Conceptual Design of an Ontological Approach to personalising GUISET Portal

Nonhlanhla Melody Gumbi

(**20050461**)

A dissertation submitted in fulfillment of the requirements for the degree of

Master of Science in Computer Science

Department of Computer Science, Faculty of Science and Agriculture

University of Zululand
Supervisor: Professor M.O. Adigun
Co-Supervisor: Mr. E. Jembere

2014

Declaration

I Nonhlanhla Melody Gumbi, declare that this research study represents my own work and has never been presented in any form for the award of any degree in any University. All the material used as source of knowledge has been acknowledged in text and references.

______________________________   ______________________________

Signature   Date

Dedication

To my mother Simangele Irene Gumbi

Acknowledgements

I would like to thank the Almighty God, without whose enduring mercy I could not have achieved anything.

I would like to thank my supervisors Mr E. Jembere and Prof M.O. Adigun for their support and guidance throughout this research work making it a reality. I would also like to thank my fellow researchers in the centre for mobile e-Service for development, M.T. Nene and T.Shezi for their support during the implementation stage of this work. N.Y.Z. Ndlela, P.T. Cwele and S. Cebekhulu, thank you for all the support and motivation you gave me during the course of this work, and to all the members of the centre for their support and assistance. Last, but not least, I would like to express my sincere appreciation to M.J Mokenela for his guidance and support he gave me throughout this study.

Table of Contents

List of Figures

List of Tables

Abstract

Searching and finding specific information from web-based information systems has become tedious and time consuming due to information overload on the web. The current adaptation and personalisation techniques employed are gradually becoming inadequate, as the information available on the web grows exponentially. Measures were taken towards improving the current solutions. The Generic Adaptation Framework (GAF) has been proposed as a standard for building personalised web-based systems. It identifies the standard adaptation components and how adaptation process should be done. The components of the GAF overlay model are only conceptualised to interact at syntax level, which limits interoperability between the components and subsequently the quality of the personalisation results (recommendations).

This work proposes semantic interaction of the GAF modelling components, which will not only improve the interaction, but also enhance the overall adaptation process. To enhance personalisation, this work introduced ontologies to model the knowledge about the components and ontology mapping to support meaningful inter-component interaction. The proposed solution, Ontology-based GAF (O-GAF), was applied in a recommender system and compared with the classical GAF recommender system. The experiments carried out showed the O-GAF performs better than the classical GAF in terms of its recommendation accuracy.

- CHAPTER ONE -

Introduction

1.0 Overview

Providing affordable IT based infrastructure is an area that has received a lot of attention within the research community in recent years. The challenge is aggravated if one considers low income communities such as rural based Small Medium and Macro Enterprises (SMMEs). The lack of IT infrastructure limits SMMEs from making their services easily available to their clients/customers.

Grid based Utility Infrastructure for Small medium and macro enterprises Enabling Technology (GUISET), is a research project at University of Zululand and its goal is to provide affordable IT based infrastructure for SMMEs (Adigun *et al*., 2006). The project aims to develop an infrastructure that would enable SMMEs to pool their resources and expertise together for the sharing and collaboration among themselves and their partners. The infrastructure is based on the concept of Service Oriented Computing (Web Services and Grid computing), Utility Computing and e-commerce (Adigun *et al*., 2006). The SMMEs deploy their services in the GUISET infrastructure as e-commerce applications. SMME clients utilise available services through a GUISET portal interface (Adigun et al., 2006). The GUISET portal provides an entry interface for the diverse SMME clients with different goals and different interest. To efficiently serve these clients, there is a need for the GUISET portal to be adaptive and personalise its content and services for each user. Personalisation in the presentation layer will enable services and information retrieval to be personalised to each SMME client. This shields the SMME clients from information overload and confusion through reducing the resource space to only what is relevant to the client. Relevance in GUISET context concept includes: relevant to the user and computing context, domain application and the current user task. So GUISET requires that relevance of personalised contents be with respect to the client's preference, user and computing context, domain of application and the current user task.

In recent years portals have been gaining importance in many usage scenarios. They represent a single point of access to personalised resources by incorporating various applications and processes into one homogeneous user interface (Andreas *et al*., 2008). Nowadays, organisations use portals comprehensively as a complete e-business solution providing users with a single-point of access to the enormous amount of company resources. Furthermore, with the arrival of Web 2.0 (Xiaobo *et al*., 2006), portals are gaining popularity as a gateway to community-driven resources.

However, with the continuously expanding growth of resources presented on the Web, it has become more problematic and time consuming to find relevant resources. In the case of a portal consisting of several hundreds of pages and links contributed by different members of a community, traditional hierarchical navigation is no longer efficient. This is due to the fact that it is not feasible for the administrator to come up with one optimal topology that would fit the navigation patterns of every individual user (Fedor *et al*., 2009). Therefore, the next generation of portals need to be more adaptive. Instead of providing all possible relevant information, portal systems should only present the interface relevant to the current needs of the user. There is, therefore, a need to develop portals that can reduce this large resource space to a smaller set that is more relevant to the user (Thanh *et al*., 2006). The corresponding process is called adaptation or personalisation when geared towards the user's needs.

Solutions to reduce the resource space to a smaller set relevant to the end user in the current web-based systems have been proposed over the years. However, most of these approaches depend on statistical and artificial intelligence techniques to provide personalisation. These solutions can be categorised into Collaborative filtering (Balabanovic and Shoham, 1997) and Content-based filtering (Joachims *et al*., 1997). Although these solutions have been tested and found to produce the best results so far, they have been criticised for being complex and computerised oracles that use heuristics to make recommendation hence cannot be interrogated (Aerts *et al*., 2003). In addition, these solutions also do not cope well with (i) the increasing amount of content on the internet, and (ii) the dependence of user preferences and needs on the context.

One of the solutions that have widely been adopted in order to provide the end user with content/resources relevant to the user task is user modelling. User modelling has been proposed as an integral part of any system that provides the systems with the information and criteria that can be used to personalise content to a given user. Taking user profile only into consideration neglects the context under which the user is interacting with the system and the domain of application, which makes this solution fall short of the GUISET requirements. In reality, needs usually change as user context and domain of discourse changes (Andreas et al., 2008). The adaptation process needs to take the user context and application domain into consideration as well. In a business context a user might organise travels, e.g. booking flights, hotels and cars and do his travel expenses. In a private scenario he might plan spare-time events, like checking the cinema program. Of course, his interests and preferences would be totally different in both context and obviously he needs access to totally different resources. A lot of solutions to enable the use of context and domain aspects in the personalisation process have been proposed over past years (Moore and Hu, 2006; Kramár and Bieliková 2012; Codina and Ceccaron 2009). Given these solutions, there have also been efforts to conceptualise a generic and comprehensive framework that defines the components needed to have efficient personalisation. Such efforts are defined in the evolution of Hypertext reference model, from Dexter model, Adaptive Hypermedia Model (AHAM) to the current Generic Adaptation Framework (GAF) (Knutov, 2009). The GAF advocates that for efficient personalisation of content in an adaptive system, the system should be aware of the user goal, domain of application, the user profile and the context. The GAF presents the adaptation process and the overlay model consisting of a goal model, user model, domain model and context model, which can be argued to be the key components necessary for efficient personalisation and more specifically for GUISET. Although the GAF meets the GUISET requirements, it also has its drawbacks. The GAF components are only conceptualised to interact at syntactic level, which this dissertation argues is not sufficient for efficient personalisation of web-based system in the presence of increasing information overload, the need to recommend new items and the openness of web-based systems.

To enable efficient personalisation for the GUISET portal, this work adopted the GAF as the base solution to achieving personalisation. However, owing to the interoperability problems the GAF has this work proposes the use of semantic web technologies such as ontologies to enable the use of rich semantics in the personalisation process to enhance interaction of the GAF modelling components. Different ontologies may not fully understand each other. This lack of understanding results in ontologies developed for each component in the GAF not able to efficiently interoperate. To enable ontologies developed for each component to understand each other, ontology mapping was used. We envisage that this intervention increases interoperability between GAF components and subsequently enable GAF to achieve a higher degree of adaptivity.

1.1 Statement of the problem

The GAF provides a holistic and generic solution to personalisation in adaptive hypermedia systems. It reiterates that for efficient personalisation any hypermedia adaptive system should have four key components; the goal model, the user model, the domain model and the context model. However, the state of the art on personalisation in the web is moving towards use of rich semantics in the personalisation process. This is believed to improve inter-component interaction between the software components making up the recommender system and subsequently improve the quality of recommendation (Sosnovsky and Dichava 2010). Lack of use rich semantics in the GAF has been identified as one of the major weaknesses of the GAF (Knutov, 2009). This can be addressed by adopting techniques emanating from semantic web technologies to provide the rich semantics. In an adaptive hypermedia system the data used to effect personalisation, and possibly the components themselves, may come from different entities. This leads to a lot of heterogeneities between the components and the data, which results in interoperability problems between the components.

In an effort to include semantics and address interoperability problems in the GAF, this research work seeks to answer the following research questions:

i. How can knowledge representation of the GAF modelling components be achieved in a way that can be processed and reasoned about by computers to provide adaptation effect?

ii. How do we enable GAF adaptation components to interoperate in a way that enhances the adaptation process?

1.2 Rationale

Nowadays, as more organisations use portals as a business solution, searching for the right information is a widespread problem for service consumers in e-business portals. The significance of this work lies in how personalisation can be enhanced to reduce the issues of information overload in portals. The beneficiaries of this work are e-business portal users such as GUISET users (SMMEs and SMME client) as service providers and service consumers.

Service providers: the service productivity is increased by cutting maintenance cost, the use of ontologies improves the way services are maintained; furthermore, services are dynamically generated at runtime. Service consumers: searching for the right services is a widespread problem in portals especially in e-business portals. The average user of an e-business portal is usually not an expert but rather a novice regarding finding the right web services. However this work will help improve the service consumer's Quality of Experience (QoE). First, by the portal infer user objectives. Second, help the user to discover the scope of information or services available. Third, delineate a relevant path to get to the resources required.

1.3 Goal and Objectives

1.3.1 Research Goal

The goal of this work is to provide semantic interoperability in the adaptation process with the aim to enhance personalisation on web portals.

1.3.2 Research Objectives

The goal of this research work was decomposed into the following list of objectives:

i. To review existing literature on adaptive systems and knowledge representation with the aim to identify the modelling components involved in the adaptation process.

ii. To review and reuse existing interoperability approaches to establish a way the GAF modelling components can interoperate and share knowledge in a meaningful way.

iii. To propose and design a model that will allow GAF components to semantically interoperate with each other.

iv. To build an experimental Prototype in order to evaluate the proposed model.

1.4 Research Methodology

In fulfilling the goal of this research work the main research method used is conceptual design. In achieving this various frameworks and models used in adaptive hypermedia systems were studied. This framework allowed the identification of necessary modelling components in an adaptive system. This knowledge was then used to propose a model.

The conceptual design method was supported by secondary research methods such as; literature survey, conceptual analysis and model formulation. The proof of concept was supported by three methods; prototyping, benchmark and controlled experimentation. Each of the secondary methods is discussed in the following section.

1.4.1 Literature survey

An extensive literature review was conducted and it provided the current state of the art in personalisation of adaptive systems and the prevailing issues on personalisation. The literature was also used in identifying knowledge representation approaches such as the ontologies for user modelling and interoperability. The results of literature survey were used to propose the ontology-based approach to personalisation in adaptive systems.

1.4.2 Conceptual analysis

A critical conceptual analysis of existing interoperability approaches was done with an aim to reuse existing approaches. The analysis focused on concepts such as knowledge representation and ontology mapping for a possible approach to interoperability of the modelling components.

1.4.3 Proof of Concept/ Model Evaluation

The design of the model was formulated based on two design criteria: knowledge representation and interoperability. The knowledge in the modelling components was represented in such a way that the system uses it to reason, infer new knowledge and make personalisation decisions. Interoperability of the modelling components that partake in the process of adaptation allowed the components to share knowledge unambiguously and interact in a meaningful way.

The model was implemented as a prototype, another non ontology-based prototype was implemented for the purpose of evaluating the proposed model against current existing personalisation systems. Controlled experiments were conducted evaluating the ontology-based prototyped model against the non-ontology-based system. The classification accuracy metrics were used to evaluate the two systems (i.e. f-measure, recall and precision)

1.5 Research Scope

This research was conducted under the general aim of enabling the GUISET portal interface to personalise its content to end users. A lot of solutions to personalisation in GUISET-like infrastructure have been proposed over the years. The researcher's investigation of literature identified the Generic Adaptation Framework as the foundational concept that a solution for personalisation in the GUISET portal can be built upon. Contemporary solution to personalisation in open systems emphasise the need for use of semantics in the form of ontologies in order to achieve a higher degree of personalisation. However, the GAF does not consider the use of ontologies, which is argued in this work to be one of the weaknesses the GAF has. The particular aim of this work was therefore to investigate how ontologies can be incorporated into the Generic Adaptive Framework in order to improve its performance in terms of the quality of its recommendation.

A solution that supports semantic inter-component interaction in the GAF was provided. However, due to the fact that getting real life data that can be used to test this solution was infeasible, the researcher resorted into using publicly available datasets that have over the years been used by a lot of similar research studies to validate solutions proposed for personalisation. One of these dataset is the movieLens dataset (http://www.grouplens.org/node/12). The data set dictated that the solution can only be tested using only two components of the GAF; the user model and the domain model. This does not compromise the validity of the conclusion drawn from the work since the solution provided is in the context of inter-component interaction, which was comprehensively evaluated in the interaction of the user model and the domain model component interaction.

Another aspect which is also important to note in this study is the fact although the solution is proposed to work in a GUISET environment, the discussion of the solution was not done in the GUISET context. The discussion of the solution was given in the context of the open problem within a solution (GAF) that can be adapted to address GUISET personalisation problems. As way of showing how the solution will be useful to GUISET, Section 5.5 discusses the findings of this study in the GUISET context.

1.6 Organisation of the Dissertation

The remainder of this dissertation is organised as follows:

Chapter 2 presents the background on the important concept which forms the foundation of this work. These concepts include: the adaptation process, Generic Adaption Framework, personalisation, user modelling and ontologies. This chapter also presents the adaptation components that are part of the process of providing personalisation. Chapter 3 starts by introducing personalisation in the context of this work and analysis on the state of the art on personalised systems and user modelling. It also reviews ontology-based approaches to user modelling and semantic interoperability. Chapter 4 presents the solution to the research challenges raised in this work. The ontology-based representation of the modelling components are presented and ontology mapping as a means to bring semantic interoperability in the components. Chapter 4 also presents the proposed model design and the model architecture. To evaluate the effectiveness of our solution approach, experiments were conducted. To achieve this, our proposed model is instantiated as a prototype. The design, implementation and evaluation of the prototype are presented in chapter 5. Chapter 6 concludes this work and presents limitations and future work.

– CHAPTER TWO –
Background

2.0 Introduction

As indicated in the previous chapter, this research work aims to enhance personalisation in adaptive systems, more specifically for GUISET. In this work, this goal is realised by utilising semantic technologies such as ontology based user modelling and ontology mapping. In this chapter we present important background concepts related to this study.

The chapter is organised as follows; Section 2.1 starts by identifying the adaptation process and the need for personalisation in web-based systems. User modelling and user modelling dimensions are presented Section 2.2. Section 2.3 discusses Generic Adaptation Framework and its modelling components. Section 2.4 presents knowledge representation formalisms and ontology languages. Section 2.5 discusses semantic interoperability approaches such as ontology matching.

2.1 The Adaptation process

The World Wide Web has become, with no doubt, the best known and widely used hypermedia system. In the first generation of Web-based systems the information was presented in terms of carefully authored hypermedia documents. This process involved manual creation of a static set of Hyper Text Markup Language (HTML) pages in order to bear information to the users. The popularity of the Web brought the need for Web-based systems to be interactive and publish up-to-date content. As a result, the Web Information Systems (WIS) emerged. These information systems use the Web to present information to the users (Isakowitz *et al*. 1998).

The Web's success enabled the world-wide publications of huge amounts of inter-linked information, allowing heterogeneous groups of users to navigate in a non-linear manner. This rapid growth of information and heterogeneity of users showed a shortcoming of traditional hypermedia systems, they offer the same page content and the same set of links to all the users. It turned out to be obvious that the "one size fits all" approach was not sufficient, demanding hypermedia systems to alter (adapt) themselves towards its user to better simplify this navigation over the information space. This necessity is addressed by adaptive hypermedia or adaptive Web-based systems (De Bra *et al*., 2004), (Brusilovsky 1996).

Adaptive Hypermedia Systems (AHS) have drawn substantial attention since the introduction of the Web. Today, there exist numerous AHS in several application domains with great diversity of capabilities (see, Brusilovsky and Mill`an, 2007; Ardissono and Goy *et al*., 2007; Kobsa, 2001; Henze, 2001). Some of the major categories of AHS include on-line help systems, educational hypermedia and information retrieval systems. This work adopted the definition of adaptive hypermedia systems (AHS) by Brusilovsky (Brusilovsky 2001) which states:

"Adaptive Hypermedia Systems are all hypertext and hypermedia systems that reflect some features of the user in the user model and apply this model to adapt various aspects of the system to the user."

Classification of AHS is further made between adaptivity (also called dynamic adaptation) and adaptability (static adaptation). The adaptable systems allow users to make certain changes to the system parameters and adapt their behaviour accordingly. On the other hand, adaptive systems adapt to the user automatically based on assumptions it has about the user needs. In this dissertation the following descriptions of adaptability and adaptivity will be used; in adaptability the presentation process is based on existing information that defines the state in which the user will use the generated presentation (Frasincar and Houben 2002), and adaptivity adapts its content automatically while the user is browsing.

Adaptive systems are further classified into two: hard-wired adaptation systems and system supported adaptation (Alexandros, 2009). The hard-wired adaptation allows users to select alternative characteristics for presentation and other features, among those already built in the system. Customisation is a typical example of hard-wired adaptation. The system supported adaptation is based on the principle that the system should be capable of identifying the situations in which adaptation is necessary, and further selecting and activating an appropriate course of action. This category of systems coincides with “self-adaptation”. The preceding description incorporates the notion that the system may also be capable of monitoring the user interaction so as to extract information about the user, verifying, improving, reassessing, and, if necessary, retracting assumptions known to be true for a given user, and which leads to the area of personalisation.

2.1.1 Personalisation

The growing usage of the Web accelerates the pace at which information becomes available online. Searching and finding specific information from the web-based information systems is becoming tedious and time consuming due to information overload on the web. Personalisation is seen as one major concept to solve the problem of information overload.

Web Personalisation is defined as the task of making Web-based information systems adaptive to the needs and interests of individual users (Mobasher *et al.*, 2000). Usually, a personalised Web-based system collects information about its users' preferences and adapts the services/content in order to match the users' needs. Personalisation aims to select content that is most relevant to the user from a greater volume of content and presenting the content in a way most suitable to the user. Figure 2-1 shows the process of personalising content to a given user.

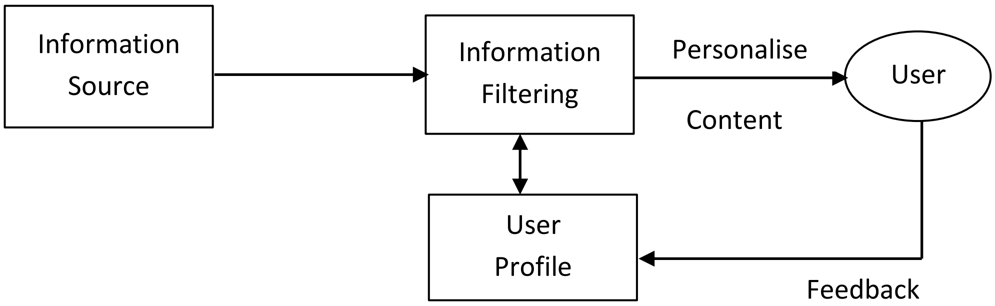


Figure 2-1 Structure of a personalised system

In e-business, personalisation offers mechanisms for the system to learn about its customer needs identify future trends and eventually increase customer loyalty. The personalisation of content and services is an important step in the direction of addressing information overload.

Personalisation is usually based on user models (see Section 2.2). A user model contains information about an individual's preferences, context and characteristics. The interpretation of user and usage data is derived from this information. This interpreted data is used by the system to infer suitable content for a given user.

2.2 User Modelling

User modelling has more than 30 years of research. It originated from lab projects and developed to the expansion of commercial user modelling servers supporting millions of users (Fink and Kobsa, 2000).

Many authors have proposed their explanations for the concepts of user models and user modelling. By restating (Wahlster and Kobsa, 1989) to adjust to a broader class of user-adaptive systems, in this work we define a user model as a knowledge base in a system that contains expectations on different aspects of the user that may be applicable to the system's adaptive behaviour. These expectations must be independent from the rest of system's knowledge. Consequently, user modelling is the field of Information Science that deals with elicitation, representation, and utilisation of user models. User modelling is the process performed by an adaptive system in order to create and maintain an up-to-date user model. There are two main ways the system can collect information about its users: by implicitly observing user interaction or by explicitly requesting direct input from the user. User modelling and adaptation are complementary to one another. The extent and the nature of the represented information in the user model depend to a large extent on the type of adaptation effect that the system has to provide.

This section is arranged into the following sub-sections: Section 2.2.1 discusses the different user modelling dimension such as, user knowledge, user preferences, user goal, user background and user context. Section 2.3.2 defines the Generic Adaptation Framework.

2.2.1 Dimensions of User Modelling

The basic idea of adaptation is based on the notion that different user characteristics should influence the individual effectiveness of the service/information provided; hence if system's behaviour is personalised according to these characteristics, the significance of the system will increase. Some adaptive systems store individual information only for a single characteristic, others model users along multiple dimensions. It is hardly possible to enlist all the dimensions utilised by all user modelling systems over the years. This section will describe the most important dimensions for this work.

User knowledge

The user knowledge is one of the most important user feature used by AHS types, providing adaptive navigation support, adaptive content presentation or personalised access. User knowledge is identified with the actual, acquired competencies.

