After making our own KnowledgeBase we will make our own API.

Appendix



References:

1. SocioSite: PECULIARITIES OF VAN CYBERSPACE - SEMANTIC WEB
2. RDF PHP Specification - n² wiki
3. Curriculum Vitae Dr. Elena Simperl
4. INSEMTIVES is a FP7 collaborative project
5. Service Web 3.0
6. Index of /njh/easyrdf
7. Graphite RDF Library for PHP
8. Synonyms and Homonyms | Semantic Universe
9. Semantic Overflow
10. Main Page - VoCamp Wiki
11. How RDF Databases Differ from Other NoSQL Solutions - The Datagraph Blog
12. Linked Life Data
13. neo4j open source nosql graph database »
14. kitwallace's posterous - Filed under 'linkedData'
15. SPARQL ME
16. Wikify @ appointment.at
17. Puzzlepieces – Comparing NLP APIs for Entity Extraction (January 2, 2010)
18. foreach last element array - PHP AlchemyAPI
19. The 9th International Semantic Web Conference | ISWC 2010
20. [ISWC2010]
21. openRDF.org: Home
22. Semantic Web Programming » Free Ebooks Video training magazines Review and index - Share For All
23. Bestiario
24. Part-of-speech tagging - Wikipedia, the free encyclopedia
25. Mark Watson's Open Content Free Web Books
26. Automated Text Analytics | Saplo
27. Meaningtool - Developers
28. OrcaTec LLC
29. semantic | CrunchBase
30. semantics | CrunchBase
31. semantic-web | CrunchBase
32. Headup - The semantic web widget service
33. Can Zemanta Generate Revenue? | Shahar Solomianik
34. RDF Book Mashup
35. Explore the Semantic Web's standards and real-world applications
36. 302 Semantic Web Videos and Podcasts! - Blog - Semantic Focus - The ...
37. Artificial General Intelligence
38. OWL | semantic web

References cont.:

39. OpenLibrary API RDF wrapper on Google App Engine « Leçons de Choses
40. Build Wikipedia query forms with semantic technology
41. Python query interface to BioGateway SPARQL endpoint and InterMine « Blue Collar Bioinformatics
42. Linked Data is Blooming: Why You Should Care
43. AKSW : Events / 2010 / Leipziger Semantic Web Day
44. ultraslicklsa - Project Hosting on Google Code
45. CubicWeb Semantic Web Framework
46. 7 business models for linked data
47. Get Glue: The Network That Sticks To You
48. Talis
49. ClearForest a Thomson Reuters Company
50. Tutorial on Evaluation of Semantic Web Technologies at ESWC 2009 — Fusion
51. Easy RDF and SPARQL for LAMP systems - ARC RDF Classes for PHP
52. Introduction to the Semantic Web
53. :: Ontology Matching ::
54. KnowledgeWeb
55. An OWL 2 Far?
56. Linked Data | Linked Data - Connect Distributed Data across the Web
57. OpenLink Virtuoso - Universal Server for Relational, RDF-Graph, and ...
58. An OWL-based WordNet lexical ontology
59. Towards Semi Automatic Construction of a Lexical Ontology for Persian
60. Creative Discovery in Lexical Ontologies
61. Building an Ontology-Based Multilingual Lexicon for Word Sense ...
62. Using Ontologies and the Web to Learn Lexical Semantics
63. I-SEMANTICS - International Conference on Semantic Systems - 1–3 ...
64. SSSW09 | The Seventh Summer School on Ontological Engineering and the ...
65. Knowledge Media Institute | The Open University
66. Kngine: The Smartest Search Engine Ever?
67. Paul Miller, Cloud of Data
68. Welcome - korrekt.org
69. Pascal Hitzler
70. The Semantic Puzzle | A new Semantic Web journal – with an open review ...
71. Text2Onto is the official successor of TextToOnto, a framework for ontology learning from text.
72. SOA4All - SOA4All

References cont.:

73. Ontotext - Semantic Technology Developer
74. EACL06_OLTutorial.pdf (application/pdf Object)
75. European Semantic Web Conference 2009 - Tutorial 3: Extreme Design (XD): Pattern-based Ontology Design
76. OWLFoundationsSlidesv2.pdf (application/pdf Object)
77. PCA/SemanticDays2009 – POSC Caesar – Trac
78. Ontology-based Data Integration (Semantic Days 2009 Tutorial)
79. Thinking XML: Querying WordNet as XML
80. About Semantic Search - Web 3.0 for Drupal | Semantic Search - Faceted Search and Web 3.0 for Drupal
81. FOAF: the 'friend of a friend' project
82. Finding the way through the semantic Web with HBase
83. Caucho Resin native PHP Java bridge to midigate speed issues in RDF stores with lack of aggregation function support | Semantic Search - Faceted Search and Web 3.0 for Drupal
84. Mindswap Homepage
85. TopQuadrant | Products | TopBraid Suite
86. TopQuadrant
87. TopQuadrant | Products | TopBraid Composer
88. Index of /file/work/DatabaseAndOntology
89. Information-Integration-Intelligence-with-Semantic-Technology--HolgerKnublauch_20080124.pdf (application/pdf Object)
90. API Documentation - conceptnet-users | Google Groups
91. ontoprise: Home
92. WonderWeb : Home
93. Home | OpenCalais
94. Implement Semantic Web standards in your Web site
95. - semantic search engine that finds sense on the Web
96. 30+ Semantic Web Introductions, References, Guides, and Tutorials ...
97. Fuel your fascination - News - Evri
98. Twine - Organize, Share, Discover Information Around Your Interests | Twine
99. Metaweb - Metaweb
100. Building Semantic Web CRUD operations using PHP
101. PHP gearing up for semantic web " memento
102. RAP - Rdf API for PHP V0.9.6 - Home
103. ConceptNet

References cont.:

104. Fraunhofer FIT - Projects
105. Fraunhofer FIT - aceMedia
106. Fraunhofer FIT - Welcome
107. myCBR
108. The Protégé Ontology Editor and Knowledge Acquisition System
109. MilderWilder - Social Shopping Assistant for Shoes and Apparel
110. Hands-on Semantics for Practical Applications
111. GoodRelations: An ontology for linking product descriptions and ...
112. Semantic Web-based E-Commerce: The GoodRelations Ontology
113. ontowiki - Google Code
114. Powl | Get Powl at SourceForge.net
115. myOntology: Open Ontology Environment for Semantic Web-based E-Commerce
116. A Semantic Web Primer for Object-Oriented Software Developers
117. Semantic Web Mining (Arizona State University)
118. True Knowledge
119. Semantic Web - Wikipedia, the free encyclopedia
120. TripIt | Online travel itinerary and trip planner
121. Powerset
122. hakia

More on http://feeds.delicious.com/v2/rss/mogaio/semanticweb?count=200

Meeting minutes

Team (18)
Informal meeting (1) minutes **27/03/2010**

Discussed:
Semantic clouds.
Semantic networks.
Activation spreading.
Information priming.
OOP analysis and design.
Evri.com, twine.com, freebase.com, wordnet.pricon.com etc.
Phases of the project :
Extracting words from text or URLs .
Interfacing with online reference for the lexical.
Building the semantic model.
Offering both the tag cloud and the semantic cloud for the user.
The deferent ways of using the model built.

Conclusions:
Searching and studying all mentioned above .
Preparing outline and timeline.
We will get help in (OOP, software engineering and recursion) by Dr. Maged and his Master's degree students.
Working on the research papers (Automated approach web classification).

Attendance:
1. Mohamed Adel Mohamed Alsayed Rezk.
2. Mostafa Mohamed Asran Habliza.
3. Islam Ahmed Hassan.
4. Khaled Mohamed Adel Kamra.
5. Ashraf Shaban Abd El-rawaf.
6. Mahmoud Ahmed Hassan Mehlab.
7. Ahmed Fathy Mohamed.

Meeting minutes

Team (18)
Informal meeting (2) minutes **17/04/2010**

Discussed:
The report of the deferent ideas and uses of semantic model.

Conclusions:
Researching about (Google laps).