Given a certain domain, different AHS express the user knowledge of the domain in various forms:

a) The Scalar Model approximates the level of user domain global knowledge by a single value on a particular scale, quantitative (e.g. scale ranging from 0 to 5) or qualitative (e.g. good, average, poor, none). Such models are used for example by AHS to adapt each page's content (Encarnação, 1997) or page fragments (Brailsford *et al.*, 2002) to users with dissimilar levels of knowledge. The weakness of the scalar model is its low precision. User knowledge can be rather dissimilar for different parts of the domain. For example, in the multimedia indexing domain, a user may be an expert in manually annotating the multimedia objects with some support tools, but a novice in automatic indexing algorithms. They adopted some structural models where they divide the domain knowledge into certain independent fragments and represent user knowledge of different fragments independently:

b) The Overlay Model represents an individual user's knowledge as a subset of the domain model, which reflects the expert-level knowledge of the subject. The user knowledge of each domain fragment could be expressed through a Boolean value, yes or no, indicating whether the user knows or does not know this fragment (VanLehn, 1988), through a Qualitative measure (good-average-poor), or a quantitative measure, such as the probability/measure that the user knows the fragment.

Interests and Preferences

Interests and preferences are often used as synonyms. They are the most utilised user characteristics in adaptive systems. The user interests are always considered as a part of the user profile in the adaptive information retrieval and filtering systems, as well as in the Web recommender systems.

When modelling user preferences, long-term and short-term preferences are differentiated. Long-term preferences are reasonably constant; they slowly evolve along the entire period of user’s working with the system. Short-term preferences are dynamic and usually occupied only in a single session, reflecting user’s current task. After the session is over, the influence of short-term model on the system adaptation ends. Sometimes systems need to model user’s preferences not related to the main task of the system, e.g. a preference over a certain type of the interface, or a certain language (Kay and Kummerfeld, 1994).

Goals, Plan, Tasks and Needs

The user's goal represents the immediate purpose for a user's work within an adaptive system. Modelling of user’s goals and plan has been broadly used in adaptive dialog systems.

Knowing what is the user trying to achieve (goal) and what sequences of actions will the user take (plan) is important for such system (Kass and Finin, 1988). Examples of such goals would be “To search for a certain information”, “To express an opinion”, “To get help”. User task and need are closely related to these modelling dimensions. Modelling of user’s information need or task is a very common approach in adaptive systems on the Web. Thus, for adaptive recommender systems it is argued that recommendations not taking into account user information task/need are likely to be meaningless and useless (McNee *et al*., 2006), or even harmful, since they would eventually lead to the lack of trust in the system. The following are some techniques that were adopted for detecting or approximating the user’s current goal in order to develop the user model:

i) User must select one of the predefined goals;

ii) User could declare a new goal and the system increasingly learn how to adapt according to this (Mathé and Chen, 1996);

iii) The user current goal is modeled as a probabilistic overlay of the goal catalogue, where for each goal the system maintains the probability that this goal is the current goal of the user (Encarnação, 1997);

iv) Data mining technologies could be applied in order to identify the current user task in an expected sequence of tasks and to provide personalised task-level support (Jin *et al*., 2005).

User Background

The user's background identifies some features associated with the user's earlier experience outside the main domain of a definite Web application; his or her profession, job responsibilities, work experience in related areas, his or her language ability, etc. For example, medical adaptive hypermedia systems can provide users with a different version of the content depending on their profession and responsibilities (e.g. student, nurse, doctor). Because the user's background does not change during his/her work within an AHS, the background is provided explicitly to the system. Generally, a stereotype modelling of the user's background is adopted, but also this could be modelled through one or more domain. The problem, in this case, is to find (or to define) some relations between the domain concepts and the core domain concepts (for example, what computer science competences request mandatory language competences).

Context

The user context is a broad concept; it includes any information about the user's time, location, the device being used, social and physical environment.

Once user features are known, the system needs to model them for providing personalisation in a specific domain. There are various approaches for user modelling.

The overlay model is commonly used in user modelling, more especially if a number of dimensions are to be considered. The Generic Adaptation Framework is one of the frameworks in AHS which employs the overlay model approach. The Generic Adaption Framework encapsulates these user features in four overlay modelling components (i.e. user model, goal model, domain model and context model). In the following sub-section we discuss the Generic Adaptation Framework.

2.3 Generic Adaptation Framework (GAF)

In defining the adaptation process some classification models were introduced in (Brusilovsky, 1996) and later revised in (Knutov *et al.*, 2010). A generic adaptation process was defined which serves as a process guideline and framework for defining the way an

Adaptive Hypermedia System (AHS) functions. Generic Adaptation Process means that the interaction in AHS starts with the goal statement, then exploits features of the user and domain models in different contexts and adapts several aspects of the system to the user (Knutov *et al*., 2010).

Considering a general purpose adaptive system one may think not only about the framework or the reference model but about what the adaptation process within the system looks like. Initially defined in (Brusilovisky, 1996) it is mostly known as a classical loop of 'user-modelling-adaptation.' This scheme explains how most of the adaptive systems work. Since its inception the loop has been improved and enhanced several times considering different domain areas and recent developments in the adjacent fields of Adaptive Hypermedia (AH). However Knutov (2009) looked a bit further and extended the classification of initial AH techniques and methods, with the adaptation process cycle to give the first insight of the Generic Adaptation Framework.

Figure 2-2 shows the evolution of the Hypertext reference models, from Hypertext to Adaptive Hypermedia to the new Generic Adaptation Framework (GAF) which encapsulates most recent developments in AH and adjacent fields (Knutov *et al*., 2011).

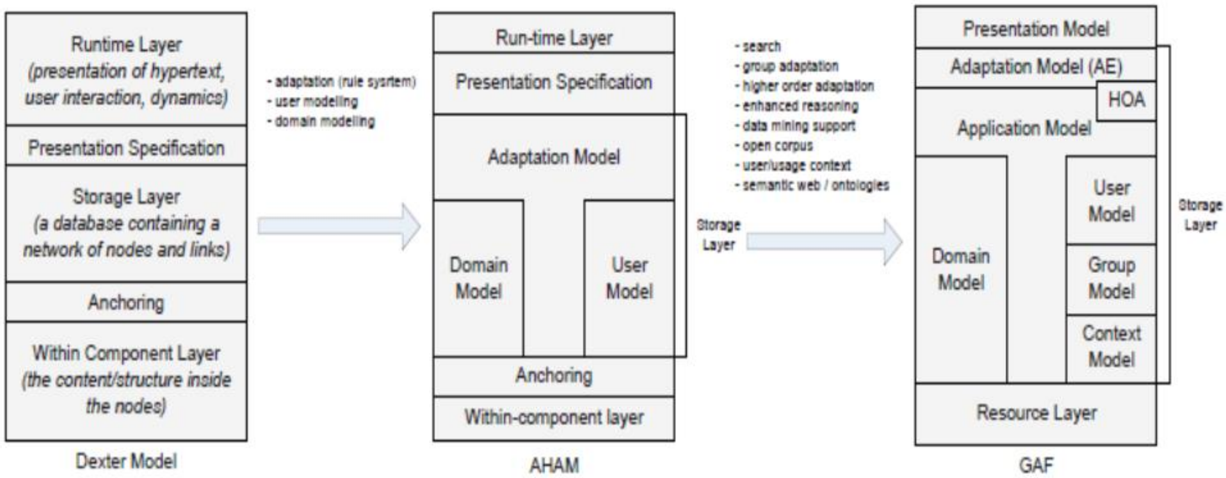


Figure 2-2 Evolution of GAF, from Dexter, through AHAM, to GAF (Knutov et al., 2011)

The GAF's evolution moved from Dexter model, Adaptive Hypermedia Model (AHAM) to the Generic Adaptation Framework (Knutov, 2009). Figure 2-3 presents the conceptual structure of GAF, the layers are aligned in the order according to the classification of AH methods and techniques. Though this order represents the basic understanding of the adaptation questions every particular system may vary or even omit some of these, thus leading to a different composition of the system layers determined by the different adaptation process. One of the goals of the GAF is to provide reference architecture for AHS describing necessary and optional elements. It, therefore, defines the criteria for distinguishing between these elements, and describes their functionality and interaction. It also defines a modular structure which can be used in describing and developing applications that fulfil different adaptation and recommendation needs. The GAF has a layered structure where the layers match with the original classification of adaptive hypermedia methods and techniques. This classification is used to describe recommendation functionality within the same adaptation system layers which depend on the requirements of the application and hence contribute to the system extensibility and heterogeneity.

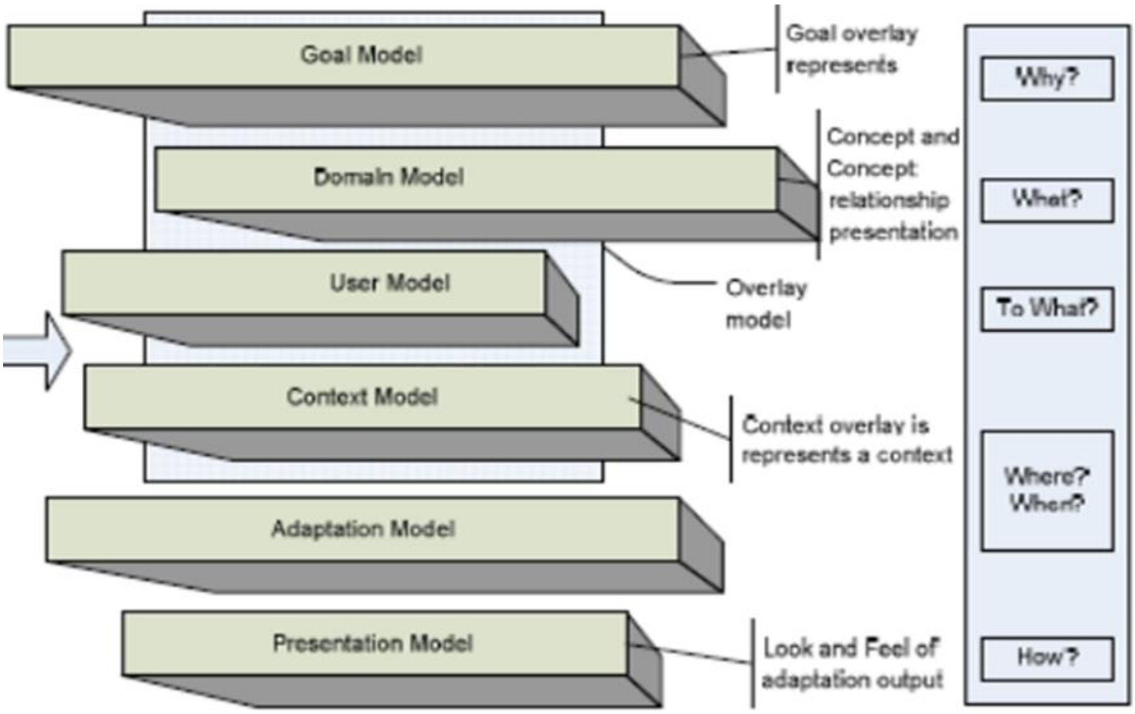


Figure 2-3 Conceptual scheme of a Generic Adaptation Framework (GAF) (Knutov, 2009)

2.3.1 GAF Modelling Components

This work focuses on the modelling components of the GAF; the GAF defines an overlay model were it defines the following components:

Goal Model – it models the instant purpose of the user task. The user goal may be stated by the user in the form of a query statement and key words, or the goal can be proposed by the system. The goal model is optional in recommender systems, they are commonly defined within the system and some applications are goal specific which may not allow the user to specify their goal. The goal is normally aligned with the domain model (availability of concepts and their structure).

User model – it models the user information including those stated in Section 2.2.1 user modelling dimensions (such as user knowledge, interests, preferences, and background). This information is obtained from the user in two ways; explicitly collected by asking the user directly, or implicitly collected by the system based on the user's activity.

Domain Model – in AHS the domain model is a conceptual representation of the real world; it defines concepts and their relationships. The concepts of a specific domain are modelled in a hierarchy.

Context Model – defines the user context properties such as time and other user activities. Modelling context gives a possibility to consider context-aware personalisation from the user point of view.

These components are referred to as overlay modelling components as they are the key components that capture knowledge in the process of providing personalisation.

In (Knutov, 2009) it is said that the GAF have some limitations and one of those limitations is interoperability in the modelling components. The modelling components interact at syntactic level, which is argued to be not sufficient for efficient personalisation in the presence of increasing information overload and openness of web-based systems. These components lack meaningful interaction giving rise to interoperability problems and consequently hinder the personalisation process. In order to provide meaningful interaction, one of the possible solution which this work is exploring is adopting semantic web technologies.

2.4 Knowledge Representation

The following definition of knowledge representation is given in (Davis and Szolovits 1993), and is adopted in this work:

“A knowledge representation (KR) is most basically a surrogate, a substitute for the thing itself, used to enable an entity to determine consequences by thinking rather than acting, i.e., by reasoning about the world rather than taking action in it; It is a set of ontological commitments, i.e., an answer to the question: In what terms should I think about the world? It is a fragmentary theory of intelligent reasoning, expressed in terms of three components: (i) the representation's fundamental conception of intelligent reasoning; (ii) the set of inferences the representation sanctions; and (iii) the set of inferences it recommends; It is a medium for pragmatic efficient computation, i.e., the computational environment in which thinking is accomplished. One contribution to this pragmatic efficiency is supplied by the guidance a representation provides for organising information so as to facilitate making the recommended inferences; It is a medium of human expression, i.e., a language in which we say things about the world.” (Davis *et al*., 1993)

When knowledge representation is applied in web-based systems the above definition is restricted (Bosch, 2006). The reason for such restriction is because knowledge representation is only used to describe content and formal aspect of web resources. The resulting description is expressed by a specific markup language for metadata: Resource Description Framework (RDF), see Section 2.4.2.1 for details.

2.4.1 Ontologies

The idea of ontology originally came from philosophy, with the meaning, study of existence. The ontology concept was then borrowed by the Artificial Intelligence and Computer Science researchers for defining knowledge representation formalism with the intention to model the view on the domain of discourse. There are a number of definitions of the formal ontology. The mostly cited definition and adopted in this work is provided in (Gruber, 1995) and was later adapted in (Guarino, 1998): referring to ontology as an explicit specification of shared conceptualisation of a domain. In other words an ontology provides a formal model of a conceptual structure of the domain. Experts use such models to form an understanding of a shared domain. Such models openly represent the details of the domain and convey the intended, essential meaning of the domain conceptual structure even if the simplified model will be adequate for the system's needs.

The development and use of ontologies in the Semantic Web is motivated by the fact that they provide foundation for knowledge mediation in between mediators, databases and systems referring to the similar or related entities. It is easier to develop, maintain and use less formal web ontologies. No organisation has the responsibility of developing a central ontology and frankly it will be virtually impossible to manage such ontology. Instead ontologies are developed to represent some specific domain based shared conceptualisation.

Ontologies on the Web are aimed at making the domain knowledge explicit and intentional, however, in practice, the difference between the ontological models used in Semantic Web applications demonstrate the absence of strict rules on this account. Some ontology developers restrict the representation to the simple hierarchy of concepts connected by the generic subsumption (or categorisation) relation usually expressed by *rdfs:subClassOf*predicate (Henze and Herrlich, 2004). Other researchers, however, represent domains as conceptual networks with complex and developed topology, with clear distinction between classes and their instances, using multiple relations, and exploiting the expressive power of description logic to enrich the ontology with axioms and constraints (Antoniou and Harmelen, 2003). Therefore, on Semantic Web there is a differentiation between the so-called lightweight and heavyweight ontologies. Models of the former kind are not semantically complete, however, are easier to develop and support. The heavyweight ontologies follow the principles of formal ontologies and provide their users with advanced inference capabilities based on well-formalised logical apparatus. This does not mean that the lightweight ontologies do not support reasoning at all. However, to benefit of advanced capabilities of the modern inference engine, one should develop a knowledge structure beyond the simple hierarchy.

2.4.2 Ontology languages on the Web

Tim Berners-Lee presented the layered architecture of representation technologies underlying Semantic Web vision (Berners-Lee, 2000). The basic standards Unicode and URI (Uniform Resource Identifier) supply the symbolic means for naming resources and referring to them. On the next level the family of XML-technologies provides general syntactic rules for resource description. The RDF-based technologies relying on XML syntax help to develop metadata models linking resources together. RDF-schema supports creation of new types of resources and links between them. Ontologies use extensions of RDF-model to define semantic vocabularies of concepts and relations. They provide the mechanism for term disambiguation and cross-model inference, hence different metadata documents can be merged together and new information about the resources could be discovered. The upper-level technologies provide tools for embedding description logic into ontologies, representing rules and, finally organising the trust-based infrastructure where the information is discovered and negotiated automatically. In this section we overview the mainstream representation technologies inhabiting the RDF.

2.4.2.1 Resource Description Framework (RDF)

Resource Description Framework (RDF) is the main representation mechanism for metadata on the web, standardised by the W3C. The basic element of the RDF model is a resource that represents which can be referred to by a URI. But a resource on its own has very little meaning. For a resource to uncover semantics it must be linked to other resources. RDF statements provide the mechanism for organising resources into triples “subject-predicate-object.” A triple represent a direct semantic link between two resources, resulting in interlinked resources from a semantic network. Figure 2-4, gives an example of such network, where the resource *http://www.w3.org/People/EM/contact#me* participates a subject in four RDF statements. Because of the information provided by these statement, we can say, that this resource designates a person with certain name, title and e-mail.

Figure 2-4 RDF Model Example, (Manola and Miller, 2004)

There exist two major serialisations for representing RDF graphs. Notation 3 (N3) format supports human-readable RDF documents (Berners-Lee, 1998). However, as the main purpose of RDF is to model machine-readable metadata, W3C proposed another RDF syntax, based on XML as the dominating format for data description on the Web (Beckett, 2004), (Lassila and Swick, 1999).

This RDF vocabulary is limited, therefore to create richer RDF there is a need for advanced tools. The W3C proposed the RDF vocabulary description language RDF Schema (RDFS), which is designed using RDF model.

2.4.2.2 Web Ontology Language (OWL)

Though RDFS is used represent ontologies, their expressiveness does not allow the formation of rich domain models outside simple class hierarchies. A more complex ontology language is required in cases where a heavy weight ontology that requires modelling deep-level domain semantics is being created. Well-defined semantics are essential for supporting machine interpretability, as they avoid possible ambiguities in the description interpretations. Hence the advanced the formality of the ontology language, the more compound reasoning it allows, the more beneficial the developed ontologies can be.

The Web Ontology Language (OWL) is recommended by W3C as the major formalism for the ontology representation on the Semantic Web (McGuinness and van Harmelen, 2004). OWL is built on top of the generic RDF model; it uses RDF/XML syntax; and the RDFS vocabulary (Connolly *et al*., 2001). It also introduces new features to handle semantically complete metadata models. OWL was built upon such rich representational formalisms as frame theory and description logics (Baader *et al*., 2003) (for a more detailed on OWL see) (Horrocks *et al*., 2003). As a consequence OWL can model all the major components of formal ontologies: classes (or concepts), individuals (or instances), properties (including class relations and instance attributes), and axioms (or constraints). OWL comes in three flavours;

OWL-Lite, OWL-DL and OWL-Full. The OWL flavours provide ontology developers with different levels of expressiveness:

OWL-Lite being simplest level, it permits simple specification of class hierarchy, properties and individuals. It tools for modelling constraints are limited (0 or 1 cardinality), property characteristics (symmetric, transitive, etc.) and class equality/inequality.

OWL-DL combines the features of OWL-Lite and description logic functionality (class conjunction, disjoint, disjunction etc.). OWL-DL provides the guaranteed computational completeness and the maximum expressive power.

OWL-Full includes the expressiveness of OWL-DL vocabulary combined with the freedom of RDF model. OWL-Full can handle the models resulted from the merging of OWL and RDF vocabularies. However, such models cannot guarantee the complete reasoning across them. While, OWL-DL is based on RDF, it introduces certain restrictions to it. RDF, on the contrary treats everything as a resource, and does not make such a distinction.

2.5 Semantic interoperability

Semantics refers to the study of meanings. It is the ability of a computer system to transmit data with unambiguous shared meaning. Semantic interoperability enables machine computable logic, inferences, knowledge discovery, and data federation between information systems (Heflin and Hendler, 2000). It is concerned not just with the packaging of data (syntax), but with the concurrent transmission of the meaning with the data (semantics). This is accomplished by adding data about the data (metadata), linking each data element to a controlled, shared vocabulary. The meaning of the data is transmitted with the data itself, in one self-describing "information package" that is independent of any information system. It is this shared vocabulary, and its associated links that provide the foundation and capability of machine interpretation, inferences and interoperability. Syntactic interoperability, which is a prerequisite for semantic interoperability, refers to the packaging and transmission mechanisms for data (Veltman, 2001). To realise semantic interoperability ontology matching approaches are used.

An ontology is shared specifications; a single ontology can be used to represent multiple data sources. The use of such shared vocabularies enables a certain degree of interaction between these data sources. However, this does not completely solve the integration problem, because it is not possible that all organisations and individuals on the Semantic Web will agree on using one common vocabulary or ontology (Visser and Cui, 1998; Uschold, 2000). It can be expected that many different ontologies will appear and, in order to enable inter-operation, differences between these ontologies have to be resolved.

2.5.1 Ontology matching

Ontology matching is the key toward semantic interoperability problem (Shvaiko and Euzenat 2011). It finds correspondences between semantically linked entities of ontologies. These correspondences can be used for different tasks, such as ontology merging, data translation or query answering. Thus, matching ontologies allows the knowledge and data conveyed with respect to the matched ontologies to interoperate (Euzenat and Shvaiko 2007). There are a number of ontology matching techniques available, they include: ontology alignment, ontology merging, ontology integration and ontology mapping.

Ontology alignment is concerned with the (semi-)automatic encounter of correspondences between ontologies; and ontology merging is concern with creating a single new ontology, based on a number of source ontologies; Ontology integration is the process of generating a single ontology in one subject from two or more existing and different ontologies in different subjects; ontology mapping is mostly concerned with the representation of correspondences between ontologies. The concept of ontology mapping is broad in scope refer to (Kalfoglou *et al.*, 2003) more details. Ontology mapping is used in the process of ontology alignment merging, and integration (Choi *et al.*, 2006).