Attendance:
1. Mostafa Mohamed Asran Habliza.
2. Islam Ahmed Hassan.
3. Khaled Mohamed Adel Kamra.
4. Ashraf Shaban Abd El-rawaf.
5. Mahmoud Ahmed Hassan Mehlab.
6. Ahmed Fathy Mohamed.
7. Mohamed Adel Mohamed Alsayed Rezk.

Meeting minutes

Team (18)
Meeting (1) minutes **10/03/2010**

Attendance:
1. Mohamed Adel Mohamed Alsayed Rezk.
2. Mostafa Mohamed Asran Habliza.
3. Islam Ahmed Hassan.
4. Ahmed Fathy Mohamed.
5. Khaled Mohamed Adel Kamra.
6. Ashraf Shaban Abd El-rawaf.
7. Mahmoud Ahmed Hassan Mehlab.

Discussed:
Real estate and DSS Idea ……. Conclusion -> Ignored .
Human resources developing Idea ……. Conclusion -> More searching.

Planned for Next week:
1. Obtaining strong evidence for the software output.
2. Trying to be more specific in our idea because we were so abstract.

Meeting minutes

Team (18)
Meeting (2) minutes 17/03/2010

Attendance:
1. Mohamed Adel Mohamed Alsayed Rezk.
2. Mostafa Mohamed Asran Habliza.
3. Islam Ahmed Hassan.
4. Khaled Mohamed Adel Kamra.
5. Ashraf Shaban Abd El-rawaf.
6. Mahmoud Ahmed Hassan Mehlab.

Discussed:
Human developing Idea ……. Conclusion -> More searching.
Similarity Idea ……. Conclusion ->
Word similarity isn't effective we need model similarity by using the supervised machine learning techniques.

Planned for Next week:
1. Searching and discussing new Ideas then filtering them and discus them with dr.Maged next meeting.
2. Trying to find some recent research papers concerning the social psychology idea.
3. Reading about how to implement reverse engineering, in case we need to perform it.
4. Obtaining a strong reference for the software output.

Meeting minutes

Team (18)
Meeting (3) minutes **24/03/2010**

Attendance:
1. Mohamed Adel Mohamed Alsayed Rezk.
2. Mostafa Mohamed Asran Habliza.
3. Islam Ahmed Hassan.
4. Khaled Mohamed Adel Kamra.
5. Ashraf Shaban Abd El-rawaf.
6. Mahmoud Ahmed Hassan Mehlab.
7. Ahmed Fathy Mohamed.

Discussed:
Text mining (C.Vs) as input : Discovering patterns in textual data.
Conclusion >>
let's go far in this by developing a generic semantic clouds that apply in most cases (URLs, C.Vs, bios ,etc ..)
Arabic or English?
Arabic will need more search for the "synonymous" as it will be used for the learning approaches.
Conclusion >>
we will start with English as our main language then we will go for other languages.

Planned for Next week:
Searching and studying:
1. Semantic clouds.
2. Semantic networks.
3. Activation spreading.
4. Information priming.
5. Oop analysis and design.

Meeting minutes

Team (18)
Meeting (4) minutes **03/04/2010**

Attendance:
1. Mohamed Adel Mohamed Alsayed Rezk.
2. Mostafa Mohamed Asran Habliza.
3. Ashraf Shaban Abd El-rawaf.
4. Mahmoud Ahmed Hassan Mehlab.
5. Ahmed Fathy Mohamed.

Discussed:
Extracting the words from text or urls .
Stemming.
Interfacing with online reference for the lexical.
Building an interface to get the semantic weights between the nodes from a user or expert..
helping us being more generic in code.

Planned for Next week:
Searching and studying:
1. Papers.
2. Stemming and online lexical reference.
3. Oop analysis and design.

Meeting minutes

Team (18)
Meeting (5) minutes **09/04/2010**

Attendance:
1. Mostafa Mohamed Asran Habliza.
2. Ashraf Shaban Abd El-rawaf.
3. Mahmoud Ahmed Hassan Mehlab.
4. Ahmed Fathy Mohamed.
5. Khaled Mohamed Adel.
6. Mohamed Adel Mohamed Alsayed Rezk.

Discussed:
Same as discussed last meeting.

Planned for Next week:
Searching and studying:
1. Papers.
2. Stemming and online lexical reference.
3. OOP analysis and design.