2.6 Summary

This chapter have presented the background in the following areas: adaptive personalisation where we looked at the origin of the need for adaptive personalisation, the adaptation process and the challenges around it; looked at the area of user modelling in relation with personalisation in adaptive systems and the different user modelling dimension used to model user profiles. The chapter also looked at semantic web technologies such as ontologies and ontology mapping which is used as a solution towards user modelling and interoperability. This chapter have highlighted important concepts in these fields which are related to this study. The following chapter presents the state of the art and analysis towards realising the goal of this work which is enhancing personalisation.

– CHAPTER THREE –
Literature Review

3.0 Introduction

The increasing amount of available information through web-based systems (such as e-commerce portals, knowledge-based recommenders, hypermedia system, information systems etc.) is a major time consuming challenge when searching for relevant information. Adaptation addresses such challenges of information overload by identifying and presenting relevant content. There are different approaches to realise adaptation and the commonly known and wildly used is adaptive personalisation. In the previous chapter the Generic Adaptation Framework (GAF) was introduced, this framework encapsulates various methods and techniques for providing adaptation and personalisation in web-based systems. Some limitations were identified in the GAF such as interoperability, i.e. the GAF lacks emphasis of component interoperability which affects personalisation performance (Knutov *et al*., 2010).

As indicated in chapter one the goal of this work is to enhance personalisation and in achieving this goal we attempt to address the issue of interoperability in the GAF. In addressing this problem the chapter performs an analytical review of existing approaches towards providing enriched interoperability in the GAF, more specifically in knowledge-based recommender systems. For providing interoperability in knowledge-based recommender system, this work first need to consider how the knowledge should be represented for each system component as a source of knowledge, in such a way that it would facilitate unambiguous information sharing. Secondly, these sources of knowledge should be able to support semantic interoperability.

The organisation of this chapter is as follows: Section 3.1 introduces current personalisation approaches in knowledge-based recommender systems. An analysis of knowledge representation approaches is presented in Section 3.2. Section 3.3 presents an analysis of relevant approaches for supporting interoperability among system components.

3.1 Recommender Systems

Adaptive recommender systems are a sub-type of adaptive web systems, which attempt to infer the user's goals, interests, preferences and similarity between users, and building applicable content which are recommended to the active user (Brusilovsky *et al.*, 2001). Different recommendation algorithms require different types of backgrounds and input data in order to offer recommendations. The GAF offer user modelling, domain modelling and context modelling techniques and are required to provide personalisation in adaptive systems. Besides the GAF being a comprehensive generic framework for supporting personalisation, this work argues that, its effectiveness can be improved if; (i) it can use some knowledge persisted in some knowledge bases in making recommendation, and (ii) communication among the GAF components is aware of the semantics of other components. Therefore, a personalised recommender system should support such techniques and exploit the knowledge in the modelling components.

In general there are a number of recommendation approaches available in literature. They all basically deal with the same problem of reducing information overload in web-based systems. However, the input data and recommendation techniques are different for each approach. This section presents an analysis of the well-known recommender systems, it first looks as the classical recommender systems and their limitations which are later addressed with knowledge-based recommenders.

3.1.1 Classical Recommender Systems

Classical recommender systems are mainly classified into two; collaborative filtering and content-based filtering (Balabanovic and Shoham 1997).

Collaborative filtering

Assumes similar users share similar interests, so recommendations are based on user similarity (Herlocker *et al*., 1999). A collaborative filtering system recommends items to a given active user based on what users similar to him or her have shown interest in. The most commonly used collaborative filtering approach calculates similarity on the basis of item ratings. Users with similar item ratings to the active user are deemed similar to the active user and hence items they have shown interest in are recommended to the active user. User similarity can also be defined in terms of user demographics, user preferences, usage patterns, or any other data that characterise the users.

Collaborative filtering only keeps data about the users which is equivalent to the user model, it neglect the data about the domain. Recommendations are only based on the data about ratings between users no other information about the user nor the domain is used.

Although collaborative filtering can persist some data about users, it is too raw for it to be deemed to be knowledge. Neither is there a mechanism capturing expert knowledge that can be reused within the system. Each time a recommendation is to be done, collaborative filtering systems re-compute the items to be recommended. Further to this, the assumption of similar users have similar interest lessens the part of individual preferences and therefore personalisation is not fully employed (Cremonesi, 2009).

Content-based filtering

Content-based filtering use the concept of similarity of available items to the items that the active user has shown interest to make recommendations to the active user (Claypool *et al*., 1999, Joachims *et al*., 1997). Content-based filtering makes use of the features of the items as the domain data and users previous interests as the user data for the recommendation. Although it makes use of both user data and domain data, it is still not at knowledge level and such a representation structure capture only certain aspect of the content (de Gemmis *et al*., 2009). Since the system continuously recommends more of what the user has previously indicated an interest for, there is a possible problem of overspecialisation, reducing the possibility of unexpected finds.

Hybrid algorithms

In an effort to leverage on the strength of collaborative and content-based filtering and reduce the effect of their weaknesses, Hybrid approaches of the two have widely been proposed in literature (Balabanovic and Shoham., 1997; Burke., 2002)**.** Hybrid algorithms combine two recommendation algorithms to deliver better results with fewer of the disadvantages of an individual algorithm. Even in such a setting the problem of raw data representation is still an issue that needs to be addressed.

Collaborative and content-based filtering are statistical-based and each has its own limitation, their major drawback is the use of raw data and their representation structure. The models they develop are black-box-like and are not open to user inspection or adjustment. However, apart from the characteristic difficulties in gathering, and especially, in exchanging model information, these approaches have been criticized to be complex (i.e. nature of the underlying algorithms), computerised oracles, which give advice but cannot be interrogated (Aerts *et al*., 2003). In view of the foregoing, it was realised that there is a need for recommendation algorithms which do recommendation in a manner that is understandable and can be interrogated to improve the recommendation accuracy. This can only be made possible if the recommender system has a mechanism of holding or representing the knowledge that is used in the recommendation process. It is in view of this, that knowledge-based recommender systems were proposed.

3.1.2 Knowledge-based Recommender Systems

Knowledge-based approach avoids some of the drawbacks in collaborative and content-based filtering. Because they are based on knowledge of the domain, they are immune to statistical anomalies (Burke, 1999).

Knowledge-based recommender systems (Burke, 1997) use knowledge about the users, domain and any other knowledge relevant to the specific user task to make recommendations to a given user. Such knowledge is used to see how a particular item meets the needs of the active user, they can therefore reason about which items/content meet the active user's requirements (Heitmann and Hayes 2010), (Bruke, 2002).

The knowledge in knowledge-based recommender systems can be in any structure that supports inference. Their knowledge structure may be as simple as in Google where the knowledge structure is formulated based on the user query or as detailed user model representation (Towle and Quinn, 2000). Other knowledge-based recommender systems make use of Case-based reasoning to make recommendation such as in (Bruke, 2002; Schafer *et al*., 1999).

Another important distinction found in knowledge-based recommender systems is their knowledge representation. There are various knowledge representation approaches available in literature. The following section discusses knowledge representation in knowledge-based systems.

3.2 Knowledge Representation in knowledge-based Recommender system

Knowledge is exchanged between system components because this way each component gets access to more than the knowledge it has been able to build up. Each component, and the system as a whole, is then more prepared to make the correct choices. It is important how this knowledge is represented since it will affect the reasoning process, how it's manipulated and how it's understood across system components. Knowledge propagation can be compromised if there is no proper representation structure (Davis and Szolovits, 1993). Therefore, knowledge representation should; (i) support semantic reasoning, (ii) allow common understanding (i.e. unambiguity) and (iii) allow sharing of knowledge across system components. This section explores knowledge representation techniques and finds the most suitable approach for the GAF knowledge-based recommender systems.

3.2.1 Knowledge Representation Techniques

Knowledge representation appears in different forms and the most dominant are based on semantic networks, rules and logic. These techniques are discussed below.

Semantic Network

Semantic networks formalisms (Mika, 2005) concentrate on expressing the taxonomic structure of categories of object and relationships between them. They provide a structural representation of statements about a domain of interest. In principle, the concepts and relations in a semantic network are generic and could stand for anything relevant in the domain of interest. For example Figure 3-1 illustrate knowledge expresses in semantic network form.

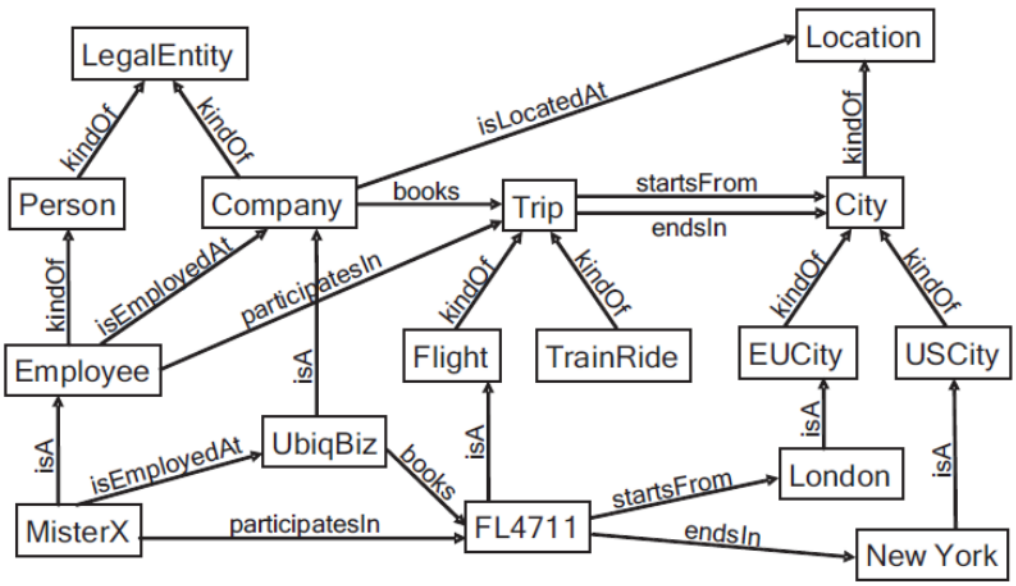


Figure 3-1 A semantic network for business trip (Mika, 2005)

The fragment $\boxed{\text{Company}} \xrightarrow{\textit{books}} \boxed{\text{Trip}}$ is read as "companies book trips". However, this statement is not clear, one cannot tell whether every company books trips or just some companies and also companies can book trips only for their employees or it can also book for employees of partner companies. Therefore, semantic networks are vague and ambiguous; their semantics are not precise (Studer *et al*., 2007).

Rules

Another natural form of expressing knowledge in some domain of interest is rules that reflect the notion of consequence (Yang *et al.,* 2003). Rules come in the form of IF-THEN-constructs and expressing various kinds of complex statements. Similar to semantic networks the set of rules expressed by rule formalism their exact meaning are still unclear (Studer *et al*., 2007), resulting in ambiguity.

Logic

Semantic networks and Rule formalism can be formalised using logic (Baader, 1999) to give them precise semantics. Without such a precise formalisation they are vague and ambiguous, and thus problematic for computational purposes.

Logic formalisms can be roughly categories in to three groups (Kutz *et al*., 2004): Very expressive but undecidable, higher-order-logic; Quantifier-free, classical propositional logic; Decidable logics, first-order-logic. The most prominent and fundamental logical formalism classically used for knowledge representation is the first-order logic (Studer *et al*., 2007). They are motivated by the fact that classical propositional logic is not sufficiently expressive and higher-order-logic are too complex for efficient reasoning in realistic application domain (Wu and Chen, 2008). The challenge in the first-order-logic approach is finding a balance between expressiveness and effectiveness, the idea is a language that is a compromise between the two.

Ontology

In logic, the existential quantifier is a notation for asserting something exists. But logic itself has no vocabulary for describing the things that exist. Ontology fills that gap; it is the study of existence (Wu and Chen, 2008). The two main sources of ontological categories are observation and reasoning. Ontology provides observation of knowledge in some domain of interest, and reasoning making sense of observation by generating a framework of abstraction called metaphysics (Wu and Chen, 2008). Ontology as a specification of conceptualisation (Guarino, 1998) provides knowledge sharing and a common understanding of the structure of information among system components. Ontologies can be represented as graphs (semantic networks), logic (description logic) and a conceptual hierarchy. The current standard for ontology representation language is OWL (see section 2.4.2), its combines formal logic, description logic, and web standards.

3.2.2 Ontology-based Knowledge Representation in Knowledge based Recommender Systems

It has become essential for a knowledge-based recommender system to personalise or adapt its content towards an individual user, in order to achieve this they must have a model of assured characteristics of their users (see Section 2.2) and possess techniques to employ the information in those models. To represent user knowledge, goals and interests adaptive systems make use of explicit user models (Bhowmick *et al*., 2010). The adaptive systems use these models in providing personalised experience for their users. Knowledge-based recommender systems also make use of such models together with domain models.

Ontologies have been the basis for the construction of a user model in personalised systems (Middleton *et al*., 2002). Ontologies are a knowledge representation technology and in knowledge-based systems are used for defining relationships between concepts and to infer facts (Stuart *et al*., 2001). The use of ontologies in user modelling has been recognised as beneficial by several researchers (Sosnovsky *et al*, 2010), (Bra *et al*., 2004), (Razmerita *et al*., 2003), (Bouzeghoub *et al*., 2003) and (Kalfoglou *et al*., 2001).

The variety of ontological approaches designed for representing user models comes from the enormous pool of research that has been done in the past for user model representations and from the prospects. Ontologies are being seen as the new generation of adaptive systems (Middleton *et al*., 2002) and (Bhowmick *et al*., 2010) . There are two principle directions that can be identified summarising the whole field. A knowledge based recommender system can use ontologies for modelling the structure of the domain and represent atomic user characteristics based on the elements of the ontology (Bhowmick *et al*., 2010). Another trend is to structure a user profile modelling several dimensions of user's state as ontologies. The first direction inherited its ideas from overlay user modelling (see section 3.2.1), while the second direction relate to the user modelling dimensions and has origins in stereotype user modelling (Sosnovsky, 2008). A small set of research focuses on enabling keyword-based user modelling employing lexical ontologies such as WordNet.

Ontologies play a prominent role on the semantic web; they could be key element in many applications such as information retrieval, web services search, knowledge-based recommender systems. It is believed that ontologies will contribute to solve the problem of interoperability between software components and applications, providing a shared understanding of common domains (Amrouch *et al*., 2011).

This approach enables introduction of ontologies in the GAF overlay model which permit extension and modularisation of the framework, allowing the system components to work together. Each component is specified by a corresponding ontology, providing common ground for knowledge sharing, exploitation and interoperability among the components. This leads to a highly modularised architecture that offers a high degree of flexibility.

3.3 Supporting interoperability between System components

The exchange of information is a crucial factor in information systems employing the GAF structure. Many activities in the business world involve different organisations that have to work together, and use existing information whenever possible, in order to reach a common goal. Similar activities can be found with the overlay modelling components in a knowledge-based system. Interoperability as defined in Section 2.5 is the ability of two or more entities (such as system components) to exchange information and to use the information that has been exchanged (Abel, 1998). There are two main levels of interoperability that can be distinguished: *signature level* (syntactic; names and signatures of operations), and *semantic level* (the meaning of operations) (Juan Hern´andez, 2000). They are also referred to as "*static*" and "*dynamic*" levels. Syntactic interoperability is well defined and understood. However, it has been recognised as insufficient for ensuring the correct development of large component based applications in open systems (Juan Hern´andez, 2000). The existing proposals at the semantic level endow component interfaces with information about their behaviour and meaning. Approaches to semantic interoperability such as ontology matching are discussed in this section.

3.3.1 Semantic Interoperability

To achieve semantic interoperability, concerned components must be able to exchange data in such a way that the precise meaning of the knowledge is readily accessible and it can be translated into a form that it understands across system components. For two components to semantically interoperate, they must be able to communicate with each in a meaningful way (Juan Hern´andez, 2000). Only then can the two components know how to use information from the other components.

The benefits of semantic interoperability are numerous. For example, search can often be frustrating because of the limitations of keyword-based matching techniques. Users frequently experience one of two problems: they either get back no results or too many irrelevant results. The problem is that words can be synonymous (that is, two words have the same meaning) or polysemous (a single word has multiple meanings). However, if the languages used to describe the content were semantically interoperable, then the user could specify a query in the terminology that was most convenient, and be assured that the correct results were returned, regardless of how the data was expressed in the sources.

Ontology, as an explicit information modelling method, is viewed as a fundamental approach to support semantic interoperability (Heflin *et al*., 2000). However, in the Semantic web community it has been agreed that a single, universal ontology cannot be built because of the large variety on information sources on the web, documents on it will inevitably result from many different ontologies (Amrouch *et al*., 2011). For these reasons, enabling interoperability among system components with an ontological knowledge representation is still a key challenge. Ontologies can only interoperate if correspondences between their elements have been identified and established. If two ontologies need to interoperate, matching is mainly achieved manually (Godugula *et al*., 2008). But this tusk is tedious, error-prone, and time consuming. Therefore, the introduction of new methodology and user-friendly tools that support the knowledge engineer in discovering semantic correspondences is crucial. There are various approaches which have been proposed and used for ontology matching in realising semantic interoperability. In the following subsection, the existing approaches to establish correspondences between different ontologies for data integration are discussed.

3.3.2 Ontology Matching

Tasks on component based systems (such as the GAF) may demand support from more than one ontology, therefore multiple ontologies need to be accessed. A system may consist of multiple ontologies, each information source described by an ontology to represent conceptualisations of a specific domain. Conceptualisation of each domain may be too disparate to be integrated into a common global ontology and various components with different ontologies do not fully understand each other. A lot of research has been done in the area of ontology matching (e.g. Godugula *et al*., 2008, Dou *et al*., 2005).Such works (propose ontology matching techniques as a means to archive data integration and interoperability. Ontology matching is the key interoperability enabler for the Semantic Web (Ritze *et al*., 2010).

Ontologies are taken as input and determine alignments as output, that is, a set of correspondences between the semantically related entities of those ontologies. Matching ontologies enables the knowledge and data expressed in the matched ontologies to interoperate. Ontology matching can be divided into sub-approaches: ontology mapping (Choi *et al*., 2006), which is mostly concerned with the representation of correspondences between ontologies; ontology alignment (Noy *et al*., 2000), which is concerned with the (semi-)automatic discovery of correspondences between ontologies; and ontology merging (Pinto *et al*., 1999), which is concerned with creating a single new ontology, based on a number of source ontologies; ontology integration (Pinto *et al*., 1999), which is the process of creating a single ontology in one subject from two or more existing ontologies from different subjects. However, as defined in Section 2.6 ontology mapping is required in the process of ontology merging, ontology alignment and ontology integration (Choi *et al*., 2006). Ontology merging and integration combines two ontologies which is not ideal when the two ontologies define conceptualisation from different domains. Ontology mapping is the dominant approach employed in the other methods; it fines similar correspondences between two ontologies similarly to the alignment approach. Therefore, the remainder of this section explores ontology mapping techniques with the aim of finding a suitable technique for supporting interoperability in the GAF overlay model.

3.3.2.1 Ontology Mappings Approaches

Ontology mappings are used to overcome interoperability issues previously raised. Ontology mapping plays a vital technical role in the course of integration of distributed information retrieval applications. Ontology mapping can bring a mediation layer from which multiple ontologies could be accessed and hence exchanging information in a semantically sound way, i.e. ontology mappings map a term T_1 of Ontology O_1 to another term T_2 of Ontology O_2, such that the axiom if $T_1 = T_2$for any axiom in O_1 with T_1, its substitution axiom with T_2 also holds. Vocabulary and semantic expressions are mapped across different conceptualisations to resolve representation transformations with different focuses.

A survey done in (Zuo, 2006) grouped related projects into three groups regarding their usage of ontology mapping solving data integration and interoperability issues: Syntactic mappings which support schematic integration of relational databases; Vocabulary mappings which support terminology integration, Semantic mappings to support the integration of different meanings. Vocabulary approaches are heuristic rather than being a formal method applied in syntactic system. It tends to focus more on solving terminology heterogeneity amongst application systems. Semantic mapping approach is more flexible and uses representation languages, such as DAML+OIL and OWL. Semantic mappings are expressed in terms of sub-sumption relations between conceptual terminologies and instance sets in the same knowledge domain (Zuo, 2006).

Several ontology mapping algorithms have been proposed in literature to address the ontology interoperability problem. Table 3-1 below presents a brief description of the most relevant algorithms for semantic interoperability. A set of evaluation criteria to compare ontology mapping tools are given in (Noy and Musen, 2002). The following are the criteria used in selecting an ontology mapping technique for the GAF:

Input requirement: This criterion is concern with what element form the source ontologies are required. This includes knowledge representation formalism, ontology language and ontology structure.

Strategy used: this concerns the strategy the approach use for mapping, where some approaches are linguistic based and others are based on analysis and evaluation of the structure and model similarity.

Level of automation: this refers to the level of automation that is used, where the approach can be manual, semi-automated or fully automated.

Type of output: This refers to, in what form is the desired output. The results of the ontology mapping process is import, the desired output depend on what is it going to be used for.

Table 3-1 shows a summary of the ontology mapping approaches and their classification along the above mentioned criterions. H-match, AnchorPROMPT and H-CONE support OWL for description logic and AnchorPROMPT has an advantage over the others for also supporting RDFS language. AnchorPROMPT, Glue and QOM support all the entities in the ontology structure i.e. mapping is done across concept, properties and instances. All approaches except Glue support semantic mapping which is important for semantic interoperability and reasoning. The level of automation depends on the involvement of the user, such that the aim is to minimise the users' involvement in the process. The output format is very important as it determined how will the mapping results be used and processed.

For supporting interoperability in the GAF modelling components semantic interoperability is employed and to achieve semantic interoperability various approaches are available. However, for the context of this work ontology mapping are the key enablers of semantic interoperability and after an analysis of ontology mapping approaches, AnchorPROMPT is found to meet the requirements for proving semantic interoperability of the GAF modelling components.