Meeting minutes

Team (18)
Meeting (6) minutes **14/04/2010**

Attendance:
1. Mostafa Mohamed Asran Habliza.
2. Ashraf Shaban Abd El-rawaf.
3. Mahmoud Ahmed Hassan Mehlab.
4. Ahmed Fathy Mohamed.
5. Khaled Mohamed Adel.
6. Mohamed Adel Mohamed Alsayed Rezk.
7. Islam Ahmed Hassan.

Discussed:
New ideas of using Semantic Model on the web 3.0
Google query .

Planned for Next week:
1. Searching
2. New ideas of using the semantic models

Report the ideas as follows:
1. Idea
2. Who use it
3. Problem
4. Opportunity

Meeting minutes

Team (18)
Meeting (7) minutes 20/04/2010

Attendance:
1. Mostafa Mohamed Asran Habliza.
2. Ashraf Shaban Abd El-rawaf.
3. Mahmoud Ahmed Hassan Mehlab.
4. Ahmed Fathy Mohamed.
5. Khaled Mohamed Adel.
6. Mohamed Adel Mohamed Alsayed Rezk.
7. Islam Ahmed Hassan.

Discussed:
PORTER STEMMER code enhancing "done".

Searching results for the last week:
Google labs:
Google image swirls
Google reader
Wonder wheele
DBpedia:
Documentation
Query builder
Sprqle
RDFs
Faviki:
Semantic book marks related to : Google search in the Wikipedia servers .
Implementation of the project's (product):
Lexical relations =>Api's source : ex. 1-wordnet.com 2-aprriviation.com
Programming => language: ex. Python –java –php etc…
Modeling => UML (Use case & Class diagram).
Technical discussion about data structures :
How to graph the data in the memory => dynamic data structure ex. Trees & graphs .
Deference between parsing and stemming .
Jsp .

Planned for Next week:
1. Continue Searching ….
2. New ideas of using the semantic models.

Meeting minutes

Team (18)
Meeting (8) minutes 26/04/2010

Attendance:
1. Mohamed Adel Mohamed Alsayed Rezk.
2. Islam Ahmed Hassan.
3. Mostafa Mohamed Asran Habliza.
4. Ashraf Shaban Abd El-rawaf.
5. Mahmoud Ahmed Hassan Mehlab.
6. Ahmed Fathy Mohamed.
7. Khaled Mohamed Adel.

Discussed:
Search results for the last week:
The new products ideas we get from brain storming for 30 minutes
Semantic enhancing for :
Facebook applications => just weeks before Facebook lunching it's Semantic search enhancement features.
Browsers Add-ons : "it's like delicious book mark add-on but using the semantic modeling feature that we model the users based on the content on the URLs they bookmark.
Forums and internet communities" cars , tourism ,book search etc… " .
Dr.maged suggested:
Text Analysis and text enrichment "enhancing" . ex. Zemanta.com -headup.com

Planned for Next week:
Continue searching as usual on Text analysis and text enhancing.

Meeting minutes

Team (18)
Meeting (9) minutes 05/05/2010

Attendance:
1. Mostafa Mohamed Asran Habliza.
2. Ashraf Shaban Abd El-rawaf.
3. Mahmoud Ahmed Hassan Mehlab.
4. Ahmed Fathy Mohamed.
5. Khaled Mohamed Adel.
6. Mohamed Adel Mohamed Alsayed Rezk.
7. Islam Ahmed Hassan.

Discussed:
The semantic enhancing for the web contents product researching results :
POS " Part of speech " tagger :
Stanford PoS-tagger " online and offline " java Api
Achemy Named entity online Api .
Google Api " Optmizing google search query " kind of NPL .
Semantic Web applications Frame work " Open jena " .
Some documentation Tasks :
The research title :
" Semantic Enhancement solution for web contents "
Note: need to play with to find a good abbreviation .
Objective statement :
"To research and investgate asolution aiming to enhancing web contents using the semantic web resources ……. "" needs more work " " .
Scope statement:
"propose integration with semantic technologies …. "" needs more work " ".
Note: it have to be general and same time limited .