Table 3-1: features of ontology mapping approaches (Noy and Musen, 2002)

Approach	Ontology Language	Ontology structure	Mapping strategy	Level of automation	Type of output
H-Match	OWL	concept	Semantic contextual	automated	
Anchor-prompt (Noy & Musen 2000a)	OWL/RDFS	Concept, properties, instances	Semantic heuristic	Semi-automated	Set of pair anchors (concept), can be Merged to ontology
H-CONE (Kalfoglou et al., 2003)	OWL	Concept, properties	Semantic reasoning	Semi-automated	
IF-Map (Kalfoglou et al., 2003)	RDF, KIF, Ontolingua, ProtegeKB, Prologo	Concept,	Semantic Heuristic reasoning	automated	RDF format Relations and concepts.
ONION (Kalfoglou et al., 2003)	DAML+OIL, IDL, XML based	Concept	Semantic	automated	Sets of Articulation rules between two ontologies
Glue (Doan et al. 2004)		Concept, properties, instances	Probability	automated	A set of pairs of similar concepts
QOM (Ehrig & Staab 2004)	RDF	Concept, properties, instances	Semantic heuristic	Semi-automated	table of correspondences between the ontologies

3.4 Summary

The chapter presented an analysis of existing interoperability approaches and selecting a suitable solution for providing interoperability in the GAF more specifically knowledge-based recommender system. In providing interoperability among system component some factures needed to be considered such as knowledge representation formalism of the sources of knowledge and the representation language. Various approaches for providing personalisation in recommender systems were discussed. Ontologies were identified as a knowledge representation tool and the OWL ontology language supports description logic for semantic reasoning.

Therefore, semantic interoperability was seen as the possible solution for supporting interoperability in the GAF knowledge-based recommender system. In providing semantic interoperability various approaches were reviewed and ontology mapping was found to be the most adopted because of its advantages over the other approaches. In realising semantic interoperability the Anchor-prompt mappings approach was adopted. The literature results obtained in this chapter informed the design of our solution approach presented in the following chapter (Chapter Four).

– CHAPTER FOUR –

Design of the Ontology-Based Generic Adaptation Framework

4.0 Introduction

In Chapter three, the Generic Adaptation Framework (GAF) was discussed as a generic and extensible building block for Adaptive Hypermedia System (AHS).

The GAF introduced the overlay modelling components which were also discussed in Section 3.2.1, as the key components that capture knowledge in the process of providing personalisation in adaptive systems. These modelling components interact at a syntactic level, which was identified in literature to be insufficient for effective personalisation of adaptive web-based systems in the presence of increasing information overload and openness of web-based systems. This work argues that these components lack meaningful interaction giving rise to interoperability problems and consequently hinders personalisation.

In order to address these previously mentioned challenges, research questions were raised in Section 1.1, which were to achieve knowledge representation in the GAF modelling components and provide semantic interoperability within the systems' components. In order to achieve knowledge representation and meaningful interaction in the GAF, concerned components must agree upon concept structure and meaning. To realise this goal, ontology-based approach to knowledge representation was employed. This provided the necessary semantics to contextualise information and ontology mapping to enable semantic interoperability across heterogeneous components. In the rest of this thesis we refer to this solution as the Ontology-based GAF (O-GAF).

The chapter is organised as follows: Section 4.1 presents the design requirements for the O-GAF. Section 4.2 discusses how the first design requirement was met which is knowledge representation and what informs it. Section 4.3 presents semantic interoperability and how it was achieved and Section 4.4 presents a similarity score algorithm. The O-GAF architecture is presented in Section 4.5.

4.1 Design Requirements

The Generic Adaptation Framework (GAF) overlay modelling components were addressed in literature and its prevailing issues. This section addresses the research issues raised in Section 1.1, which were to achieve knowledge representation in the GAF modelling components and such representation should address the GAF interoperability issues raised in the previous chapter. The O-GAF is meant to address these research issues and following are the design requirement considered:

i. **Knowledge representation:** in order to represent knowledge in a way that the system can process and reason about to provide adaptation effect, there is a need to use a specific knowledge representation technique (Chandrasekaran *et al*., 1999). Therefore, knowledge representation should be provided in order to allow knowledge to be successfully shared and understood across the systems components and allow reasoning. In literature review a number of knowledge representation techniques were discussed and reviewed. Ontologies (OWL/RDF) were exploited as the basis for knowledge representation allowing semantic reasoning.

ii. **Interoperability:** The GAF identified some modelling components where each component was represented with ontology. The O-GAF needs support from more than one ontology and therefore accessing multiple ontologies, however different ontologies do not fully understand each other (Choi *et al.*, 2006). The distributed nature of ontology development process leads to dissimilar ontologies. Therefore, there is a need to have a mechanism that allows different ontologies to cooperate in addressing interoperability problem between components while keeping the components independent from one another. Component interoperability should also be provided at a semantic level allowing the components to successfully interoperate in a way that enhances the adaptation process. Among other interoperability solutions ontology mapping was employed to allow the ontologies to meaningfully share knowledge unambiguously and allowing semantic interoperability in the modelling components.

4.2 Knowledge Representation

As previously discussed in Section 3.3, knowledge representation and reasoning aims at designing computer systems that reason about a machine-interpretable representation of the world, similar to human reasoning. Knowledge-based systems have a computational model of some domain of interest in which symbols serve as surrogates for real world domain artefacts, such as physical objects, events, relationships, etc. (Wu and Chen, 2008).

In order to address the first research issue (in Section 1.1) ontologies were realised as knowledge representation and this section discusses how this was achieved. Ontology-based modelling requires each constituent components be modelled using an ontology and in this work four modelling components which the O-GAF defines each having an ontological representation. The following subsection presents what informed knowledge representation of the O-GAF.

4.2.1 Knowledge representation formalism

Knowledge representation appears in different forms (see Section 3.2.1). However, ontologies are the core of knowledge representation in Semantic Web and other related fields. The knowledge in GAF modelling components was represented with ontologies. Ontologies as a knowledge representation in the GAF overlay model allowed; structural and natural form of expressing domain knowledge, expressing taxonomic structure of classifications of objects and their relationships in the user model and context model.

The benefits of using ontologies in personalised web-based systems have been recognised in many research works (Sosnovsky *et al*., 2010; De Bra et al., 2004; Razmerita *et al*., 2003). The following benefits were recognised in the ontology-based GAF in providing personalisation:

Explicit semantics– The domain model which was used to describe and index information resources provided semantics which assisted personalisation process in better understanding how domain knowledge aligned with user query and user's interests.

Formal representation–The knowledge sources of the O-GAF (i.e. ontologies) provided means of formalising information resources about the specific domain knowledge. As it is on the Web, in the O-GAF knowledge sources each information resource had its own identifier provided, specified as a *Unified Resource Identifier (URI)* which is globally unique.

Formal reasoning – The formal representations of knowledge in the O-GAF enabled formal reasoning on top of the ontologies, enabling the O-GAF in making personalisation decisions (see Section 4.2.3).

4.2.2 Ontology Language

To make ontologies available to knowledge representation various concrete ontology languages have been designed and proposed for standardization. Such as RDF and RDFS, are well established and widely accepted standard for encoding meta-data (Grosof *et al*., 2003). On top of RDFS there is OWL family of languages, which comes in several levels with increasing expressiveness. OWL-DL is currently the most prominent Semantic Web ontology language following the description logic paradigm built on top of RDFS.

The current ontology language standard OWL was used in the O-GAF ontologies; as it combines formal logic, description logic, and web standards. OWL-DL's less expressiveness and being decidable allowed reasoning on the ontologies which is of more importance in personalisation than expressiveness. Such representation allowed knowledge sharing and common understanding of the structure of information in the domain among other entities (i.e. user model, goal model and context model).

4.2.3 Reasoning

Ontologies as knowledge representation in the O-GAF provide the opportunity for personalised systems to reason about their knowledge, inferring new knowledge to make personalisation decisions. The knowledge in the knowledge-base is processed and new statements are derived for drawing conclusions.

To archive knowledge reasoning in the O-GAF, SPARQL11 was used. SPARQL is an ontology query language and the most recent RDF query languages, which is used to query meta-data. Therefore, SPARQL was used to query the source of knowledge for each modelling component providing efficient and effective access to the knowledge. In addition to query language, different reasoning technologies for are available to check for inconsistences in the ontologies. The most common used reasoners are those that use Description Logics reasoning (DL) such as the OWL-DL ones: Pellet12, Racer13 and Fact++14 (Codina and Ceccaron, 2009). The ability to process explicit knowledge computationally allowed O-GAF system to reason over a domain of interest by deriving implicit knowledge that follows from what has been told explicitly.

4.3 Semantic interoperability

The previous section presented the modelling components and their ontological representation. As in any personalised system the modelling components need to interact. While traditional approaches such as the GAF interact at syntactic level, it was recognised in the previous section that there are limitations with syntactic interaction that are addressed with semantic interaction. To achieve semantic interaction or a commonly used term semantic interoperability in the overlay modelling components ontology-based approach was employed as representation of these components, allowing the meaning of information that is interchanged to be understood across the system’s components.

Tasks on component based systems that make use of ontology may demand support from more than one ontology and, therefore, multiple ontologies may need to be accessed. As presented in section 4.2, each component has an ontological representation. However, different ontologies do not fully understand each other. The distributed nature of ontology development process has led to dissimilar ontologies within a system. To address this problem, it is necessary to have a mechanism that allows different ontologies to cooperate in addressing an interoperability problem between components while keeping the components independent from one another. Against this background our solution adopts the use of ontology mapping as a tool for achieving interoperability among the modelling components. We envisage that this intervention will increase interoperability between these components and subsequently enable the system to achieve a higher degree of personalisation.

4.3.1 Ontology mapping

There are various approaches towards achieving interoperability, as discussed in Section 3.0 this work employs semantic interoperability which can be archived through ontology mapping and reasoning. Various ontology mapping methods are also available and the choice of using one method over the other depends on the ontology language and structure of the ontologies this means the level of mapping (i.e. concept, properties and individuals) see Section 3.4.2 Table 1. Anchor-Prompt approach is based on OWL and RDFS ontology languages and the mapping is across concept, properties and individuals, therefore Anchor-Prompt was the suitable ontology mapping method to be employed in the O-GAF.

An ontology is define as a pair O = (S, A), where S is the ontology signature which describes the vocabulary, and A is a set of ontological axioms, these specify the intended interpretation of the domain. The ontological signature is modelled by a mathematical structure. The vocabularies of two ontologies are related in such a way that the structures of the ontology signatures are respected. These structure preserving mappings are called morphisms. Therefore, ontology mapping are modelled as follows (Godugula *et al*., 2008):

A total ontology mapping from $O_1 = (S_1, A_1)$ *to* $O_2 = (S_2, A_2)$ is a morphism $f: S_1 \rightarrow S_2$ of ontological signatures, such that, $A_2 \models f(A_1)$, which means that all the interpretations that satisfy O_2 satisfy O_1. Of course, in reality, it is difficult to attain these total mappings and therefore there is the notion of a *partial ontology mapping* from $O_1 = (S_1, A_1)$ *to* $O_2 = (S_2, A_2)$ if there exists a sub-ontology $O'_1 = (S'_1, A'_1)$ (S'_1 subset of S_1 and A'_1 is a subset of $A1$) such that there is a total mapping from O'_1 to O_2.

To realise semantic interoperability in this work an ontology mapping technique adopted that attempts to automatically find semantic similarity between two concepts from different ontologies using partial ontology mapping:

Similarity $S(C_1, C)$ where C_1 and C are two terms (C_1 is from ontology O_1 and C is from ontology O). Figure 4-1 shows the ontology mapping algorithm. It takes two ontologies as input and finds a set of pairs of related terms by computing their similarities. The tokenisation-based syntactic matching is employed to check whether concepts from two source ontologies are similar according to a pre-specified threshold θ. If not, the concepts are further processed with the semantic matching algorithms based on a thesaurus (e.g. WordNet). Otherwise, similarities are determined with specific properties and relationships between their parent/child nodes. As output the algorithm returns a resulting aligned ontology (matched ontology).

```
Inputs: Ontology1, Ontology2, SemanticDistanceThreshold (θ), Wordnet
Output: alignedOntology                          //Ontology2 aligned to Ontology1
Method:
List termsFromO1                                 //list of terms in ontology1
List termsFromO2                                 //list of terms in ontology2
For each term i in termsfromO1
{
        For each term k in termsfromO2
        {
If ( i.isSyntacticallyEqual(k) & SematicDist(k,i) < θ)
{
      Align k to i
      Add k to alignedOntology
}
Else if (i.isSyntacticallyNotEqual(k) & SemanticDist (k,i)= 0)
{
      Align k to i
      Add k to alignedOntology
}
Else if (k is synonym/hypernym/hyponym/meronym of i (based on WordNet) & SematicDist(k,i) < θ)
{
      Align k to i
      Add k to alignedOntology
}
        }
}
Return alignedOntology
```

Figure 4-1 Ontology mapping algorithm

The ontology mapping allows meaningful and unambiguous sharing of concepts. Therefore, it allows the components that are represented by the ontologies to semantically interact with each other. Achieving semantic interoperability in the modelling components was the main aim of this work allowing the components to interact in a meaningful way without ambiguity.

The following section discusses the proposed approach defined in the previous sections in model architecture. This proposed model is referred to as the Ontology-based GAF (O-GAF).

4.4 Semantic similarity and Collaborative filtering score Algorithm

Once the ontology mapping has been done, the semantic preferences are calculated using the resultant aligned ontology. This section discusses on how the semantic preferences are calculated and presents the algorithm that combines the semantic preferences and the collaborative filtering scores.

4.4.1 Semantic Preferences

Another aspect of the model that requires attention is how the user preferences are modelled in the O-GAF. This section also gives particular attention to how users' preferences are determined (i.e. what will the user likely prefer in future) given his/her profile and interest. To determine what the user is likely to prefer in the future given a specific domain that he/she has shown interest in, weights are assigned to items that the user has shown interest in. For each user, all the items the user has shown interests are assigned weight and their attributes/features. Based on the weight assigned to an item their features are assigned weights. These weights represent a user's interest on that particular feature or attribute. The user's preference on items that he/she has not previously seen or showed interest on is calculated from these features/attribute weights. The weights are defined as follows:

Let $i_{m,1}$, $i_{m,2}$,..., $i_{m,Nm}$ be the N_m items which the user has shown interest in and $r_{m,1}$, $r_{m,2}$,...,$r_{m,Nm}$[1, v] be the corresponding interest given to an item (interest are measured in ratings with the scale [1,v]). Where v is the highest rating that can be given to an item by the user. We define the weight of an item, i_n, for user u_m as (Cantador *et al*., 2008):

$$W_{m,n} = \frac{r_{m,n}}{v} \in (0,1]. \qquad (4.1)$$

The preference weight for a given feature/attribute for a given user is calculated by averaging the user preference weights on the rated items the attribute appears in:

$$W_{m,f} = \frac{1}{N_m} \sum_{n:f \in features(i_n)} w_{m,n} \qquad (4.2)$$

These weights are calculated by querying the aligned ontology and are used to determine semantic preferences in the O-GAF when providing personalisation.

4.4.2 Final Score Algorithm

Once the ontology mapping is completed and semantic similarity is computed which are the semantic weights in equation 1 and 2, the score is combined with the user-based preference score to give the final results. This section presents the algorithm that is responsible for computing the final scores. Figure 4-2 shows the algorithm where it takes the semantic preferences (see section 4.2.1.1 and 4.2.1.2) and the user-based preferences from the collaborative filtering algorithm. The item score is boosted by β, if it matches with the semantic preference.

If β= 0.1, then *CombinedPref*= *User-basedPref*, in other words the standard user-based filtering. On the other hand, if β= 1.0, then only the semantic preferences is used. Finding a suitable value for β is an important task. A proper value was obtained by performing sensitivity analysis in (i.e. β=0.4) (Mobasher *et al*., 2004).

Input: SemanticPref, User-basedPref, Score
Output: preferred items and combinePref
Compare *SemanticPref*and *User-basedPref*
If *SemanticPref* == *User-basedPref*
then
CombinePref = *Score* * *β*
End If
Return *CombinePref*

Figure 4-2 Algorithm for computing the final score

4.5 The O-GAF Architecture

The previous sections presented the modeling components that are necessary to provide personalisation and further identified limitations of syntactic interaction of these components which hinders the personalisation. To overcome this issue semantic interaction of these components was realised, employing ontologies and ontology mapping to enhance the interaction. This section presents the O-GAF architecture which offers ontological knowledge representation of the modeling components and semantic interoperability, Figure 4-3 shows the O-GAF architecture.

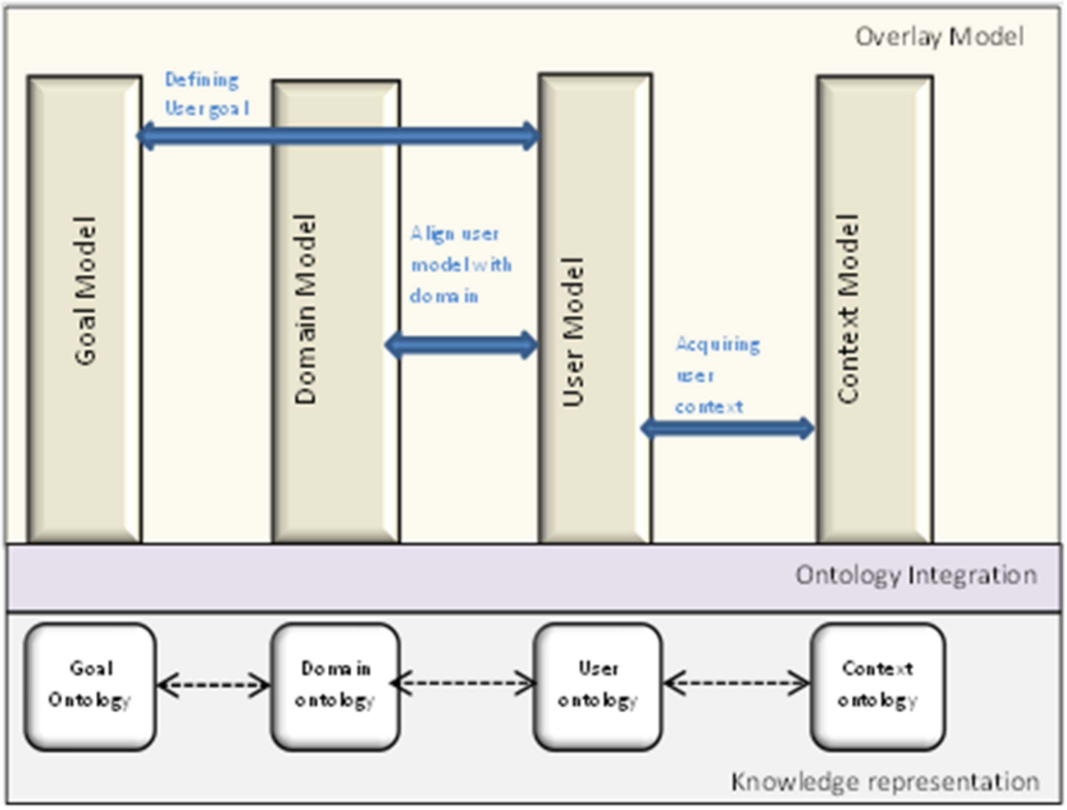


Figure 4-3 Ontology-based GAF (O-GAF) Architecture

4.5.1 O-GAF Architecture Components

This section presents the modelling components, their ontological representation and the semantic interaction which is made possible by ontology mapping. The O-GAF architecture is divided into three layers; knowledge representation layer, middle layer and overlay layer.

4.5.1.1 Knowledge representation layer

In Section 4.2.1 discussed the importance of ontology representation of the modelling components and presenting O-GAF. This section looks at the knowledge representation layer of the O-GAF architecture which is the essential part of this research work. The knowledge representation layer encapsulates modeling ontologies presented as follow:

Goal ontology; defining concept of user’s key works and aligning them to the domain ontology,

Domain ontology; defines concept and their relationships within a specific domain, see Section 4.2.1.2.

User ontology; defining concepts and their relationships in the user model,

Context ontology; defines concepts and their relationships in the context model.

The use of ontologies facilitates interoperability of the modeling components. The data structure of these models is represented by ontologies, which enables the storage of metadata. The proposed model is based on semantic data representation. Such representation is utilised by machines in process automation, integrated and reused. The O-GAF is more concerned with data integration within the system. An integration mechanism aims to resolve conflicts due to component heterogeneity.

4.5.1.2 Ontology Integration layer

The ontology integration layer is composed of ontology integration tools. In order for the overlay components to work with ontology models it is necessary to use some framework as a gateway. This work has adopted the Jena Framework to act as a gateway between the ontology representation layer and the overlay layer. This layer allows the modelling components to interact with the ontologies and acquire the semantic reasoning they require.

4.5.1.3 Overlay layer

The overlay model is composed of the four optional modeling components; Goal Model, User Model, Domain Model and Context Model. The interaction of the modelling components is defined in the context of a personalised recommender system.

Goal Model: The User states the goal thus formulating a new query for which inferences over the user preferences are made. This step can be considered as stating a goal or choosing a particular concept (set of concepts) to study/visit in an AHS. The goal can be interpreted and aligned with DM (availability of concepts, concept structures and hierarchies, etc.) and

User Model: defines user specific information (considering user competencies, preferences, experience, interests, etc.). The user model is significantly important in personalised systems, employing collaborative filtering, semantic-based approaches or combination of the two approaches.

The Domain Model (DM) is defined by the knowledge structure of the domain, representing keywords and terms together with the relationships which can be used to facilitate fast and reliable information retrieval and filtering of the concerned item space, where the resources are defined by RM. On the other hand DM may represent the feature space in case of content-based recommendations where content is again retrieved from RM. It happens more often in the recommender systems that the domain is represented by an ontology (Cantador *et al.*, 2008) in order to facilitate more elaborate reasoning over the items relationships and present more accurate recommendations.

The Context Model: defines user and usage context properties such as IP address, time, other activities of the user, etc. Modelling Context gives a possibility to consider context-aware recommendations and adaptation, both from the usage and the user point of view.

4.6 Summary

The chapter has presented the O-GAF as a solution to the research issues raised in chapter one of this study. In addressing those research issues some design requirements have been drawn which were met by the O-GAF. The O-GAF was designed and the model architecture combining the solutions presented in Figure 4-3. In the following chapter the O-GAF is implemented as a proof of concept and some experiments are performed to evaluate the solution approach provided by O-GAF.

– CHAPTER FIVE –

Model Prototyping and Evaluation

5.0 Introduction

In achieving the goal of enhancing personalisation in web-based systems, an ontology-based approach was employed. Chapter four presented the ontology-based overlay model architecture for personalisation in web-based systems, providing semantic interoperability for optimum component interaction. This solution approach is referred to as the ontology-based Generic Adaptation Framework (O-GAF) which is an extension of the GAF.

This chapter presents the implementation of O-GAF. O-GAF was applied in a personalised movie recommender system as proof of concept of this work. The O-GAF demands support from more than one modelling component (i.e. the modelling components are optional depending on the application needs). Therefore, for the purpose of this work and suitability of the application needs two ontology models were developed (i.e. user ontology model and domain ontology model). We present the design and implementation of the prototyped O-GAF recommender system.

The organisation of this chapter is as follows, Section 5.1 presents the prototype design of the ontology-based recommender system using UML modelling and Section 5.2 presents the implementation details and the environment used for the implementation. In Section 5.3

experiments and evaluation of the prototype are discussed as a means to prove the effectiveness of our proposed solution.

5.1 Prototype Design

This section presents the prototype design of the proposed model in the movie recommendation domain. The prototype design is presented in UML: the use case diagram shows a list of steps, typically defining interactions between roles and the system; the sequence diagram shows how processes operate with one another and in what order; and the activity diagram shows the overall flow of control in the system.

5.1.1 Use Case Modelling

The use case modelling of the prototype design of the semantically enhanced recommender system is shown in Figure 5-1, the system has two actors: the user and the application domain. The system has one main use case: recommendation which includes and uses sub-use cases.

The recommendation use case generates recommendations for the users, in order to generate these recommendations it uses two sub-use cases, computing user similarity and computing semantic similarity. For computing semantic similarity the ontology modeller task is invoked. The ontology modeller task manages all the ontologies in the system. The user ontology is build using the user preferences and the domain ontology using the domain data, this action is performed offline. Ontology mapping is performed between ontologies and a resultant ontology is returned. Therefore, the semantic similarity measure returns the semantic results to the recommendation use case.

User similarity is computed and returns the results to the recommendation use case. For the user similarity to be computed two tasks are included: the user preference task which supply the similarity measure with user's ratings of items in the application domain; the get neighbourhood size task which get the number of neighbours to be considered.

Actor: User

The user actor is responsible for initiating the process of recommendation by invoking the recommendation use case requesting for recommendation. The request passed by the user contains the userID and number of items to be recommended. The user is also responsible for providing user preferences to the systems; this interaction may not be at runtime.

Actor: Application Domain

The application domain is an external structure that stores domain information. This component supplies the system with domain knowledge for building a domain model.

Recommendation

The recommendation is modelled as the decision making component. It is invoked by a request from the user. The recommendation is responsible for making recommendation and returns the recommended items to the user. To make the recommendation it uses the semantic knowledge from the ontology modeller where semantic similarity is computed. Before the recommendation use case makes the final recommendation user similarity is computed comparing user's preferences with similar users. The preferred items and their score are returned. The preference score for user similarity and semantic similarly are combined and the final score is computed. The items with the highest scores are returned to the user as recommended items.

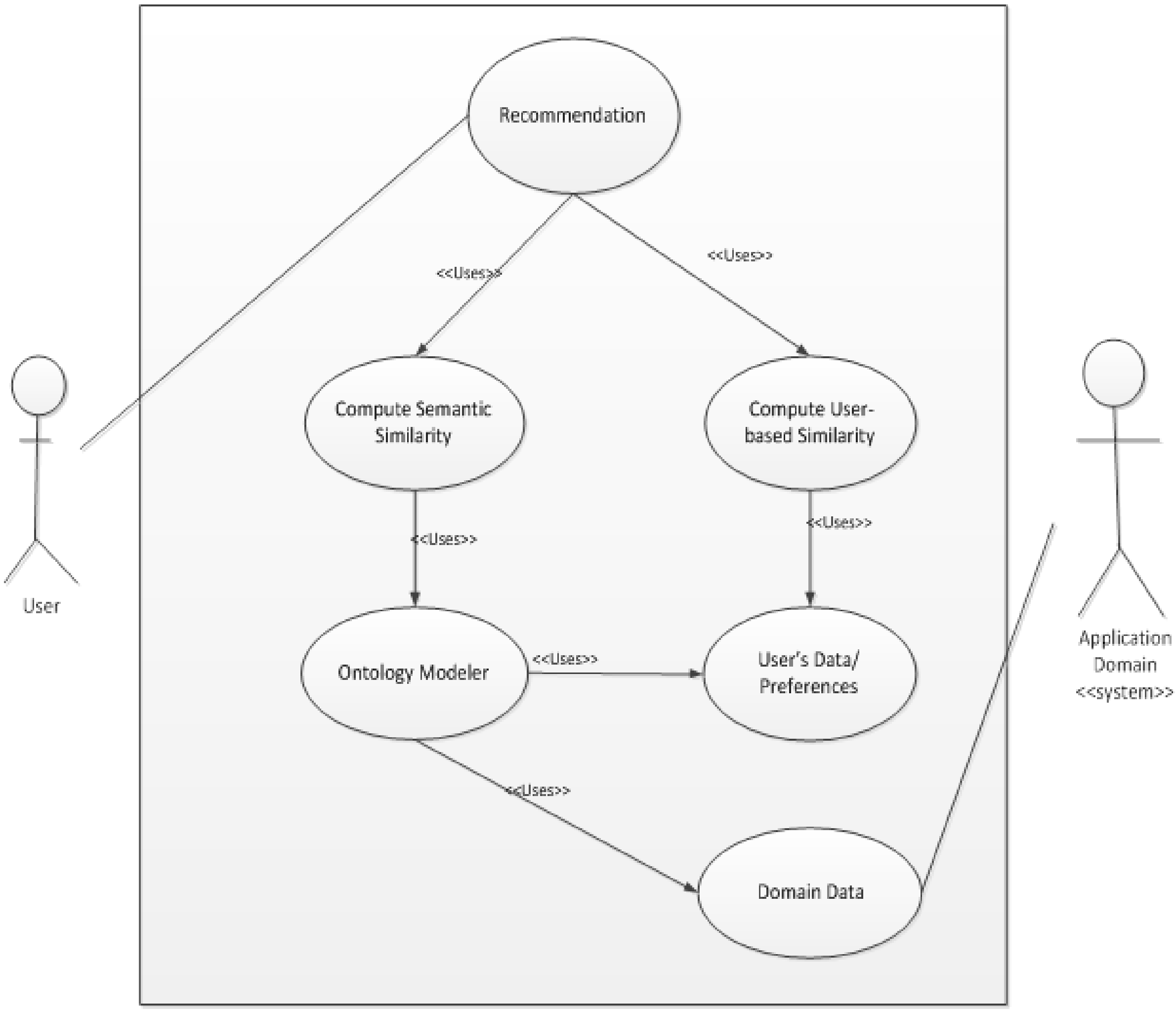


Figure 5-1 use case diagram

5.1.2 Sequence Diagram Modelling

The Sequence diagram in Figure 5-2 shows the sequence of the flow of messages among the components and actors of the prototype. The flow of messages is initiated by the user sending a recommendation request to the recommendation; the message contains request details (i.e. userID and number of items to be recommended). The recommender sends a request with the userID to the similarity measure to find nearest neighbourhood for the active user. For the similarity measure to perform this action it get the user preferences from the data models. The data model replay with required data model. After computing the user similarity a replay message is sent back to the recommendation with the active user's neighbourhood and their preferences. The recommender sends another request to the ontology manager requesting semantic knowledge of the domain with relation to the active user. The ontology manager forwards its message to the data model acquiring the user model and domain model and ontology models. The necessary models are returned to the ontology manager, where ontology mapping and semantic similarity are performed returning the results to the recommendation. The recommendation makes use of this knowledge and returns the results to the user which is a list of recommended items.

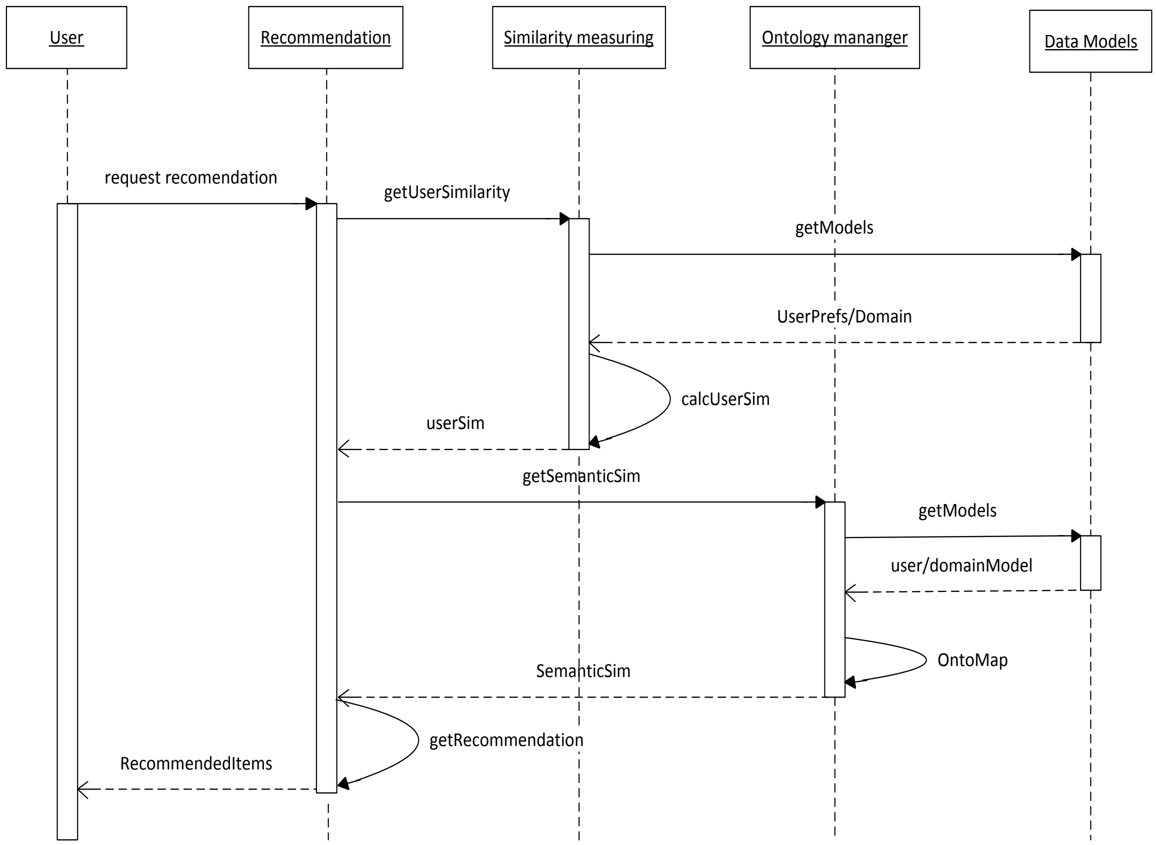


Figure 5-2 sequence diagram

5.1.3 Activity Diagram

In this section we present graphical representations of workflows of stepwise activities and actions of the system components. Figure 5-3 shows this workflow in an activity diagram of our prototyped system. The recommendation process starts with the user requesting recommendations. The recommendation invokes two methods the user similarity measure and the ontology manager. The similarity measure in receipt of the recommender's request to get user similarity, it invokes a similarity measure which requires a user model with all the user's preferences. The data model loads the required data model to the user similarity measure algorithm computes the nearest neighbours to the active user. The results are sent back to the recommender. The second method invoked by the recommendation which is the ontology manager builds the ontology models, but first request for the data models (i.e. user model and domain model).

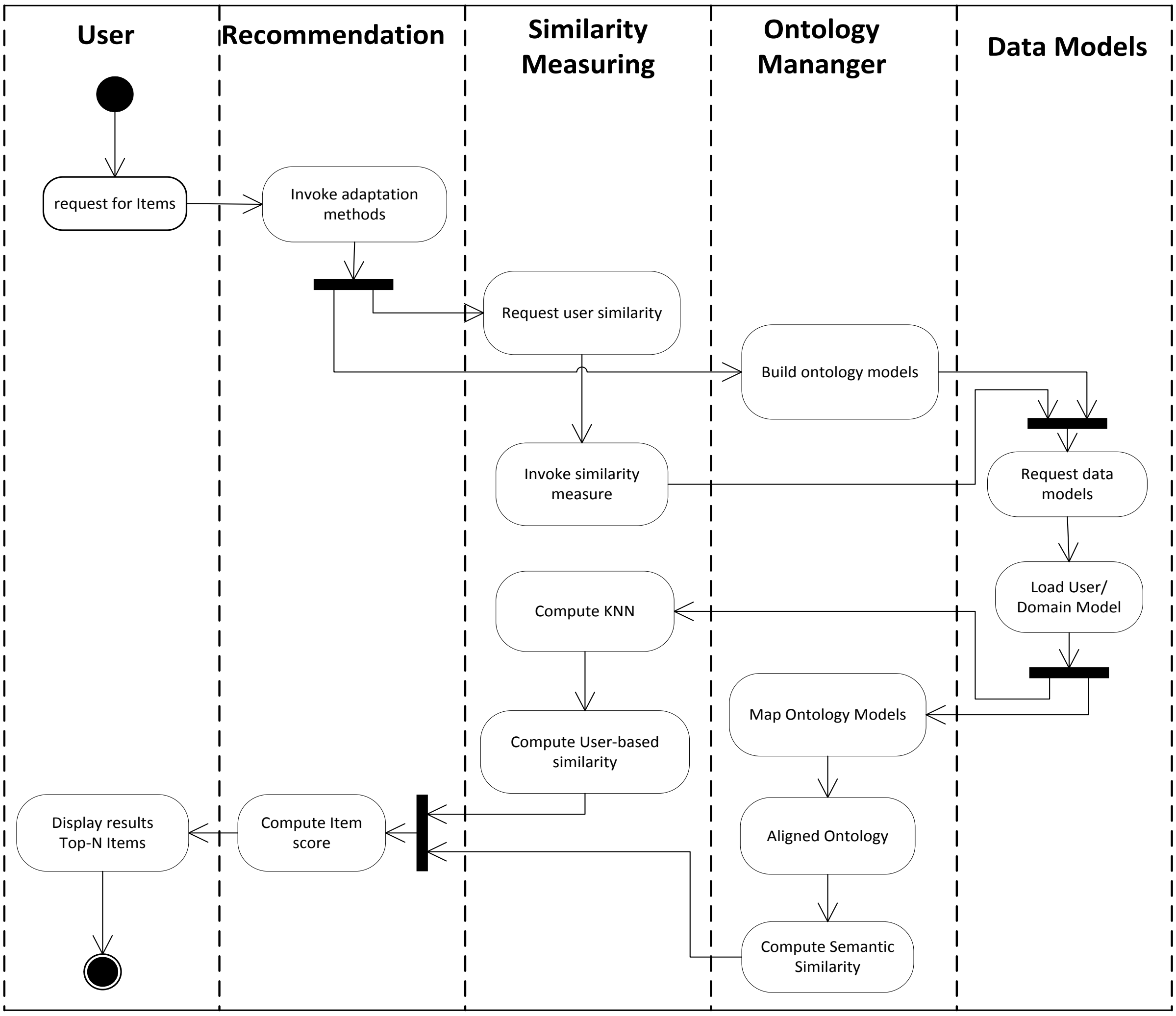


Figure 5-3 Activity diagram

The required ontologies are mapped using an ontology mapping algorithm. After the mapping the resultant ontology is used to compute semantic similarity and the results are returned. The recommendation using both the information from the similarity measure and the ontology manager computes the final score of the items to be recommended and the items with the highest score are recommended to the user in a list from highest to lowest.

5.2 Implementation

In the previous section we presented the design of the ontology-based recommender prototype, and therefore this section continues to presents the implementation and the environment of the prototyped ontology-based recommender system. The system prototype was implemented on Microsoft windows 7 machine, with 4GB of Ram and Intel(R) Core(TM) i5 CPU at 2.27 GHz processor. The tools for the development of the system were JDK 1.7, and Apache Test Mahout Framework. The tools for the development of the ontologies were Protégé 4.1 (an ontology development and editing tool) supporting RDF/OWL ontology languages, Jena Framework and SPARQL (**S**PARQL Protocol and RDF Query Language) an RDF query language allowed the system to work with ontology models from the java classes in a transparent way. The details of the prototype implementation are given in the following subsection.

The user interface of the recommender prototype is shown in Figure 5-4, where the user enters their user ID and number of items to be recommended. The user can choose which method to use for recommendation between the ontology-based methods or the non-ontology based one. Finally the user can execute to get the recommendations. Figure 5-5 shows the screen shot interface for evaluating the recommender system. The dataset for evaluation was partitioned into four groups for validity of the results (see Section 5.3.1). To evaluate the recommender system one group was selected, parameters and the recommendation method specified. The evaluation results are displayed at the bottom of the window.

Figure 5-4 Screen shot – User Interface of Movie recommender system

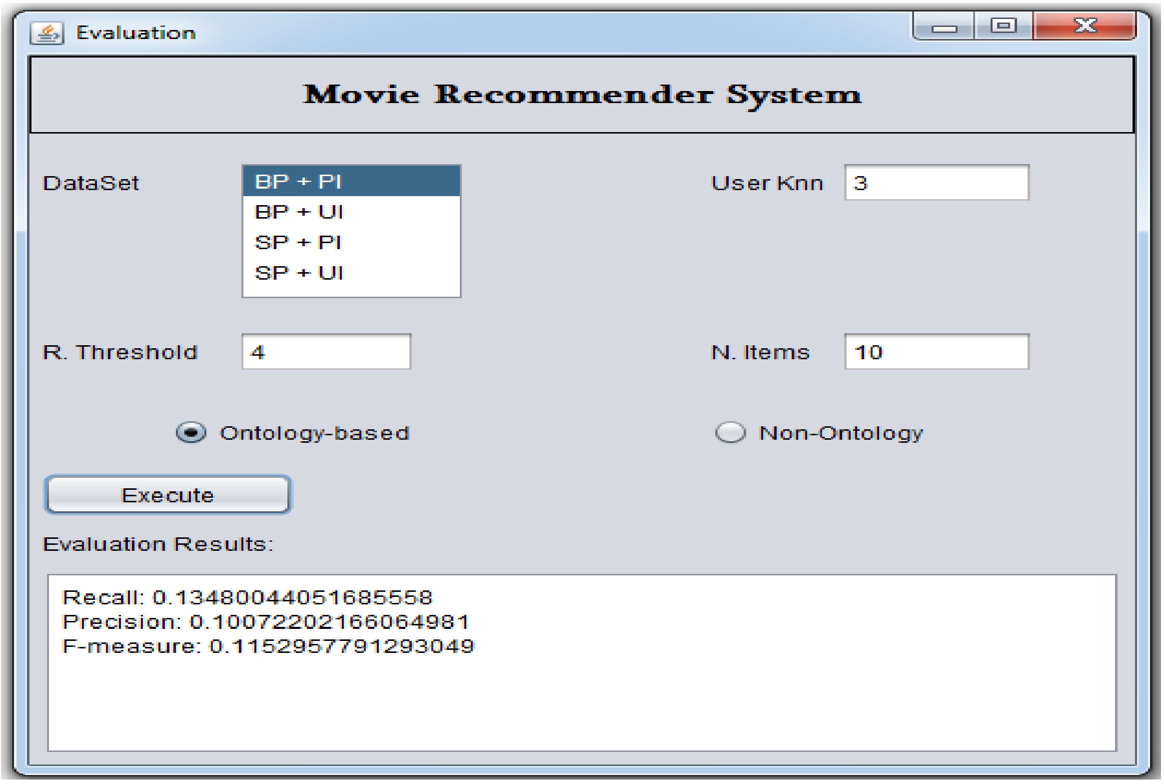


Figure 5-5 Screen shot – Evaluation Interface of Movie recommender system

5.2.1 Ontology modelling

As stated earlier, due to the suitability and application needs only two of the ontology models were developed. These ontologies where built corresponding with two modelling components which are domain model (domain ontology) and user model (user ontology). The user ontology model was developed using the protégé 4.1 and the domain ontology was built from available ontologies, available on the online ontology repositories.

Figure 5-6 and shows the user ontology, defining the concepts and their relationship in the user model. User information is divided in domain-dependent, such as user's interests, stereotypes and ratings; and domain-independent, such as user personal and demographic data such as user's occupation and gender. The knowledge in the user ontology is used to infer new knowledge about the user. As mentioned in the design chapter Section 4.2.1.1 the user interest on new items can be inferred by calculating the preference weight of feature that appear in items the user has shown interest in.

Figure 5-7 shows the domain ontology which defines the movie domain. The movie ontology is being adapted from available online ontology repository (Bouza, 2010) and was edited using protégé 4.1 for the suitability of our proof of concept. The movie ontology defines the concept of a movie and the relationship with other related concepts as shown in Figure 5-7.