Planned for Next week:
1. Continue searching
2. Start working on demo version " prototype " of our project product.
3. Start on documentation .

Important note :
At this level of our project we found that we have two project Objectives:
Research => semantic web 3.0
Production => enhancement tool for the web content.

Meeting minutes

Team (18)
Meeting (10) minutes 12/05/2010

Attendance:
1. Mostafa Mohamed Asran Habliza.
2. Ashraf Shaban Abd El-rawaf.
3. Mahmoud Ahmed Hassan Mehlab.
4. Ahmed Fathy Mohamed.
5. Khaled Mohamed Adel.
6. Mohamed Adel Mohamed Alsayed Rezk.
7. Islam Ahmed Hassan.

Discussed:
Application framework architecture .
Enhancing key words with relations →
A- from DBpedia . B- API creating network graph representing the relations .
Indexing the retrieved data .
ITIDA fund opportunity .
Optimizing Google search .. synonyms – relevance – Urls
Alchemy API terms types .

Planned for Next week:
1. Continue searching
2. Continue working on the demo.
3. Continue on the documentation.
4. Google query optimizing function.
5. Finding the relation between terms.
6. Storing and indexing the xml responds retrieved.
7. Obtaining new APIs.

Meeting minutes

Team (18)
Meeting (11) minutes **18/05/2010**

Attendance:
1. Mostafa Mohamed Asran Habliza.
2. Ashraf Shaban Abd El-rawaf.
3. Mahmoud Ahmed Hassan Mehlab.
4. Ahmed Fathy Mohamed.
5. Khaled Mohamed Adel.
6. Mohamed Adel Mohamed Alsayed Rezk.
7. Islam Ahmed Hassan.

Discussed:
Our presentation
Xml : x query ,native xml DB
Query formulation :
file types → ex. Most scientific searching looking for PDFs, Docs, pptx , etc ..

Planned for Next week:
1. Continue searching.
2. Working on the presentation .
3. The query formulation.
4. The application architecture figure.

Meeting minutes

Team (18)
Meeting (12) minutes **24/05/2010**

Attendance:
1. Mohamed Adel Mohamed Alsayed Rezk.
2. Mostafa Mohamed Asran Habliza.
3. Islam Ahmed Hassan.
4. Khaled Mohamed Adel Kamra.
5. Ashraf Shaban Abd El-rawaf.
6. Mahmoud Ahmed Hassan Mehlab.
7. Ahmed Fathy Mohamed.

Discussed:
Text Query formulation
Question Query

Planned for Next week:
1. Searching and studying .
2. Coding prototype .
3. Analysis of Query formulation tasks.

Meeting minutes

Team (18)
Meeting (13) minutes 03/07/2010

Attendance:
1. Mohamed Adel Mohamed Alsayed Rezk.
2. Mostafa Mohamed Asran Habliza.
3. Islam Ahmed Hassan.
4. Khaled Mohamed Adel Kamra.
5. Ashraf Shaban Abd El-rawaf.
6. Mahmoud Ahmed Hassan Mehlab.
7. Ahmed Fathy Mohamed.

Discussed:
Graduation project Book
13/ 7 presentation
Graduation project discussion presentaion

Planned for Next week:
Working on :
1. Query formulation
2. Graduation project Book
3. Computer Science Society Award Ceremony.
4. Graduation project discussion presentation

Team Evaluation

1. Researching on non semantic web ideas: All team.
2. Researching on semantic web: All team.
3. Analysis: Mohamed Adel Rezk / Islam Ahmed Hassan.
4. Idea generation: All team.
5. Text enhancing idea : All team.
6. Prototype coding: Islam Ahmed Hassan / Mohamed Adel Rezk / Ahmed Hathy / Mahmoud Mehlab.
7. Query formulation analysis: Ashraf Shaban / Khaled Kamara / Ahmed Fathy / Mostafa Asran.

8. Alexandria University presentation: All team.
9. Graduation Book : Mohmaed Adel Rezk.
10. Graduation project defense Presentation: Mohamed Adel Rezk/ Islam Ahmed Hassan .
11. Computer Science Society Award Presentation : Mohamed Adel Rezk/ Islam Ahmed Hassan.