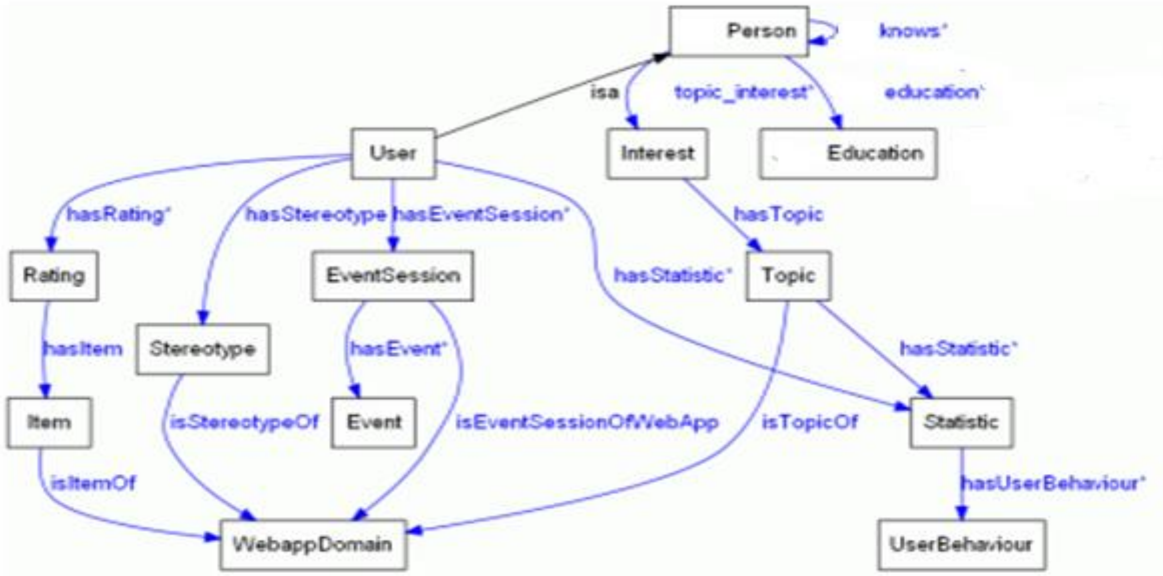


Figure 5-6 user ontology- main concept and relationship of the user model

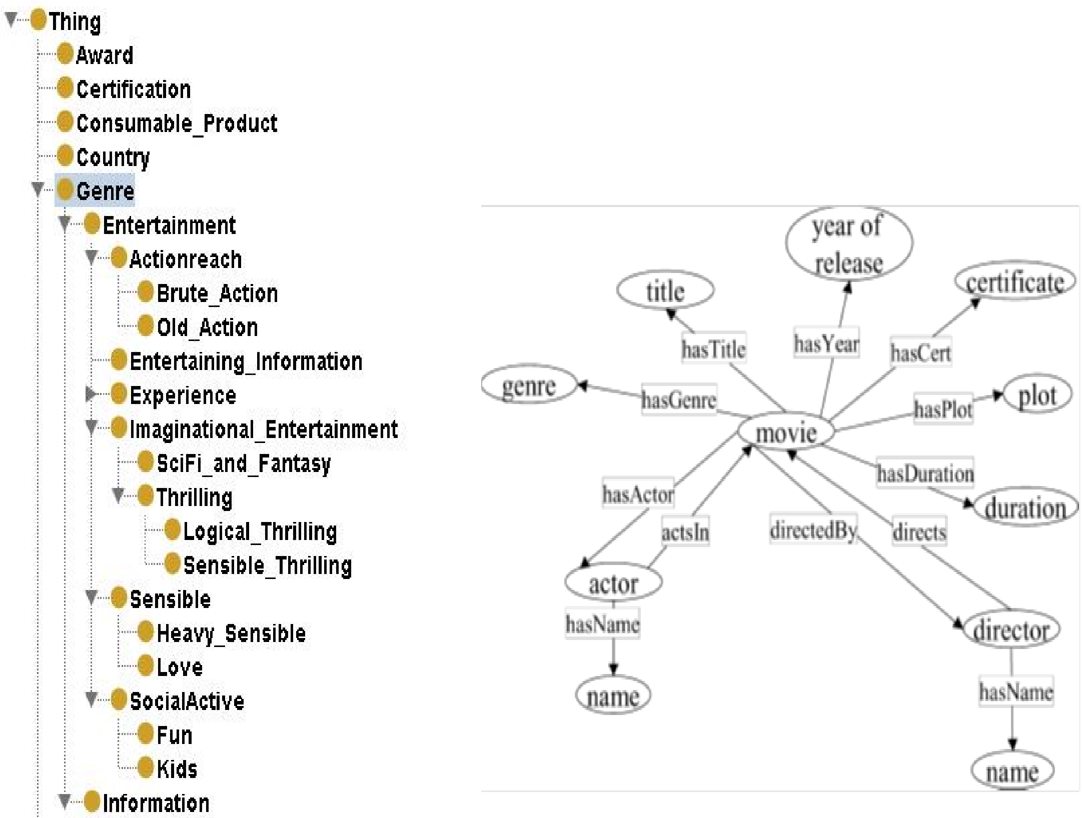


Figure 5-7 Graphical view of the movie domain ontology

5.2.2 Ontology mapping algorithm

Figure 5-8 presents the implementation of ontology mapping algorithm used to map the user ontology and the domain ontology. The algorithm takes two URL as input which are domain ontology and user ontology. The algorithm uses three different mapping methods to create alignments and computes the final mapping: **EditDistNameAlignment** uses an editing (or Levenstein) distance between (downcased) entity names. It builds a matrix of distance and chooses the alignment from the distance; **SubsDistNameAlignment** computes a substring distance on the (downcased) entity name; **NameEqAlignment** simply compares the equality of class and property names (once downcased) and aligns those objects with the same name. The output of the previous align method is passed as an argument to the next method. The finally obtained alignment is generated in RDF form resulting in a resulted ontology.

5.2.3 Collaborative Filtering Algorithm

Collaborative filtering algorithms rely on past user behaviour regardless of the domain knowledge. Section 3.0 stated the two primary approaches to collaborative filtering: the neighbourhood approach and the latent factor approach. This section presents the collaborative filtering algorithm that was adopted in the design and implementation of the personalised recommender system. The neighbourhood approach was followed and the Pearson correlation coefficient similarity measure was implemented to calculate user similarity. A user-based characteristic of a recommender algorithm was defined which is based upon some notion of similarities between users. Figure 5-9 shows the algorithm implementation.

```
…….
   URI uri1 = new URI("ontofiles/userOntology.owl");
                 URI uri2 = new URI("ontofiles/domainOntology.owl");
                 Parameters p = new BasicParameters();
             AlignmentProcess A1 = new SubsDistNameAlignment();
A1.init( uri1, uri2 );
             AlignmentProcess A2 = new EditDistNameAlignment ();
             A2.init( uri1, uri2 );
             AlignmentProcess A3 = new NameEqAlignment();
             A3.init( uri1, uri2 );
             A1.align((Alignment)null,p); A1.cut("prop", .5);
             A2.align((Alignment)null,p); A3.align(A2,p);
                 Evaluator E = new PRecEvaluator(A1, A3);
             E.eval(p);
             PrintWriterpw = new PrintWriter (new BufferedWriter(
                                  new OutputStreamWriter(System.out, "UTF-
       8" )), true);
             AlignmentVisitor V = new SWRLRendererVisitor(pw );
             if ( ((PRecEvaluator)E).getPrecision() > .6 ) A3.render(V);
       ……
```

Figure 5-8 Ontology Mapping Algorithm

```
……
while (true) {
int compare = xIndex<yIndex ? -1 :xIndex>yIndex ? 1 : 0;
if (compare == 0) {
double x = xPrefs.getValue(xPrefIndex);
double y = yPrefs.getValue(yPrefIndex);
sumXY += x * y;sumX += x;sumX2 += x * x;sumY += y;sumY2 += y * y;
double diff = x - y;
        sumXYdiff2 += diff * diff;count++;}
if (compare <= 0) {
if (++xPrefIndex == xLength) {
break;
……
double result;
if (centerData) {
doublecenteredSumXY = sumXY - meanY * sumX;
double centeredSumX2 = sumX2 - meanX * sumX;
double centeredSumY2 = sumY2 - meanY * sumY;
result = computeResult(count, centeredSumXY, centeredSumX2,
centeredSumY2, sumXYdiff2);
    } else {
result = computeResult(count, sumXY, sumX2, sumY2, sumXYdiff2);}
if (similarityTransform != null) {
result = similarityTransform.transformSimilarity(itemID1, itemID2,
result);}
if (!Double.isNaN(result)) {
result = normalizeWeightResult(result, count, cachedNumUsers);}
return result;
  }
….
```

Figure 5-9 Pearson correlation coefficient algorithm.

5.3 Comparison of O-GAF against the GAF

This section presents how the experiments to evaluate the O-GAF were conducted and the results that came out of the experiments. The overall goal of the experimentation was to investigate whether the introduction of ontologies and ontology mapping in GAF enhances personalisation. To achieve the goal the O-GAF was benchmarked against the classical GAF. The metrics that were used for the evaluation were Recall, Precision, and the F-measure. These metrics are defined as shown in equations 1, 2 and 3. The O-GAF was also compared to the GAF with respect to the response time.

$$\text{precision} = \frac{number\ of\ relevant\ items\ \cap\ number\ of\ retrieved\ items}{number\ of\ retrieved\ items} \qquad [5.1]$$

$$\text{recall} = \frac{number\ of\ relevant\ items\ \cap\ number\ of\ retrieved\ items}{number\ of\ relevant\ items} \qquad [5.2]$$

$$\text{F} - \text{measure} = (2 \cdot recall \cdot precision)/(recall + precision) \qquad [5.3]$$

In order to do fair comparisons of the O-GAF against the GAF using the afore-discussed metrics, some key decisions had to be made and key parameters determined. The decision included the choice of symmetric measure, the optimal neighbourhood size (k) for collaborative filtering in order to find optimal size for a given dataset (see Section 5.3.1) some experiments were to be carried out.

After experimentally determining the optimal neighbourhood size for collaborative filtering, experiments to compare the effect of increasing the number of recommendations on Recall on the solutions was conducted. This was followed with the experiments to compare the effect of increasing recall on the precision of the solution. These experiments are discussed in Section 5.3.3. To determine whether the O-GAF significantly outperforms the classical GAF a randomised block design experiment was conducted in Section 5.3.4.

The use of ontologies is known to be computationally intensive. Hence, the proposed solution is expected to have poor response time. In order to verify this, an experiment, presented in Section 5.3.5, was carried out.

To evaluate the performance of an ontology-based approach classification accuracy metrics were used. Classification accuracy metrics are suitable for domains with binary rating. The experiments that were carried out aimed to investigate the accuracy of recommendations in two models: one is the model presented in this work ontology-based model; and the other non-ontology model. The non-ontology-based recommender system was implemented based on the Apache Taste Mahout Framework (http://mahout.apache.org/), (Owen *et al*., 2011). Accuracy metrics that were chosen are f-measure, recall and precision.

5.3.1 Experimental data set

For testing the recommender system a movie dataset with ratings from the MovieLens recommender was used (http://www.grouplens.org/node/12). The dataset contains 100k ratings for 1 682 movies by 943 users on a 1-5 rating scale. The MovieLens is based on the

Internet Movie Database (IMDb) containing catalogues of every pertinent detail about a movie. The combinations of the two data sources were explored. Specifically, the IMDb information was used to define a domain ontology describing the fundamental concepts involved in IMDb, including classes such as movies, actors, directors, genres, languages, countries, keywords, etc., and relations among them.

According to literature, given a dataset it is important to take note of the following (Cremonesi *et al*., 2008): i) the quality of the recommendation may depend on the length of the user profile, ii) the quality of the recommendation may be different according on the popularity of the items rated by a user, i.e., the quality of the recommendation may depend on the fact that a user prefers popular items against unpopular ones. Taking into consideration these two facts, the dataset was partitioned into four groups as follows (Cremonesi *et al*., 2008): users were sorted according to the length of their profiles (i.e., the number of ratings) which resulted in two groups; items were sorted according to their popularity which also resulted in two groups (i.e. popular and unpopular items)

In the MovieLens dataset the users were split into two groups sorting users with length of profile:

[1…99] rating small profile (SP); [100… ∞] rating big profile (BP).

Similarly, in order to find the most popular items; items were sorted according to their popularity (i.e. number of ratings):

[0… 99] unpopular items (UI); [100… 495] popular items (PI)

After the dataset was partitioned four groups were obtained: (BP+PI); (BP+UI); (SP+PI); (SP+UI). This grouping helps to identify the behaviour of the model according to the length of the user profile and to the popularity of the items of the recommendation.

5.3.2 Finding the optimal neighbourhood size

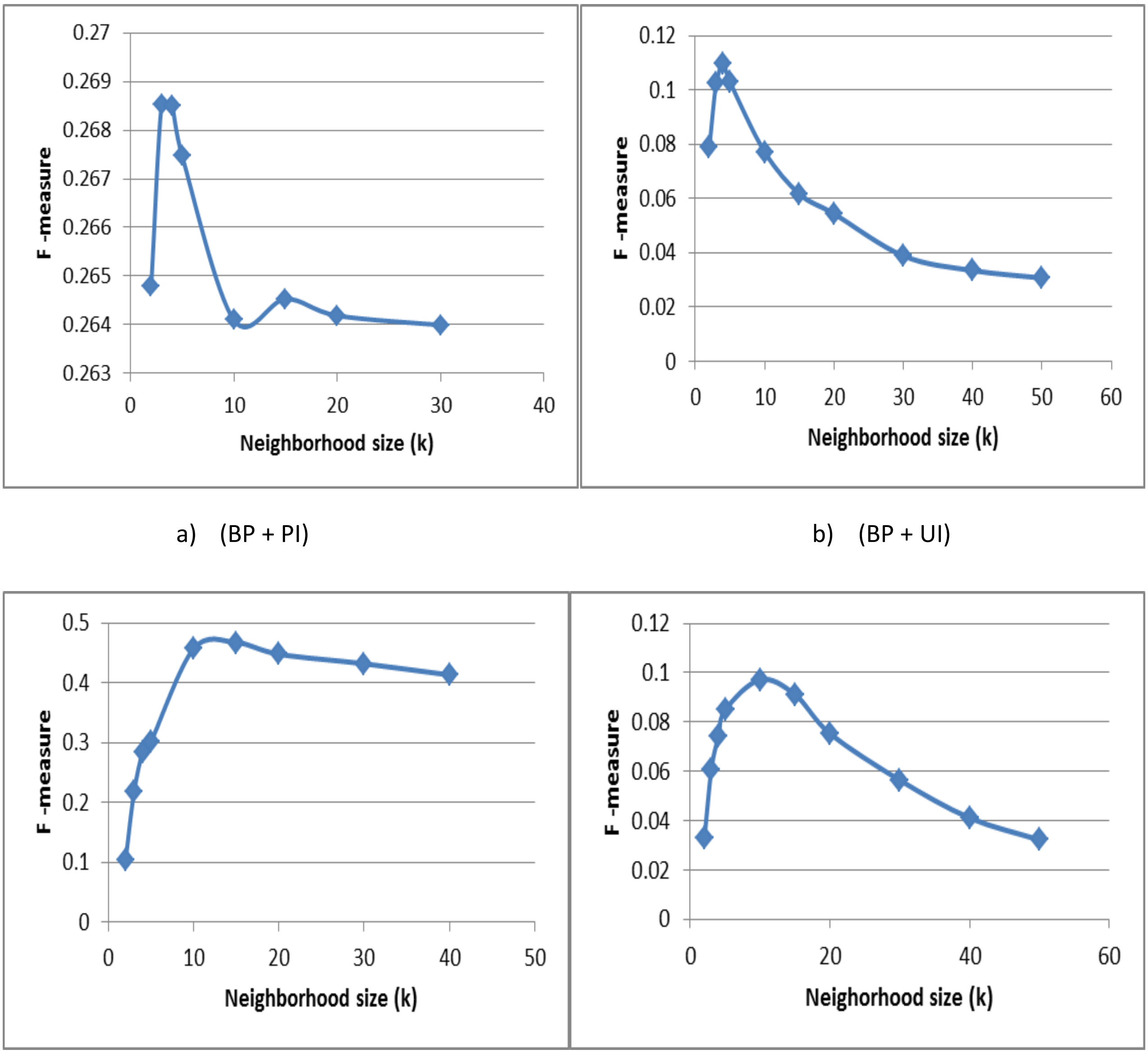
The size of the neighbourhood has a significant impact on the recommendation quality (Cremonesi et al, 2008). In order to determine the optimal neighbourhood size some experiments were carried out.

5.3.2.1 Experimental Design

To determine the optimal size of k for each of the datasets, the following experiments were performed: the number of recommendations was held constant at top-N =10, the neighbourhood size k being varied and for each point the F-measure (F1) was computed. The neighbourhood size (k) that maximises the F-measure was taken to be the optimal k for a given dataset.

5.3.2.2 Experimental Results

This section presents the results of determining the optimal k in the recommendation process using F1, in order to find the optimal k given the four groups of dataset presented in (section 5.3.1.). Figure 5-10 shows the results of all four datasets partitions where a) big profiles plus popular items (BP + PI) having [100 - ∞] ratings and [100 - 495] rated items. In the first group the optimal number of k is at 3. Figure 5-10, b) presents the second case where the dataset has big profiles plus unpopular items (BP + UI) having [100 - ∞] ratings and [0 - 99] rated items. The optimal k is at 4 where f-measure is optimal at 0.11. Figure 5-10, c) third case where the dataset has small profiles plus popular items (SP + PI) having [1 - 99] ratings and [100 - 495] rated items and figure 5-10, d) where the dataset has small profile plus unpopular items (SP + UI) having [1 -99] ratings and [0 - 99] rated items the optimal neighbourhood sizes is found at 10.



a) (BP + PI)

b) (BP + UI)

b) (SP + PI)

d) (SP + UI)

Figure 5-10 optimisation of k, using f-measure in a. (BP + PI), b. (BP + UI), c. (SP + PI), and d. (SP + UI)

5.3.3 Recall and Precision

In this section the experimental design and evaluation results of the personalised ontology-based recommender in comparison with the non-ontology-based recommender is presented.

The objective of this experiment is to evaluate the top-N recommendation recommended by the two models.

5.3.3.1 Experimental Design

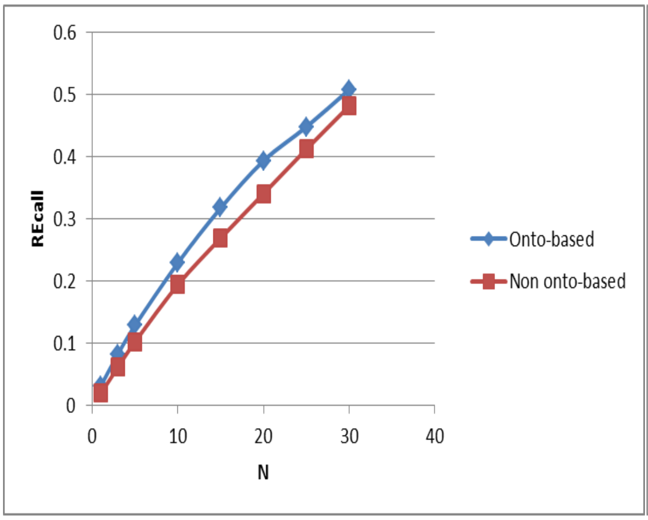
In recommender systems, precision and recall are defined respectively as: precision: the proportion of the top-N recommended items that are relevant to user; recall: the proportion of the relevant items that are recommended to user.

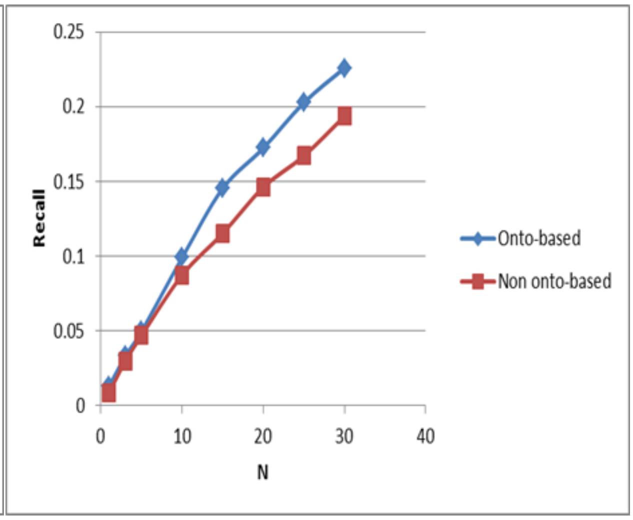
This experiment was aimed at investigating how recall behaves as the number of recommendation increases for each of the models being compared. The rating scale of [1-5] 1 being least relevant and 5 being highly relevant were used, therefore the relevant threshold was set at 4. The range of Top-N was {1, 3, 5, 10, 15, 20, 25, 30} and precision at N as well as recall at N were computed.

5.3.3.1 Experimental Results

This section reports the performance of the two recommender systems on the four cases of the dataset presented in (section 5.3.1). For each dataset group two set of experiments were performed. Recall at N results are reported followed by precision as a function of recall results.

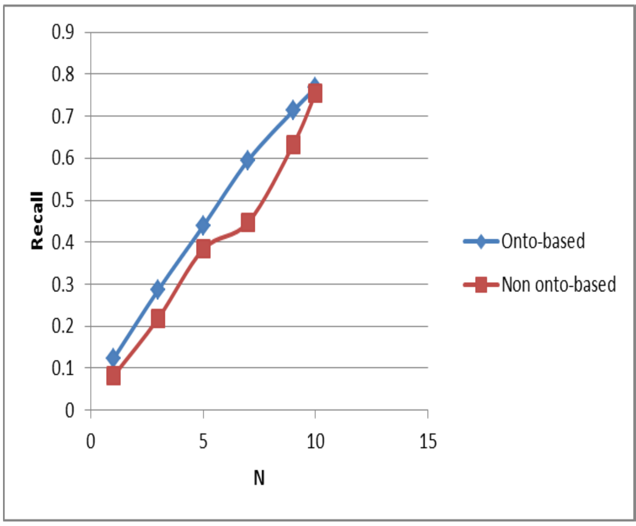
The results of all dataset partitions are shown in Figure 5-11. Where Figure 5-11, a) (BP + HR) it is apparent that there is not much difference between the ontology-based and the non-ontology-base recommender at the beginning where N = 1,3 and 5, but as N increases a slight difference can be observed, where the ontology-based recommender performs better than the non-ontology-based. For instance, recall of the ontology-based approach at N= 20 is about 0.4, i.e., the model has a probability of 0.4 to place relevant items on the top-20, while the non-ontology approach has about 33%. In Figure 5-11, b) reports on the results of the second case of the dataset (BP+UP). Similar to the previous results as the number of recommendations was increased recall increases and the ontology-based recommender outperforms the non-ontology recommender. Figure 5-11, c) and d) exhibited the same behaviour. In all the recall at N experiments it was observed initially the two recommender models performed more or less the same but as the number of recommendation increased the ontology-based approach outperforms the non-ontology approach.

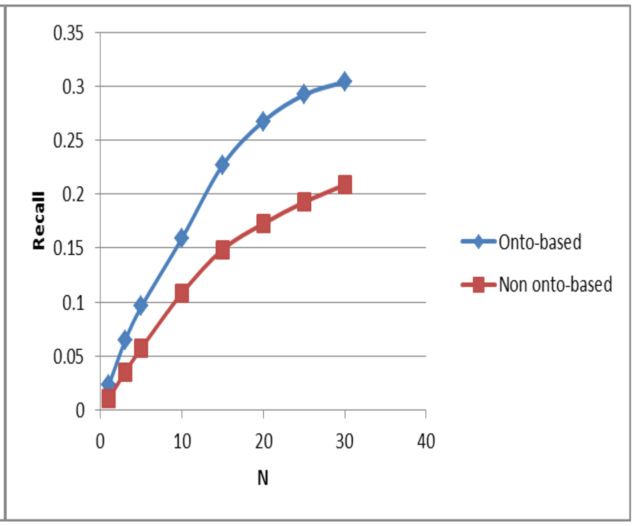




a) (BP + PI)

b) (BP + UI)



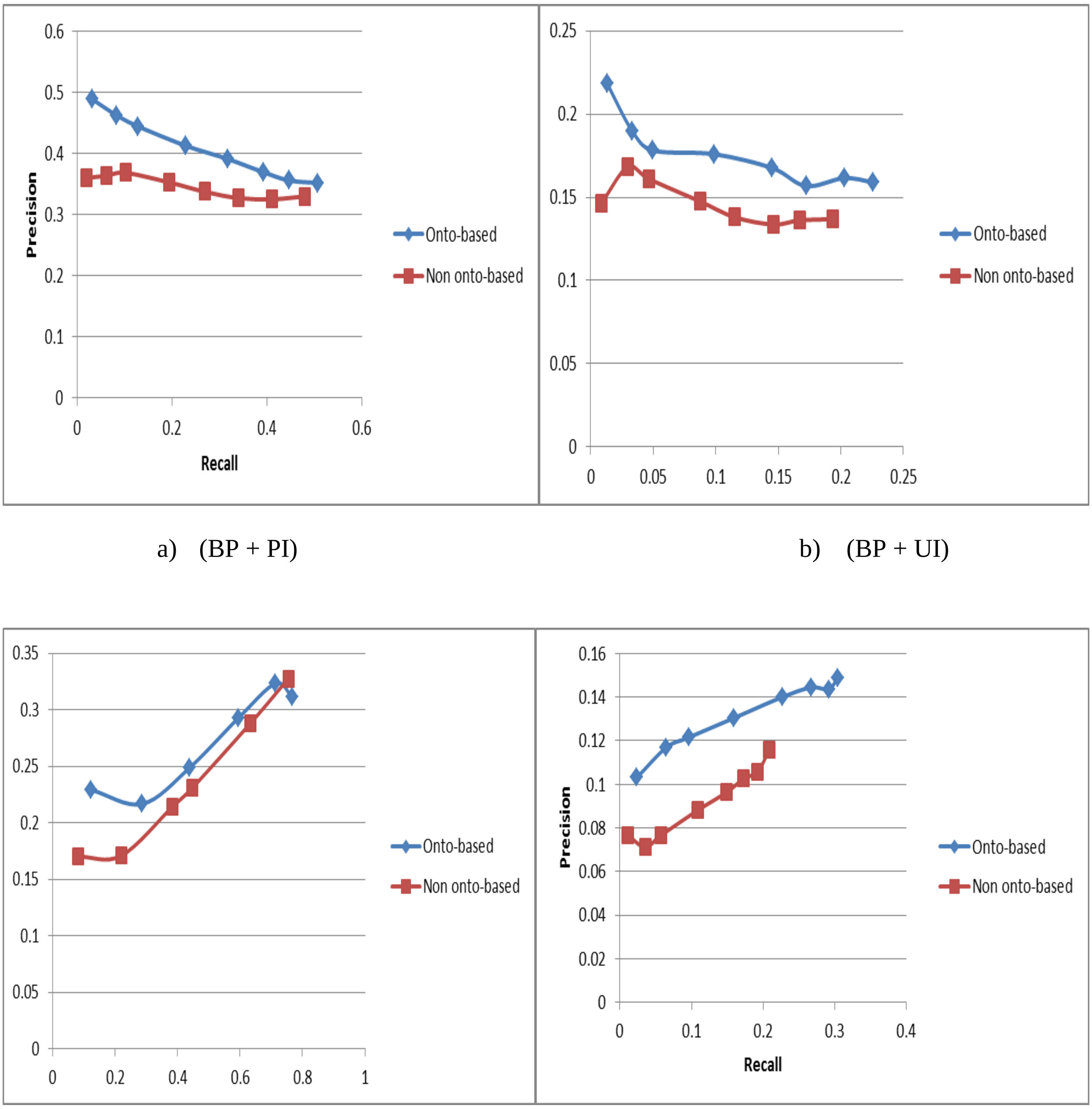


c) (SP + PI)

d) (SP + UI)

Figure 5-11 ontology-based and non-ontology-based model, precision and recall, in: a. (BP+PI), b. (BP + UP), c. (SP + PI), and d. (SP + UI)

When observing precision vs. recall, Figure 5-12, a) confirms that the ontology-based recommender outperforms the non-ontology recommender in terms of precision metrics. Each line represents the precision of the recommender system at a given recall. As recall increases precision decreases as the two metrics tread-off each other. For example, when recall is about 0.2, precision of the ontology-based approach is about 0.43. A behaviour can be observed in the recall precision graphs, as recall increases the two models are converging towards the same value. This behaviour is due to the fact that as recall is being increased by increasing the number of recommendations, all the relevant items being retrieved until the two models will reach the same value where all relevant items are exhausted. In figure 5-12, b) the second dataset, were precision and recall graphs for the two recommender models are compared. It can be observed that the ontology-based recommender performs better than the non-ontology based recommender exhibiting the same behaviour as a). In the last two cases Figure 5-12, c) and d) where the profiles are small the recall vs. precision graphs exhibits an unexpected behaviour. The reason for this behaviour is the small number of profiles the users, which affect the relevant items for each user when recommending (de Gemmis *et al*., 2009), and therefore, for good recommendation to be made the system requires a large amount of data from the user.



a) (BP + PI)

b) (BP + UI)

c) (SP + PI)

d) (SP + UI)

Figure 5-12 ontology-based and non-ontology-based model, precision and recall, in: a. (BP + PI), b. (BP + UI), c (SP + PI), and d. (SP + UI)

5.3.4 Comparison of the O-GAF to the Classical GAF

This section presents the statistical significance test, which was used to compare the performance of the ontology-based recommender system against that of the classical non-ontology-based recommender system. The Randomised Block Design statistical experimental design was used.

5.3.4.1 Experimental Design

The aim of the randomised block design experiment was to control factors that affect the quality of the recommendation, while they are not of primary relevance to what is being investigated in this research work. These factors where identified in Section 5.3.1 as the length of the users profile and the popularity of rated items. The blocking factors represent the combination of profile length and item popularity, which was grouped as follows: (BP+PI), (BP+UI), (SP+PI), and (SP+UI). Blocking these factors in the randomised block experimental design removes the effect of the factors on the comparison of the treatment effects. Because this design reduces variability and potential confounding, it produces a better estimate of treatment effects. The participants within each block are randomly assigned to treatment conditions (i.e. the two systems being compared). The optimal neighbourhood sizes for group obtained from the experiments carried out in Section 5.3.2 were used. To get a sizable amount of data to be used in the comparison the optimal neighbourhood size (k) for each group and two values in its integer neighbourhood were used in the experimentation. Thus the optimal neighbourhood sizes used were as follows: (BP+PI) k=[3,5,5], (BP+UI) k=[4,5,6], (SP+PI) k=[10,11,12], and (SP+UI) k=[10,11,12]. The number of recommendation was held constant at N= 10 in all cases.

Hypotheses:

H_0: the f-measure score of O-GAF = f-measure score of classical GAF,

H_1: the f-measure score of O-GAF ≠ f-measure score of classical GAF.

Decision Criteria

The null hypothesis (H_0) is rejected in favour of the alternative hypothesis (H_1) with 95% confidence (0.05 significance level) if the p-value $\leq$ 0.05.

5.3.4.2 Statistical analysis of Results

This section presents the results of the final experiment that was performed to test whether ontology-based approach is significantly better than the non-ontology-based approach to recommendation, using a metric that combines recall and precision, the F-measure. When the analysis was run in SPSS, the following results shown in Table 5-1 and 5-2 were obtained

Table 5-1 present the ANOVA test of between-subject effects, where the output, F = 50.292 with1 and 19 degrees of freedom. The p-value (Sig.) $\approx$ 0.000. Since the p-value is less than 0.05 we reject the H_0 and conclude that there is a statistically significant difference between the mean F-measures of the ontology-based and the non-ontology based approach. The result in Table 5.1 also shows the blocks had a significant effect on the F-measures, since the p-value for the blocks is less than 0.05. This justify why it was necessary for the dataset to be partitioned such that the length of the users profile and the popularity of rated items will not have an effect on the comparison of the O-GAF to GAF.

Table 5-1: Estimates, Tests of Between-Subjects Effects (Dependent Variable: F-measure)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	.540[a]	4	.135	755.839	.000	.994
Intercept	1.258	1	1.258	7047.408	.000	.997
Treatment	.009	1	.009	50.292	.000	.726
Block	.531	3	.177	991.022	.000	.994
Error	.003	19	.000			
Total	1.801	24				
Corrected Total	.543	23				

Table 5-2 presents the ANOVA multiple comparison analysis which shows which treatments are similar and which are not. This experiment aimed to compare two treatments. Therefore, there are only two combination of comparison: the ontology-based approach and the non-ontology-based approach. The significant test results show that the two approaches (treatment) are statistically significant at the predetermined level of significant ($p < 0.05$). Figure 5-13 shows the results illustrated in an error graph.

Table 5-2: Pairwise Comparisons (Dependent Variable: F-measure)

(I) Treatment	(J) Treatment	Mean Difference (I-J)	Std. Error	Sig.[a]
Ontological Approach	Non-Ontological Approach	.039*	.005	.000
Non-Ontological Approach	Ontological Approach	-.039*	.005	.000

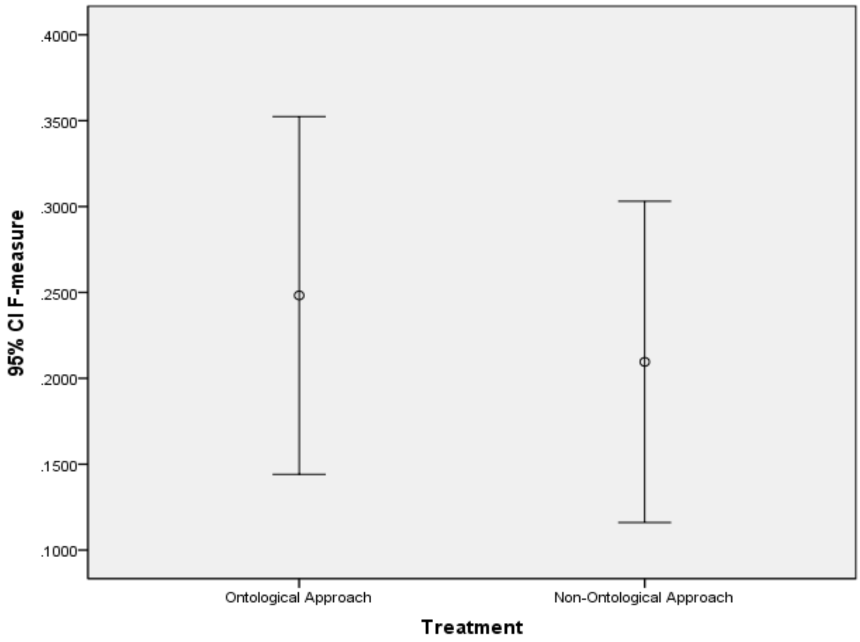


Figure 5-13: 95% CI error graph for the comparison of ontology-based and non-ontology approach

5.3.5 Response time

This section presents the performance experimental design and results of the ontology-based model compared to the non-ontology model. The aim of this experiment was to evaluate the performance of the ontology-based model in terms of its response time.

5.3.5.1 Experimental Design

To evaluate the response time of the ontology-based model compared to the non-ontology model the movieLens dataset was used without the partitions. The threshold was varied appropriately for the active user. This experiment was performed by recommending items to the active user and calculating the system's response time. The experiment was randomised and ten runs performed with the average being the final response time.

5.3.5.2 Experimental Results

Figure 5-14 shows that the ontology-based model does not perform well. The non-ontology model takes less time to respond compared to the ontology-model. Such results where expected. The ontology-based recommender system is likely to consume more time during the process of recommending because of the ontology mapping process and computing semantic similarity. In other words it was observed that the ontology-based model trades-off time for accurate recommendation.

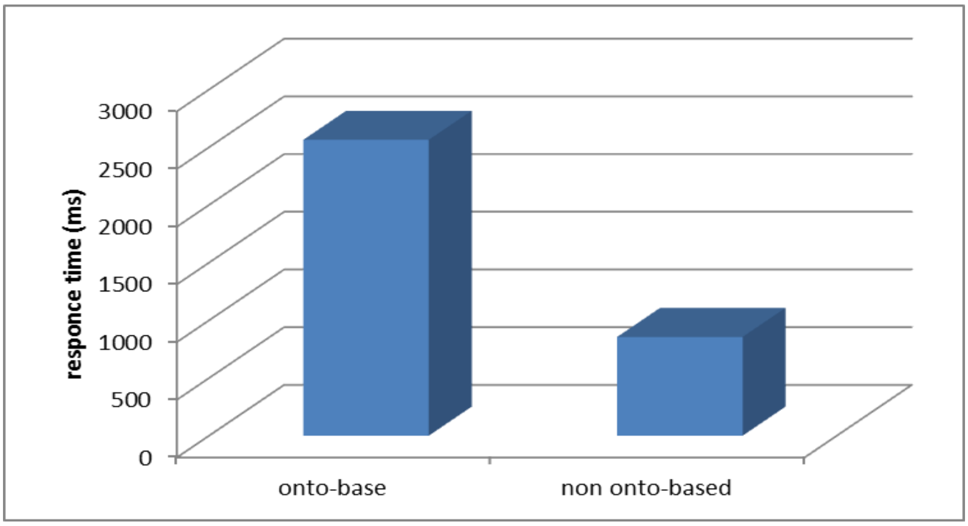


Figure 5-14 Ontology-based and non-ontology performance with response time

5.4 Discussion of Results

The current state of the art in Semantic Web and other related fields take the use of ontologies as one of the key solutions to a number of challenges dogging network-centric computing environments and as one of the key paradigm towards the future of computing. This study showed one way in which ontologies can be used to address an open challenge in personalisation and interoperability in a network-centric computing environment. This study reports the effectiveness of using ontologies as knowledge representation mechanism, as well as to facilitate semantic interoperability in the modelling components of a computing system. To demonstrate its effectiveness, experiments were designed and conducted comparing the proposed O-GAF to the classical GAF.

The O-GAF is an extension of the GAF, where the O-GAF makes use of semantic interaction in the modelling components. The results show a significant improvement in the O-GAF-based recommender system. When recommending using the proposed O-GAF it was observed that the recommendation accuracy was improved compared to the GAF. Therefore, there are three major findings from this study:

1. User modelling is an important part of any personalised adaptive system. The modelling components are a way of representing the user's interest (user model), user's goal (goal model), and user's current needs (context model) and to represent the application domain the user is using (domain model). In order for the system to make better use of the information within these components, the data models for the components were represented as ontologies. These ontological representations allow the system to meaningfully process and reason about the data in the modelling components;
2. Based on literature, the syntax interaction of the modeling component in the GAF has argued to hinder the adaptation process as well as personalisation (Knutov *et al*., 2010). In the process of providing personalisation it is important that the interacting components can successfully interoperate with each other, in order achieve complete personalisation. Therefore, an alternative interaction was proposed. Semantic interaction through ontology mapping allowed the components to meaningfully and unambiguously share knowledge hence improving personalisation;

3. The results of the study showed that the proposed O-GAF does not perform well in terms of response time. The use of ontology mapping for semantic interoperability consumes more computational time hence trading off response time for accurate personalisation.

Literature states that ontologies are being used as a solution in many research areas and more especially as a solution to interoperability problems (Stoimenov *et al.*, 2005; Heflin and Hendler, 2000; Wong *et al.*, 2008).

5.5 Discussion of the findings in the context of GUISET

Since the GUISET portal act as a single point of access to different resources provided by various SMMEs, this makes it a big challenge for the clients to find relevant resources. The findings of this study have shown that having ontological representations of the GAF modelling components and semantic interoperability within these components have a positive effect on personalisation. In the GUISET portal, these components represent SMME client (portal users) background information (such as user profile, goal, context and domain). Such knowledge is used in the process of finding relevant sources to recommend to the client (i.e. to personalise the user's information space) given his or her task at hand.

The results obtained showed that the O-GAF has better recall than the GAF. However, the GUISET portal users are more likely to access the portal through mobile devices. The small display sizes on mobile devices require that the portals recommend a small set of highly relevant resources. A higher value of recall implies the developed solution recommended a higher percentage of relevant items. However, high recall may come at the cost of having a lot of irrelevant items among the recommendations. This will not be good for any user, worse still for mobile users. To check whether the higher recall is not coming along with a lot of irrelevant recommendations precision is assessed as a function of recall. Precision measures how much proportion of the recommended items is relevant. It was discovered that the O-GAF gives better precision compared to the GAF. This implies that the O-GAF is better suited for mobile users than the GAF. .

The better recall and precision come at the cost of computational intensiveness. This is not envisaged to be a challenge, because GUISET designed to have most computations done in the infrastructure GUISET provides. End user client devices are expected thin clients.

5.6 Summary

In this chapter the O-GAF was applied in a movie recommender prototype. Experiments were conducted comparing the O-GAF prototype with the classical GAF recommender system. Conclusions were drawn based on the results obtained, the results of the recommender systems showed that the O-GAF approach performed better than the GAF. The following chapter 6 concludes the study and suggests future work.

– CHAPTER SIX –
Summary, Conclusion and Future Work

6.0 Introduction

The chapter concludes this research study by first presenting the summary of this work; summarising on how the research questions were answered and how the objectives were met. The chapter further discusses the limitations and future research directions of this work.

The rest of the chapter is organised as follows; Section 6.1 presents a summary of the research conducted in this study and the highlights of study. Section 6.2 presents the limitations of the study and the future research directions.

6.1 Summary

Retrieving personalised information from the ever increasing information of the Web has been a challenge since the explosion of the Internet. The idea is to retrieve specific information tailored towards the needs of the user, in the way that the user is shielded from information overload and confusion. In order for a system to adapt to the user's needs, the system needs to know certain information about the user, which can be used in the process of personalising the user information space. Our investigation shows that various approaches have been proposed and implemented towards realising adaptation and personalisation. These currently employed approaches such as; collaborative filtering and content-based filtering are found to give the best results so far but yet they have their limitations. The limitations in these approaches includes: (i) complexity and being computerised oracles, (ii) the involvement of the user in the personalisation process is limited, (iii) syntactic interaction within the essential modelling components.

This research addressed these challenges with the aim of enhancing personalisation performance. The Generic Adaptation Framework (GAF,) discussed in Chapter three was adopted. This framework offers a reference architecture for adaptive web-based systems. The advantages of using GAF were also discussed in chapter three. Essential and optional components were identified and the criteria to differentiate between these elements, describing their functionality and interaction. This work is more concern with the overlay modeling components provided in the GAF, which are: Goal Model, User Model, Domain Model and Context Model. Ontology-based knowledge representation of these modeling components was proposed and implemented with the following objectives:

a) Allowing the system to reason and infer new knowledge about its users when making recommendations.

b) Ontologies provide a platform for semantic interoperability of the modeling components through ontology mapping.

A personalised recommender system prototype based on ontology-based approach, O-GAF, was implemented and to test its performance it was compared to the classical GAF. The results showed that the O-GAF has better recommendation accuracy compared to the classical GAF. The results showed that ontology support improves the system performance in the process of personalisation. Therefore, based on these results our ontology-based approach can be employed into GUISET.

The objectives of this study as stated in chapter one have been perceptibly met in the following ways:

i. The literature review that was conducted reported on the state of the art in the research area of adaptive personalisation and assisted this study in identifying the Generic Adaptation Framework (GAF) which was further studied and extended. Based on this framework the essential modelling components were identified. The overlay components in the GAF were modelled with ontologies for knowledge representations enabling knowledge reasoning.
ii. Semantic interoperability was proposed in the modelling components to improve the way they interact with each other. For semantic interoperability to be realised the ontological representations of the components where mapped.
iii. The proposed model was designed based on the criteria identified in objective ii.
iv. An O-GAF recommender system prototype was implemented and experiments were designed and conducted with the aim of evaluating the recommender system compared with the classical GAF. The results obtained showed evidence of improvement in personalisation.

6.2 Limitations and Future work

Based on the adopted GAF framework there are four essential modelling components for personalisation (i.e. goal model, user model, domain model and context), but due to the limitations of the application domain and the dataset used for experimentation some components where not implemented. The four components are optional depending on the application domain being implemented, since some components may not be applicable for other domains. Therefore, the movie recommender system limited us to using the user model and domain model components. For future work, the ontology-based approach should be tested on other application domain that may require all the components to be implemented.

The results obtained from the experiments showed that the use O-GAF improved personalisation performance. However, this was achieved at a cost of longer response time. There is a need for future work to investigate this matter, which aims to find the cost of using ontology mapping in the GAF overlay model and establishing a balance between computational cost and personalisation performance.

Future work is required to test the O-GAF in an adaptive portal application such as GUISET e-commerce portal to observe it performance under real life conditions.

Bibliography

Abel, D.J, Ooi, B.C., Tan, K.L., and Tan, S.H., (1998). Towards integrated geographical information processing. Int. *Journal of Geographical Information Science* 12 (4), 353-371.

Adigun, M.O., Emuoyibofarhe, O. J., and Migiro, S.O., (2006). Challenges to Access and Opportunity to use SMME enabling Technologies in Africa, *1st All Africa Technology Diffusion Conference*, June 12-14, 2006, Johannesburg South Africa.

Aerts, A., Berden B., de Lange, B., Rousseau B., Santic, T., Smits, D., Stash, N., (2003). “Aha! the adaptive hypermedia architecture”, *Proceedings of the ACM Hypertext Conference,* Nottingham, UK, 2003, pp. 81-84.

Alexandros Paramythis. 2009. Adaptive Systems: Development, Evaluation and Evolution. PhD Thesis. pp. 212. Johannes Kepler University Linz.

Andreas, N., Stefan, P., and Rene, P. (2008). Social Recommendation Adaptation in Web Portals, In: Proceedings of AH, July

Antoniou, G. and Van Harmelen, F., (2004). OWL Web Ontology Language, 157-160. *In Handbook on Ontologies in Information*

Aroyo, L., Dolog, P., Houben, G.-J., Kravcik, M., Naeve, A., Nilsson, M., and Wild, F., (2006) Interoperabilityin Personalized Adaptive Learning. *In EducationalTechnology& Society*, 9 (2), pages 4–18,.

Baader, F. (1999). Logic-based knowledge representation. *In Artificial Intelligence Today, Recent Trends and Developments, M. Wooldridge and M. Veloso, Eds. Lecture Notes in Computer Science*, vol. 1600. Springer-Verlag, Berlin, Germany, 13--41.

Baader, F., Calvanese, D., McGuinness, D. L., Nardi, D., and Patel-Schneider, P. F. (Eds.). (2003). The description logic handbook: theory, implementation, and applications. *New York, NY, USA: Cambridge University Press.*

Baader, F., Calvanese, D., McGuinness, D., Nardi, D., and Patel-Schneider, P., The Description Logic Handbook (*2003*). *Cambridge University Press.*

Bal´ık, M. and Jel´ınek, I. (2007). “Towards semantic web-based adaptive hypermedia model” 5th Annual European Semantic Web Conference (ESWC), Tenerife, Spain.

Balabanovic, M. and Shoham, Y., (1997). Content-based, collaborative recommendation. Communications of the ACM, 40(3):66.72.

Beckett, D. (2004). RDF/XML Syntax Specification. W3C Recommendation, [online]: http://www.w3.org/TR/rdf-syntax-grammar/

Berners-Lee, T. Axioms of Web Architecture: Metadata Architecture. W3C Web site 1997. [online]: 2http://www.w3.org/DesignIssues/Metadata.html. Last accessed 2011.

Berners-Lee, T. Semantic Web - XML2000. W3C Web site 2000. URL: http://www.w3.org/2000/Talks/1206-xml2k-tbl/slide10-0.html. Last accessed 2012.

Bhowmick, P. K., Sudheshna, S. and Anupam, B., (2010). "Ontology based user modelling for personalized information access",*Int'l Journal of Computer Science and Applications*,pp.1-22

Bosch, M., (2006). Ontologies, Different Reasoning Strategies, Different Logics, Different Kinds of Knowledge Representation: *Working Together. Knowledge Organization, ErgonVerlag*. 33(3) 153-159, 2006. ISSN 0943-7444

Bouza, A.: "MO – the Movie Ontology". MO – the Movie Ontology. 2010. [online] movieontology.org

Bouzeghoub, M., FariasLóscio, B., Kedad, Z., and Salgado, A. C. (2003) "Managing the evolution of mappings". *In Proc. of the 11th. Int. Conf. on Cooperative Information Systems* (CoopIS'03), Catania, Italy. 22-37

Brailsford, T.J., Stewart, C.D., Zakaria, M.R., and Moore, A., (2002). Autonavigation, links, and narrative in an adaptive Web-based integrated learning environment. *In: Proc. of The 11th International World Wide Web Conference*.

Brusilovsky, P. (1996). Methods and Techniques of Adaptive Hypermedia. *User Modeling and User Adapted Interaction*, 6(1-2):87–129.

Brusilovsky, P. and Mill`an, E., (2007). User Models for Adaptive Hypermedia and Adaptive. *The Adaptive Web: Methods and Strategies of Web Personalization*, volume 4321, pages 3–53. Springer, Berlin, Heidelberg, 2007.

Brusilovsky, P., (2001). Adaptive Hypermedia. User Modeling and User-Adapted Interaction, 11(1-2):87{110, 2001}.

Burke, R. (1999). Integrating knowledge-based and collaborative-filtering recommender systems. In *Proceedings of the AAAI Workshop on AI and Electronic Commerce.*

Burke, R., (2002). Hybrid recommender systems: Survey and experiments, *in User Modeling and User-Adapted Interaction,* no. 12, pp. 331–370, 2002.

Cantador, I., Bellogín, A., and Castells, P. (2008). A Multilayer Ontology-based Hybrid Recommendation Model. AI Communications, 21(2-3), 203-210.

Cawsey, A., Floriana, G., and C´ecile, P., (2007). Adaptive Information for Consumers of Healthcare. *In The Adaptive Web, Methods and Strategies of Web Personalization*, pages 465–484, 2007.

Chandrasekaran, B. Josephson J. R. and Benjamins, R. (1999). What are ontologies, and why do we need them?, *IEEE Intelligent Systems*, Vol.14, No.1, 20-26.

Choi, N., Son, I.Y and Han, H., (2006) A survey on ontology mapping *ACM SIGMOD Record*, 35 (3) (2006), pp. 34–41.

Claypool, M., Gokhale, A., Miranda, T., Murnikov, P. and Sartin, M., (1999). Combining content-based and collaborative filters in an online newspaper, *Proceedings of ACM SIGIRWorkshop on Recommender Systems*, August 1999.

Codina, B. V. and Ceccaron, L., (2009). Design, Development and Deployment of an Interlligent, Personalized Recommender Sysyem, *PhD Msc Thesis, Technical University of Catalonia*, 2008-2009.

Connolly, D., van Harmelen, F., Horrocks, I., McGuinness, D.L., Patel-Schneider F.P., and Stein, L.A., (2001). DAML+OIL reference description. *W3C note 18 December 2001. [online]:http://www.w3.org/TR/2001/NOTEdaml+ oil-reference*-20011218.

Cremonesi, P. Turrin, R. (2009). Analysis of Cold-start Recommendations. *in IPTV Systems, ACM*. pp. 233-236.

Cremonesi, P., Turrin, R., Lentini, E., Matteucci, M., (2008). An Evaluation Methodology for Collaborative Recommender Systems, Proceedings of the 2008 International Conference on Automated solutions for Cross Media Content and Multi-channel Distribution, p.224-231,

Davis, R. H. S. and Szolovits, P., (1993). What is knowledge representation? *.AI Magazine,* Vol. 14, No. 1, pp. 17-33,

De Bra, P., Aroyo, L., and Chepegin, V., (2004). The Next Big Thing: Adaptive Web-Based Systems, *Journal of Digital Information*, vol. 5 issue 1 article 247.

deGemmis, M., Iaquinta, M., Lops P., Musto, C., Narducci, F., and Semeraro G., (2009). Preference learning in recommender systems. *In: ECML/PKDD-09 Workshop on Preference Learning (PL 09)*

Doan, A., Madhaven, J., Domingos, P. and Halevy, A. (2004). Ontology matching: A machine learning approach, *in S. Staab& R. Studer, eds,* '*Handbook on Ontologies in Information Systems', Springer-Verlag*, pp. 397–416.

Dou, D., McDermott, D. V., & Qi, P. (2005). Ontology translation on the Semantic Web. *Journal of Data Semantics*, 2, 35–57.

Ehrig, M. and Staab, S. (2004). QOM - quick ontology mapping, in F. van Harmelen, S. McIlraith& D. Plexousakis, eds, '*Proceedings of the Third International SemanticWeb Conference (ISWC2004)', LNCS, Springer, Hi- roshima, Japan*, pp. 683–696.

Encarnação, L.M., (1997). Multi-level user support through adaptive hypermedia: A highly application-independent help component. *In: Moore, J., Edmonds, E., Puerta, A. (eds.) Proc. of 1997 International Conference on Intelligent User Interfaces. ACM* (1997) 187-194.

Euzenat, J. and Shvaiko, P., (2007) Ontology matching. *Springer journal*.

Fedor, B., Birgitta, K., Andreas, N., and Martin, W., (2009). A Hybrid Approach to Identifying User Interests in Web Portal. *Proceedings of the 9th International Conference on Innovative Internet Community Systems*.

Ferreirós, J., (2001); The Road to Modern Logic - An Interpretation, *Bulletin of Symbolic Logic,* Volume 7, Issue 4, 2001, pp. 441–484

Fink, J., &Kobsa, A. (2000). A Review and Analysis of Commercial User Modeling Servers for Personalization on the World Wide Web. *International Journal on User Modeling and User-Adapted Interaction*, 10(2-3), 209-249.

Frasincar, F. and Houben, G. J. (2002). Hypermedia presentation adaptation on the semantic web. *In Adaptive Hypermedia and Adaptive Web-Based Systems (AH 2002)*, volume 2347 of Lecture Notes in Computer Science, pages 133-142. Springer.

Godugula, S., (2008). Survey of Ontology Mapping Techniques.

Goy, A., Liliana, A., and Giovanna, P., (2007). Personalization in E-Commerce Applications. *In The Adaptive Web, Methods and Strategies of Web Personalization*, pages 485–520,

Grosof, B. N., Horrocks, I., Volz, R, and Decker, S., (2003). Description Logic Programs: Combining Logic Programs with Description Logic". *In: Proc. 12th Intl. Conf. on the World Wide Web (WWW-2003), ACM Press.*

Gruber, T. R. (1995). “Toward Principles for the Design of Ontologies Used for Knowledge Sharing”, *International Journal of Human and Computer Studies*, 43(5/6), 907–928.

Guarino, N., (1998). Formal Ontology in Information Systems, *Amsterdam, Berlin, Oxford: IOS Press. Tokyo, Washington, DC: IOS Press (Frontiers in Artificial Intelligence and Applications).*

Heflin, J. and Hendler J. (2000). Semantic interoperability on the web. In *Extreme Markup Languages*.

Heitmann, B., Hayes C., (2010). Using linked data to build open, collaborative recommender systems. *In: AAAI Spring Symposium: Linked Data Meets Artificial Intelligence'*.

Henze, N. (2001). Open Adaptive Hypermedia: An approach to adaptive information presentation on the Web. *In C. Stephanidis (Ed.), Universal Access in HCI: Towards an Information Society for All – Volume 3 of the Proceedings of the 9th International Conference on Human-Computer Interaction (HCI International 2001), New Orleans, Louisiana, USA*, 5-10 August (pp. 818-821).

Henze, N., and Herrlich, M. (2004). The personal reader: A framework for enabling personalization services on the semantic web. *In Proceedings of the Twelfth GI Workshop on Adaptation and User Modeling in Interactive Systems (ABIS 04).*

Herlocker, J. L., Borchers, J. A. and Riedl J., (1999). An Algorithm framework for performing collaborative filtering, *Proceedings of SIGIR '99 Conference on Research and Development in Information Retrieval*, ACM Press, NewYork, NY, 1999, pp. 230-237.

Horrocks, I., Patel-Schneider, P. F., and van Harmelen, F. (2003). From SHIQ and RDF to OWL: The making of a web ontology language. *Journal of Web Semantics*, 1(1), 7-26.

IF-Map from The University of Southampton,[online]: http://www.aktors.org/technologies /ifmap/, accessed on July 27, 2011.

Isakowitz, T., Bieber, M., and Vitali, F. (1998). Web Information Systems. *Communications of the ACM*, 41(7), 78-80.

Jin, X., Zhou, Y., and Mobasher, B. (2005). Task-Oriented Web User Modeling for Recommendation. In: L. Ardissono, P. Brna& A. Mitrovic (eds.), *Proceedings of the 10th International Conference on User Modeling (UM'2005)* (pp. 109-118), Edinburgh, UK: Berlin: Springer

Joachims, T., Freitag, D., Mitchel T., (1997). A Tourguide for the World Wide Web. *In Proceedings of the International Joint Conference in Al. Los Angeles*, CA.

Kalfoglou, Y., and Schorlemmer, M., (2003). Ontology mapping: the state of the art, *The Knowledge Engineering Review, 18(1), pp. 1–31.*

Kass, R., andFinin, T. (1988). Modeling the user in natural language systems. *Computational Linguisticfs*, 14(3), 5-22.

Kay, J., and Kummerfeld, R. J. (1994). An individualised course for the C programming language. *In: Proceedings of the Second World Wide Web Conference (WWW'94), Chicago, Illinois, USA.*

Klyne, G. and Carroll, J., (2004). RDF Concepts and Abstract Syntax. http://www.w3.org/TR/rdf-primer/

Knutov, E. (2009). GAF: Generic Adaptation Framework. *Hypertext 2009 SRC poster*, 2009.

Knutov, E., Bra, D. P., and Pechenizkiy, M. (2010). Generic adaptation process, Proceedings of the *International Workshop on Architectures and Building Blocks of Web-Based User-Adaptive Systems.* (CEUR Workshop Proceedings, 609, pp. 13-24).

Knutov, E., Bra, P. D., and Pechenizkiy M., (2011). Generic adaptation framework: a process-oriented perspective. *J. Digit. Inf., 12(1)*, 2011.

Kobsa, A. (2001). Generic User Modeling Systems. *User Modeling and User-Adapted Interaction*, 11. 49-63.

Konstan, J., Miller, B., Maltz, D., Herlocker, J., Gordon L., and Riedl J., (1997). GroupLens: applying collaborative filtering to Usenet news. *Communications of the ACM*, 40(3):77.87, 1997.

Kowalski, R.A., (1986). "Logic for Problem Solving". *North Holland Publishing Co.*, 1986

Kramár, T. and Bieliková, M. (2012). Dynamically selecting an appropriate context type for personalisation. Proceedings of the sixth ACM conference on Recommender systems, 321-324.

Kutz, O., Lutz, C., Wolter, F. and Zakharyaschev, M. (2004). Connections of Abstract Description Systems. *Artificial Intelligence* 156, 1–73.

Lang, K., (1995). NewsWeeder: learning to filter netnews. *In Proceedings of the 12th International Conference on Machine Learning (ICML)*, pages 331.339. Morgan Kaufmann, 1995.

Lassila, O., and Swick, R. (1999). Resource Description Framework (RDF) Model and Syntax Specification. W3C Recommendation, From: http://www.w3.org/TR/1999/REC-rdfsyntax-19990222.

Manola, F., and Miller, E. (2004). RDF Primer. W3C Recommendation, [online]: http://www.w3.org/TR/rdf-primer/

Mathé, N. and Chen J., (1996). User-centered indexing for adaptive information access. *User Models and User Adapted Interaction 6.*

McNee, S. M., Kapoor, N., and Konstan, J. A. (2006). Don't Look Stupid: Avoiding Pitfalls when Recommending Research Papers. *In: Proceedings of the International Conference on Computer Supported Cooperative Work (CSCW'2006)* (pp. 171-180), Banff, Alberta, Canada: ACM Press.

Middleton, S. E., Alani, H., Shadbolt, N. R., and Roure, D. C. D. (2002). Exploiting synergy between ontologies and recommender systems. *Proceedings of Semantic Web Workshop 2002 At the Eleventh International World Wide Web Conference*, 41–50.

Mika, P. Social Networks and the Semantic Web: (2005) The Next Challenge. IEEE Intelligent Systems 20.

Mobasher, B., Cooley, R., and Srivatsava, J., (2000). Automatic Personalization Based on Web Usage Mining. *ACM., 43*:142-151.

Mobasher, B., Jin, X., and Zhou, Y., (2004). Semantically Enhanced Collaborative Filtering on the Web'. In B. Berendt, et al. (eds.): *Web Mining: From Web to Semantic Web. LNAI* Volume 3209, Springer. (2004).

Moore, P. and Hu, B. (2006), A context framework for entity identification for the personalization of learning in pedagogic systems. *Proceedings of the 10th International Conference on Computer Supported Co-operative Work in Design* (CSCWD 2006), Nanjing, China, 2006; 1381–1386.

Namyoun, Choi, Il-Yeol Song, Hyoil, Han, (2006) A survey on ontology mapping, *ACM SIGMOD Record*, v.35 n.3, p.34-41, September 2006.

Noy, N. and Musen, M., (2003) "*PROMPT:* Algorithm and Tool for Automated Ontology Merging and Alignment. *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, 2000.

Noy, N. F. and Musen, M. A. (2000). Anchor-PROMPT: Using non-local context for semantic matching, *in 'Proceedings of the Workshop on Ontologies and Information Sharing at the Seventeenth International Joint Conference on Artificial Intelligence (IJCAI-2001)', Seattle, WA, USA.*

Owen, S., Anil, R., Dunning, T., and Friedman E., (2011). Mahout in Action. *Manning Publications Co*.

Pinto, S. H., Gomez-Perez, A., Martins J. P., (1999). Some Issues on Ontology Integration, *In Proc. Of IJCAI99's Workshop on Ontologies and Problem Solving Methods*: Lessons Learned and Future Trends,

Razmerita, L., Angehrn, A., Maedche A., (2003). Ontology based user modeling for Knowledge Management Systems, *Proceeding of the User Modeling Conference, Pittsburgh, USA, Springer Verlag*, pp. 213-217,

Ritze, D., V¨olker, J., Meilicke, C., and ˇSv´ab-Zamazal, O. (2010). Linguistic analysis for complex ontology matching. Ontology Matching 1.

Sah, M., Hall, W., (2007). Building and Managing Personalized Semantic Portals. *In: WWW'07 (ACM)* 1227-1228

Schafer, B.J., Konstan, J., Riedi, J., (1999). Recommender systems in e-commerce, *Proceedings of the 1st ACM conference on Electronic commerce*, p.158-166, November 03-05, Denver, Colorado, United States of America.

Shakhnarovish, Darrell, and Indyk, (2005). Nearest-Neighbour Methods in Learning and Vision. *The MIT Press*.

Shvaiko, P. and Euzenat, J., (2011). 'Ontology matching: State of the art and future challenges', *IEEE Trans. Knowl. Data Eng., 99*,

Sosnovsky, S. and Dicheva, D. (2010). Ontological technologies for user modelling. *Int. J. Metadata Semant. Ontologies, 5*(1):32–71,

Sosnovsky, S., (2008). Ontological technologies for user modelling. *Working paper.*

Sowa, J.F., (2000). "Knowledge Representation. Logical, philosophical and computational foundations". *Brooks Cole Publishing Co.,*

Stoimenov L., Stanimirović A., and Đorđević-Kajan S., (2005). Semantic Interoperability using multiple ontologies, *Proceedings printed as book, Eds. Fred Toppen, Marco Painho, AGILE 2005, 8th AGILE Conference on GIScience, Estoril,* Portugal, 26-28.5.2005. , pp.261-270.

Stuart, E., Middleton, D.C., De Roure, Nigel, R., and Shadbolt, (2001). Capturing knowledge of user preferences: ontologies in recommender systems, *Proceedings of the 1st international conference on Knowledge capture*, October 22 -23, 2001, Victoria, British Columbia, Canada.

Studer,R., Grimm, S., Abecker A., (2007). Semantic Web Services - Concepts, Technologies and Applications, *Springer*.

Terveen, L. and Hill, W., (2001). Beyond Recommender Systems: Helping People Help Each Other. *HCI In The New Millenium, Carroll, J. ed. Addison-Wesley*, 2001.

Thanh, T., Philipp, C., and Anupriya, A., (2006). Rules for an Ontology-based Approach to Adaptation. *In Proceedings of the 1st International Workshop on Semantic Media Adaptation* *and Personalization. Athen, Greece*.

Tomas, K., and Ivan, J., (2007). Ontology-based Approach to Adaptation,

Towle, B. and Quinn, C., (2000). 'Knowledge Based Recommender Systems Using Explicit User Models'. *In Knowledge-Based Electronic Markets, Papers from the AAAI Workshop*, AAAI Technical Report WS-00-04. pp. 74-77. Menlo Park, CA: AAAI Press.

Uschold, M. (2000). Creating, integrating, and maintaining local and global ontologies, *in 'Proceedings of the First Workshop on Ontology Learning (OL-2000) in conjunction with the 14th European Conference on Artificial Intelligence (ECAI-2000)*', Berlin, Germany.

Vallecillo, A., Hern´andez J., and Troya, J.M., (2000). Component interoperability. *Technical Report ITI200037,* Universidad de M´alaga, Spain,

VanLehn, K., (1988). Student models. In: Polson, M.C., Richardson, J.J. (eds.): Foundations of Intelligent Tutoring Systems. *Lawrence Erlbaum Associates, Hillsdale* (1988) 55-78.

Veltman, K.H., (2001). Syntactic and Semantic Interoperability: New Approaches to Knowledge and the Semantic Web. *The New Review of Information Networking* 2001; 7: 159-84.

Visser, P. and Cui, Z. (1998). On accepting heterogeneous ontologies in distributed architectures. *In Proceedings of the ECAI'98 workshop on Applications of Ontologies and Problem-solving methods,* p. 112-119, UK, Brighton.

Vrieze, P., Fundamentals of Adaptive Personalization, PhD Thesis, Radboud University Nijmegen (2006).

W3C Recommendations, [online] http://www.w3.org/TR/2004/REC-owl-features-20040210/#s1.3, accessed on July 27, 2009.

W3C Recommendations, [online] http://www.w3.org/TR/rdf-schema/#ref-rdf-concepts, accessed on July 27, 2009.

Wahlster, W., and Kobsa, A. (1989). User Models in Dialog Systems. In A. Kobsa& W. Wahlster(eds.), User Models in Dialog Systems. *Heidelberg: Springer*.

Wong, A., Ray, P., Parameswaran, N., Strassner, J., (2008). Ontology mapping for the interoperability problem in network management, *Journal on Selected Areas in Communications*, Oct, Vol. 23, Issue 10, page(s): 2058- 2068.

Wu, Z. and Chen, H., (2008). Semantic Grid: Model, Methodology, and Applications, Jointly published with Zhejiang University Press, Series: *Advanced Topics in Science and Technology in China.*

Xiao Dong, W., Xiaobo, Y., and Rob A., (2006). Top ten question to design a successful grid portal, *CCLRC e-Science Center*.

Xiaobo, Y., and Rob, A., (2006). Bringing AJAX to Grid Portals. *CCLRC e-Science Center*

Yang, G., Kifer, M., and Zhao, C., (2003). FLORA-2: A rule-based knowledge representation and inference infrastructure for the semantic web. *In Proc. of the Second Int. Conf. on Ontologies, Databases and Applications of Semantics (ODBASE),.*

Yehuda, K., (2008). Factorization meets the neighborhood: a multifaceted collaborative filtering model, *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, August 24-27, Las Vegas, Nevada, USA.

Zuo, L. A. (2006). Semantic and Agent-Based Approach to Support Information Retrieval, *Interoperability and Multi-Lateral Viepoints for Heterogeneous Environmental Databases*. PhD Thesis, The Department of Electronic Engineering, Queen Mary, University of London.